



UNIVERSITAT DE  
BARCELONA

## Estudio acústico de las vocales del español a partir de grabaciones de habla infantil

María Manjavacas Ramírez Cruz



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 3.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 3.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 3.0. Spain License.**

**Tesis doctoral**

*Estudio acústico de las vocales del español a partir de grabaciones de habla infantil*

María Manjavacas Ramírez Cruz

Programa de doctorado de Ciencia Cognitiva y Lenguaje

Universidad de Barcelona

Director de tesis

Dr. Eugenio Martínez Celdrán

Catedrático de Lingüística

Universidad de Barcelona

2017



**Tesis doctoral**

*Estudio acústico de las vocales del español a partir de grabaciones de habla infantil*

## **AGRADECIMIENTOS**

Mi gratitud en primer lugar al profesor Eugenio Martínez Celdrán, director de esta tesis, por haber aceptado la dirección de la misma constituyendo su motivación principal el apoyo a la investigación y la apuesta por los jóvenes estudiantes. Me gustaría darle las gracias de manera muy especial por su orientación, sus relevantes aportaciones y sus interesantísimas sugerencias durante el desarrollo de esta investigación así como por su interés en la finalización y defensa de la misma aun cuando las circunstancias nos eran adversas.

Mis agradecimientos más sinceros a la profesora Juana Gil por haberme escuchado y aconsejado, por su presencia incondicional, su apoyo constante y por la total confianza que ha depositado en mí a lo largo de mi trayectoria investigadora alentándome en las distintas fases de la investigación incluso cuando los ánimos flaqueaban.

Asimismo quiero expresar mi gratitud al profesor Joaquim Llisterri no solo por la ayuda, ánimo y cariño brindados sino por haberme tendido desde el primer momento su mano amiga.

Quizás todo comenzó en primero de carrera con la profesora Alicia Mellado cuyas clases despertaron en mí el interés por la Fonética.

Obligados son mis agradecimientos al profesor Sergio Cuesta no tanto por sus consejos informáticos sino por haberme infundido el ánimo y la serenidad necesarios para haber podido llevar a término esta investigación.

No puedo olvidar a los profesores José María Lorenzo y Ángel Rodríguez que de forma altruista y desinteresada me han ayudado con los análisis estadísticos y matemáticos.

A los miembros del tribunal, por la lectura y tiempo que han dedicado a esta investigación y a los miembros de la Comisión investigadora de la Universidad de Barcelona por haber aceptado el proyecto y la tesis doctoral.

Inevitable es mi agradecimiento a los niños y a sus padres por haberse prestado de manera desinteresada a ayudarme en este proyecto y permitir así que la ciencia avance.

Por último, expreso mi más profundo agradecimiento a toda mi familia y en especial a mis padres, sin cuyo esfuerzo no hubiera podido comenzar esta tesis y sin cuyo cariño no hubiera podido acabarla.

## RESUMEN

A pesar de que al estar presentes en el habla temprana las vocales son hitos importantes en el desarrollo del lenguaje (Vorperian y Kent [2007]), la mayor parte de los trabajos que han abordado el estudio fonético de las emisiones infantiles se ha centrado en la descripción de los sonidos consonánticos debido, entre otras razones, a falsas creencias como la de que las vocales rara vez se articulan mal y son aprendidas con poco esfuerzo (Davis y MacNeilage [1990], Stokes y Wong [2002] y Van der Stelt et al. [2005]).

Dentro la reducida bibliografía que se ha acercado al estudio de los sonidos vocálicos en edades tempranas, son muy pocos los trabajos que han emprendido la tarea de registrar y transcribir fonéticamente estos segmentos a lo largo de los tres primeros años de vida (periodo de tiempo en el que se producen los mayores cambios anatómicos del tracto vocal según Selby et al. [2000] y en el que aparecen las características fundamentales de las vocales según Templin [1957]); y más escasos aún son aquellos que han analizado acústicamente los dos primeros formantes de estos sonidos vocálicos pese a la importancia que este índice acústico tiene en el reconocimiento y diferenciación de la vocal.

Entre las causas que determinan la escasez del primer tipo de trabajos se encuentran, según Palethorpe et al. (1996), la dificultad de la obtención de los datos puesto que resulta complicado encontrar sujetos de tan corta edad a los que estudiar por un periodo de tiempo tan extenso, así como la dificultad que presenta el transcribir muestras de habla de sujetos tan jóvenes, tanto por lo complejo que resulta la identificación de las vocales como por las carencias que presentan los actuales sistemas de transcripción que están diseñados para la transcripción de las emisiones adultas (Kent y Bauer [1985], Oller [1986], Gilbert et al. [1997], etc.). En cuanto a las investigaciones que han analizado la estructura formántica de las vocales tempranas, y frente a aquellas dedicadas al análisis de otros parámetros acústicos del habla infantil como la duración, la  $f_0$ , o la intensidad, la escasez de publicaciones reside en la complejidad del examen acústico que viene dada, según Kent (1976), por las peculiaridades que presenta el habla de los bebés y de los niños pequeños como la tendencia de estos sujetos a emitir sonidos con voz no modal, la nasalización de las vocales debido a la configuración particular del tracto vocal y la alta frecuencia fundamental de las emisiones.

La exigüidad de este tipo de estudios en la literatura internacional y la ausencia de trabajos que ahonden en estas cuestiones en el campo del español (al menos durante un periodo de estudio tan extenso), han constituido la motivación principal que nos ha llevado a emprender esta tesis doctoral en la que los objetivos principales son el análisis acústico de los dos primeros

formantes de las emisiones vocálicas aisladas de dos sujetos hablantes de castellano entre los 0;4 y los 3;0 años de vida, y la transcripción fonética de dichas emisiones. Estos objetivos generales pueden concretarse en los siguientes:

- a) Clasificar las vocales extraídas según las categorías vocálicas del español mediante un test de percepción.
- b) Hallar qué vocales emiten los bebés y los niños pequeños con mayor frecuencia a lo largo del periodo de estudio.
- c) Determinar los valores de los dos primeros formantes de los sonidos vocálicos aislados extraídos de las emisiones de un niño y una niña hablantes de español en los tres primeros años de vida.
- d) Comparar los valores medios de los dos primeros formantes de las vocales del español del habla infantil con los valores dados para las vocales de hablantes adultos de esta lengua.
- e) Observar si se produce una expansión o disminución del espacio vocálico durante el periodo de observación a partir del cálculo del área vocálica con ayuda de los valores formánticos.
- f) Observar si se produce una reducción o ampliación de los valores de las dos primeras resonancias formánticas a lo largo del tiempo.
- g) Averiguar si existen diferencias en el número de emisiones y en la cualidad vocálica de estas entre ambos sujetos.

Las hipótesis de partida son las ya comprobadas en otros estudios que se han centrado en la transcripción y el análisis acústico de los segmentos vocálicos de las emisiones de los niños que adquieren una lengua distinta a la española, y son las que nosotros pretendemos corroborar con este trabajo:

1. Los valores formánticos de los bebés superan con creces los valores de los formantes de los adultos (Irwin y Chen [1946], Peterson y Barney [1952], Fant [1960], Eguchi y Hirsh [1969], Boysson-Bardies et al. [1980], Stoel-Gammon y Herrington [1990], Kuhl y Meltzoff [1996], Ménard et al. [2004], Chen y Kent [2010], etc.).
2. Las primeras vocales en aparecer y las vocales más frecuentes en los primeros meses de vida, y al menos durante el primer año, son las vocales centrales medias y bajas y las

vocales anteriores (Irwin [1948], Lindblom y Sundberg [1969], Boysson-Bardies et al. [1980], Van der Stelt et al. [2006], Kern y Davis [2009], etc.).

3. Algunos sonidos vocálicos pueden articularse al poco de nacer porque el bebé posee las herramientas necesarias para ello, pero otros como [u] necesitan de una maduración anatómica y de una mejora de las habilidades de control motor para poder ser articulados (Kent [1976], Buhr [1980], Vorperian et al. [2005], etc.).
4. En los primeros meses de vida se produce una superposición de las vocales en el espacio vocálico que irá desapareciendo según aumente la edad del niño (Buhr [1980], Kent y Murray [1982], Kuhl y Meltzoff [1996], Rvachew et al. [2006], etc.).
5. A pesar de la inversa relación establecida por Fant (1960) entre la longitud del tracto vocal y las frecuencias de los formantes (de forma que según aumenta la longitud de aquel las frecuencias disminuyen), los cambios dramáticos que sufren las estructuras del tracto vocal durante el primer año de vida no provocan que los valores del F1 y del F2 descendan de manera significativa (Buhr [1980], Kent y Murray [1982], Gilbert et al. [1997], Chen [2006], etc.).
6. Aunque los niños que adquieren la misma lengua meta tienen inventarios vocálicos parecidos, pueden encontrarse diferencias individuales incluso en las primeras etapas del desarrollo (Lieberman [1980], Kent y Bauer [1985], Davis y MacNeilage [1995], Van der Stelt et al. [2006], etc.).

En cuanto al método de trabajo, se grabaron semanalmente las emisiones espontáneas de un niño y una niña hablantes de español entre los cuatro meses y los tres años de vida, en diferentes contextos familiares. Después se escucharon las grabaciones y se seleccionaron los sonidos vocálicos que se ajustaron a criterios previamente establecidos tales como: sonidos vocálicos aislados, voz modal, etc. A continuación se midieron los dos primeros formantes y se crearon cartas de formantes que nos ayudaron a situar las señales en el espacio vocálico y nos permitieron observar la evolución de las mismas a lo largo del tiempo. Posteriormente, cinco adultos hablantes nativos de español realizaron un test de percepción en el que debían clasificar los sonidos escuchados según las categorías vocálicas de la lengua española para que nosotros pudiésemos cotejar después estos resultados con los arrojados por el análisis acústico. Finalmente, y para comprobar si los valores de los dos primeros formantes disminuyeron o aumentaron de manera significativa a lo largo del tiempo o si se amplió o se redujo el área vocálica, realizamos una serie de análisis estadísticos en los que calculamos el promedio y la varianza de los mismos.

Los resultados arrojados por los análisis anteriores fueron comprobando una a una todas las hipótesis de partida (tanto aquellas que habían sido demostradas en los trabajos de investigación en los que se analizaban acústicamente los dos primeros formantes, como aquellas que habían sido verificadas en los estudios en los que se registraba el inventario vocálico de los bebés y niños pequeños que estaban adquiriendo una lengua distinta a la lengua española).

Finalmente, las conclusiones alcanzadas tras la realización de esta tesis doctoral fueron las siguientes:

1. Los valores de los dos primeros formantes de las vocales analizadas en los bebés y niños pequeños superan con creces a los de los adultos.
2. La mayoría de los sonidos vocálicos producidos por un bebé en los primeros meses de vida son articulados en la parte central y anterior de la cavidad bucal y poseen, al menos al principio, una altura media y baja.
3. Dentro del ángulo posterior, la vocal [u] es la última vocal en aparecer y no lo hace al menos hasta que el bebé ha cumplido los seis meses de vida.
4. Las producciones vocálicas de los bebés están determinadas, al menos en parte, por la lengua ambiente en la que estos están inmersos.
5. El bebé es capaz de producir cualquier vocal del español una vez que se superan los primeros siete u ocho meses de vida.
6. Se crearon campos de dispersión alrededor de todas las vocales analizadas debido a la articulación de los niños de una misma vocal con pequeñas diferencias en los valores formánticos.
7. La variabilidad de los valores formánticos dentro de una vocal puede deberse a que el tracto vocal del bebé no está aún configurado como el del adulto y sufre cambios continuos tanto en su longitud como en su anchura y volumen.
8. Los valores de los dos primeros formantes no sufren ninguna alteración significativa durante el primer año de vida a pesar de los cambios dramáticos que experimenta el tracto vocal en este periodo temporal.
9. Existe una variedad inter-individual tanto en el número de vocales extraídas a lo largo del periodo de estudio como en su cualidad vocálica.





## ÍNDICE

Agradecimientos.....	3
Resumen .....	4
Índice.....	9
Lista de abreviaturas.....	13
Lista de tablas.....	15
Lista de gráficos .....	22
1. Introducción. ....	28
2. Estado de la cuestión. ....	46
2.1. Configuración anatómica del bebé y del niño pequeño.....	47
2.1.1. Características del tracto vocal.....	47
2.1.2. Desarrollo del control motor. ....	63
2.1.3. Control motor y modelos artificiales de reproducción de habla.....	68
2.2. Capacidades perceptivas del bebé y del niño pequeño.....	77
2.2.1. Habilidades perceptivas y evolución de estas.....	77
2.2.2. Influencia de la lengua ambiente.....	92
2.2.3. Habla dirigida al bebé.....	116
2.3. Revisión bibliográfica sobre el análisis vocálico. ....	119
2.3.1. Consideraciones iniciales. ....	119
2.3.2. Revisión de estudios de inventario vocálico.....	126
2.3.3. Revisión de estudios de análisis acústico. ....	132
3. Objetivos e hipótesis de partida.....	167
4. Caracterización articulatoria y acústica de los sonidos vocálicos del español. ....	170
5. Metodología. ....	188
5.1. Sujetos.....	188

5.1.1.	Nacimiento y entorno de la niña.....	189
5.1.2.	Nacimiento y entorno del niño. ....	189
5.1.3.	Evolución de los niños a lo largo del estudio. ....	190
5.2.	Características del estudio. ....	191
5.2.1.	Sesiones de grabación. ....	191
5.2.2.	Criterios de selección. ....	195
5.2.3.	Dificultades encontradas en la grabación. ....	200
5.3.	Análisis realizados.....	201
5.3.1.	Análisis acústicos. ....	201
5.3.2.	Test de percepción.....	204
5.3.3.	Análisis estadísticos. ....	205
6.	Resultados. ....	206
6.1.	Resultados del análisis acústico. ....	206
6.1.1.	Datos obtenidos en la niña.....	207
6.1.1.1.	Primer año de vida.....	207
6.1.1.2.	Segundo año de vida.....	215
6.1.1.3.	Tercer año de vida. ....	217
6.1.2.	Datos obtenidos en el niño. ....	219
6.1.2.1.	Primer año de vida.....	219
6.1.2.2.	Segundo año de vida.....	225
6.1.2.3.	Tercer año de vida. ....	227
6.2	Resultados del test de percepción.....	229
6.2.1.	Timbres vocálicos identificados en las emisiones de la niña. ....	230
6.2.1.1.	Primer año de vida.....	230
6.2.1.2.	Segundo año de vida.....	232
6.2.1.3.	Tercer año de vida. ....	234

6.2.2.	Timbres vocálicos identificados en las emisiones del niño.....	235
6.2.2.1.	Primer año de vida.....	235
6.2.2.2.	Segundo año de vida.....	237
6.2.2.3.	Tercer año de vida.....	239
6.3.	Comparación de los resultados obtenidos en el análisis acústico y perceptivo en la niña y el niño.....	241
6.3.1.	Comparación de los resultados obtenidos tras el análisis acústico.....	241
6.3.2.	Comparación de los resultados obtenidos tras el test de percepción.....	244
7.	Análisis estadísticos.....	248
7.1.	Análisis estadísticos realizados sobre las señales que poseen el mismo timbre vocálico en cada uno de los meses analizados en los tres primeros años de vida.....	248
7.1.1.	Evolución de la media y de la desviación estándar de los dos primeros formantes en las vocales de la niña.....	249
7.1.1.1.	Vocal central baja, [a].....	249
7.1.1.2.	Vocal anterior semialta, [e].....	257
7.1.1.3.	Vocal anterior alta, [i].....	267
7.1.1.4.	Vocal posterior semialta, [o].....	276
7.1.1.5.	Vocal posterior alta, [u].....	285
7.1.1.6.	Evolución del área vocálica a partir de los valores obtenidos en las señales de la niña.....	291
7.1.2.	Evolución de la media y de la desviación estándar de los dos primeros formantes en las vocales del niño.....	294
7.1.2.1.	Vocal central baja, [a].....	294
7.1.2.2.	Vocal anterior semialta, [e].....	299
7.1.2.3.	Vocal anterior alta, [i].....	307
7.1.2.4.	Vocal posterior semialta, [o].....	316

7.1.2.5. Vocal posterior alta, [u].	322
7.1.2.6. Evolución del área vocálica a partir de los valores obtenidos en las señales del niño.	329
7.1.3. Comparación de los resultados obtenidos en cada uno de los sujetos.	333
8. Discusión.	342
9. Conclusiones.	396
10. Epílogo.	400
Bibliografía citada.	403

## LISTA DE ABREVIATURAS

FFT: *Fast Fourier Transform.*

LPC: *Linear Predictive Coding.*

IPA/AFI: *International Phonetic Alphabet/Alfabeto Fonético Internacional.*

F1, F2, F3: *Formante 1, Formante 2, Formante 3.*

$f_0$ : *frecuencia fundamental.*

VLAM: *Variable Linear Articulatory Model.*

DIVA: *Directions Into Velocities of Articulators*

ANOVA: *Analysis Of Variance.*

ANCOVA: *Analysis of Co-Variance.*

VRISD: *Visual Reinforced Infant Speech Discrimination.*

IDS: *Infant Direct Speech.*

ADS: *Adult Direct Speech.*

VOT: *Voice Onset Time.*

SMA: *Supplementary Motor Area.*

ANS-PNS: *Anterior-Posterior Nasal Spine.*

VT-H: *Vocal Tract-Horizontal.*

VT-V: *Vocal Tract-Vertical.*

VT-O: *Vocal Tract-Oral.*

LTh: *Lip-thickness.*

ACL: *Anterior Cavity Length.*

PCL: *Posterior Cavity Length.*

OPhW: *Oropharyngeal-width.*

NphL: *Nasopharyngeal-length*.

HG: *Hyoglossus*.

GGp: *Genioglossus posterior*.

GGa: *Genioglossus anterior*.

OOS: *Orbicularis oris superior*.

Tp: *Tongue position*.

Jp: *Jaw position*.

Ts: *Tongue shape*.

Tt: *Tongue tip*.

CPG: *Central Pattern Generators*.

C1, C2, C3, C4: *vértebras cervicales 1, 2, 3 y 4*.

## LISTA DE TABLAS

Tabla 2.1. Valores de los dos primeros formantes de las vocales de Eguchi y Hirsh (1969).....	133
Tabla 2.2. Valores de los dos primeros formantes de las vocales de Clement y Wijnen (1994).....	142
Tabla 2.3. Valores de los dos primeros formantes de las vocales de Kuhl y Meltzoff (1996).....	143
Tabla 2.4. Valores de los dos primeros formantes de las vocales de Gilbert et al. (1997). .....	144
Tabla 2.5. Valores de los dos primeros formantes de las vocales de Rob et al. (1997). .....	147
Tabla 2.6. Valores de los dos primeros formantes de los sujetos femeninos del trabajo de Huber et al. (1999).....	150
Tabla 2.7. Valores de los dos primeros formantes de los sujetos masculinos del trabajo de Hubert et al. (1999).....	150
Tabla 5.1. Sesiones de grabación de la niña. ....	192
Tabla 5.2. Sesiones de grabación del niño .....	193
Tabla 6.1. Sonidos vocálicos modales identificados en la niña en el primer año de vida (0;4-1;0 año).....	231
Tabla 6.2. Sonidos vocálicos nasales identificados en la niña en el primer año de vida (0;4-1;0 año).....	231
Tabla 6.3. Sonidos vocálicos modales identificados en la niña en el segundo año de vida (1;0-2;0 años). ....	232
Tabla 6.4. Sonidos vocálicos nasales identificados en la niña en el segundo año de vida (1;0-2;0 años). ....	233
Tabla 6.5. Sonidos vocálicos modales identificados en la niña en el tercer año de vida (2;0-3;0 años). ....	234
Tabla 6.6. Sonidos vocálicos nasales identificados en la niña en el tercer año de vida (2;0-3;0 años). ....	235
Tabla 6.7. Sonidos vocálicos modales identificados en el niño en el primer año de vida (0;4-1;0 año).....	236
Tabla 6.8. Sonidos vocálicos nasales identificados en el niño en el primer año de vida (0;4-1;0 año).....	236
Tabla 6.9. Sonidos vocálicos modales identificados en el niño en el segundo año de vida (1;0-2;0 años). ....	237



Tabla 6.10. Sonidos vocálicos nasales identificados en el niño en el segundo año de vida (1;0-2;0 años). .....	238
Tabla 6.11. Sonidos vocálicos modales identificados en el niño en el tercer año de vida (2;0-3;0 años). .....	240
Tabla 6.12. Sonidos vocálicos nasales identificados en el niño en el tercer año de vida (2;0-3;0 años). .....	241
Tabla 7.1. Valores de los promedios de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año). .....	252
Tabla 7.2. Valores de los promedios de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años). .....	253
Tabla 7.3. Valores de los promedios de los dos primeros formantes de la vocal [a] de la niña en el tercer año de vida (2;0-3;0 años). .....	253
Tabla 7.4. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año). .....	256
Tabla 7.5. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años). .....	256
Tabla 7.6. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el tercer año de vida (2;0-3;0 años). .....	257
Tabla 7.7. Valores de los promedios de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año). .....	261
Tabla 7.8. Valores de los promedios de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años). .....	261
Tabla 7.9. Valores de los promedios de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años). .....	261
Tabla 7.10. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año). .....	266
Tabla 7.11. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años). .....	266
Tabla 7.12. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años). .....	267
Tabla 7.13. Valores de los promedios de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año). .....	271
Tabla 7.14. Valores de los promedios de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años). .....	271
Tabla 7.15. Valores de los promedios de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años). .....	272

Tabla 7.16. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año).....	275
Tabla 7.17. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años).....	275
Tabla 7.18. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años).....	276
Tabla 7.19. Valores de los promedios de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año). ....	280
Tabla 7.20. Valores de los promedios de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años). ....	280
Tabla 7.21. Valores de los promedios de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años). ....	280
Tabla 7.22. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año).....	284
Tabla 7.23. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años).....	284
Tabla 7.24. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años).....	285
Tabla 7.25. Valores de los promedios de los dos primeros formantes de la vocal [u] de la niña en el primer año de vida (0;4-1;0 año). ....	288
Tabla 7.26. Valores de los promedios de los dos primeros formantes de la vocal [u] de la niña en el segundo año de vida (1;0-2;0 años). ....	288
Tabla 7.27. Valores de los promedios de los dos primeros formantes de la vocal [u] de la niña en el tercer año de vida (2;0-3;0 años). ....	288
Tabla 7.28. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña en el primer año de vida (0;4-1;0 año).....	290
Tabla 7.29. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña en el tercer año de vida (2;0-3;0 años).....	290
Tabla 7.30. Valores de los promedios de los dos primeros formantes de la vocal [a] del niño en el primer año de vida (0;4-1;0 año).....	297
Tabla 7.31. Valores de los promedios de los dos primeros formantes de la vocal [a] del niño en el segundo año de vida (1;0-2;0 años).....	297
Tabla 7.32. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] del niño en el primer año de vida (0;4-1;0 año).....	298
Tabla 7.33. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] del niño en el segundo año de vida (1;0-2;0 años).....	298

Tabla 7.34. Valores de los promedios de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año).....	302
Tabla 7.35. Valores de los promedios de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años).....	302
Tabla 7.36. Valores de los promedios de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años).....	302
Tabla 7.37. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año).....	306
Tabla 7.38. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años).....	307
Tabla 7.39. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años).....	307
Tabla 7.40. Valores de los promedios de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año).....	311
Tabla 7.41. Valores de los promedios de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años).....	311
Tabla 7.42. Valores de los promedios de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años).....	311
Tabla 7.43. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año).....	315
Tabla 7.44. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años).....	315
Tabla 7.45. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años).....	315
Tabla 7.46. Valores de los promedios de los dos primeros formantes de la vocal [o] del niño en el primer año de vida (0;4-1;0 año).....	319
Tabla 7.47. Valores de los promedios de los dos primeros formantes de la vocal [o] del niño en el segundo año de vida (1;0-2;0 años).....	319
Tabla 7.48. Valores de los promedios de los dos primeros formantes de la vocal [o] del niño en el tercer año de vida (2;0-3;0 años).....	320
Tabla 7.49. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en el primer año de vida (0;4-1;0 año).....	321
Tabla 7.50. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en el segundo año de vida (1;0-2;0 años).....	321
Tabla 7.51. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en el tercer año de vida (2;0-3;0 años).....	322

Tabla 7.52. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el primer año de vida (0;4-1;0 año). .....	325
Tabla 7.53. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el segundo año de vida (1;0-2;0 años). .....	325
Tabla 7.54. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años). .....	326
Tabla 7.55. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en el segundo año de vida (1;0-2;0 años). .....	329
Tabla 7.56. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años). .....	329
Tabla 7.57. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña. ....	334
Tabla 7.58. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [a] del niño. ....	334
Tabla 7.59. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña. ....	335
Tabla 7.60. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [e] del niño. ....	335
Tabla 7.61. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña. ....	337
Tabla 7.62. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [i] del niño. ....	337
Tabla 7.63. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña. ....	338
Tabla 7.64. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [o] del niño. ....	338
Tabla 7.65. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña. ....	339
Tabla 7.66. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [u] del niño. ....	340
Tabla 8.1. Promedios de los valores formánticos de las vocales de un adulto mujer (Martínez Celdrán y Fernández Planas, 2007). ....	342
Tabla 8.2. Promedios de los valores formánticos de las vocales de la niña entre los 0;4-1;0 año. ....	343
Tabla 8.3. Promedios de los valores formánticos de las vocales de la niña entre los 1;0-2;0 años. ....	343

Tabla 8.4. Promedios de los valores formánticos de las vocales de la niña entre los 2;0-3;0 años. ....	343
Tabla 8.5. Promedios de los valores formánticos de las vocales de un adulto varón (Martínez Celdrán y Fernández Planas, 2007). ....	344
Tabla 8.6. Promedios de los valores formánticos de las vocales del niño entre los 0;4-1;0 año. ....	344
Tabla 8.7. Promedios de los valores formánticos de las vocales del niño entre los 1;0-2;0 años. ....	344
Tabla 8.8. Promedios de los valores formánticos de las vocales del niño entre los 2;0-3;0 años. ....	345
Tabla 8.9. Promedios de los valores formánticos de las cinco vocales del español en los seis sujetos de desarrollo normal cuyos formantes se compararon con los de los niños con labio partido (Casal et al., 2002). ....	347
Tabla 8.10. Promedios de los valores formánticos de las cinco vocales del español en los siete sujetos de desarrollo normal cuyos formantes se compararon con los de los niños con paladar partido (Casal et al., 2002). ....	347
Tabla 8.11. Promedios de los valores formánticos de las cinco vocales del español en los siete sujetos de desarrollo normal cuyos formantes se compararon con los de los niños que presentaron labio y paladar partido de manera unilateral (Casal et al., 2002). ....	347
Tabla 8.12. Promedios de los valores formánticos de las cinco vocales del español en los dos sujetos de desarrollo normal cuyos formantes se compararon con los de los niños que presentaron labio y paladar partido de manera completa (Casal et al., 2002). ....	348
Tabla 8.13. Promedios de los valores formánticos de la vocal [a] de un niño (Buhr, 1980). ....	349
Tabla 8.14. Promedios de los valores formánticos de la vocal [e] de un niño (Buhr, 1980). ....	349
Tabla 8.15. Promedios de los valores formánticos de la vocal [i] de un niño (Buhr, 1980). ....	349
Tabla 8.16. Promedios de los valores formánticos de la vocal [o] de un niño (Buhr, 1980). ....	350
Tabla 8.17. Promedios de los valores formánticos de la vocal [u] de un niño (Buhr, 1980). ....	350
Tabla 8.18. Promedios de los valores formánticos de las vocales de los extremos (Eguchi y Hirsh, 1969). ....	351
Tabla 8.19. Promedios de los valores formánticos de las vocales de los extremos (Kuhl y Meltzoff, 1996). ....	351

Tabla 8.20. Valores formánticos mínimos y máximos de las vocales de un niño (Kent y Murray, 1982). .....	353
Tabla 8.21. Valores formánticos mínimos y máximos de las vocales de la niña de nuestro estudio. ....	354
Tabla 8.22. Valores formánticos mínimos y máximos de las vocales del niño de nuestro estudio. ....	354
Tabla 8.23. Promedios de los valores formánticos de las vocales de cuatro niños (Gilbert et al., 1997). ....	355
Tabla 8.24. Promedios de los valores formánticos de las vocales de los sujetos de nuestro estudio. ....	355
Tabla 8.25. Promedios de los valores formánticos de las vocales de veinte niños (Robb et al., 1997). ....	356
Tabla 8.26. Promedios de los valores formánticos de las vocales de los dos sujetos de nuestro estudio. ....	357
Tabla 8.27. Frecuencia de los fonemas vocálicos en español (Quilis y Esgueva, 1980). ....	367

## LISTA DE GRÁFICOS

Gráfico 6.1. Carta de formantes de la niña 0;4-0;5 años. ....	208
Gráfico 6.2. Carta de formantes de la niña 0;5-0;6 años. ....	209
Gráfico 6.3. Carta de formantes de la niña 0;6-0;7 años. ....	210
Gráfico 6.4. Carta de formantes de la niña 0;7-0;8 años. ....	211
Gráfico 6.5. Carta de formantes de la niña 0;8-0;9 años. ....	212
Gráfico 6.6. Carta de formantes de la niña 0;9-0;10 años. ....	213
Gráfico 6.7. Carta de formantes de la niña 0;10-0;11 años. ....	214
Gráfico 6.8. Carta de formantes de la niña 0;11-1;0 año. ....	215
Gráfico 6.9. Carta de formantes de la niña 1;0-1;6 años. ....	216
Gráfico 6.10. Carta de formantes de la niña 1;6-2;0 años. ....	217
Gráfico 6.11. Carta de formantes de la niña 2;0-2;6 años. ....	218
Gráfico 6.12. Carta de formantes de la niña 2;6-3;0 años. ....	219
Gráfico 6.13. Carta de formantes del niño 0;4-0;5 años. ....	220
Gráfico 6.14. Carta de formantes del niño 0;5-0;6 años. ....	221
Gráfico 6.15. Carta de formantes del niño 0;6-0;7 años. ....	221
Gráfico 6.16. Carta de formantes del niño 0;7-0;8 años. ....	222
Gráfico 6.17. Carta de formantes del niño 0;8-0;9 años. ....	223
Gráfico 6.18. Carta de formantes del niño 0;9-0;10 años. ....	224
Gráfico 6.19. Carta de formantes del niño 0;10-0;11 años. ....	224
Gráfico 6.20. Carta de formantes del niño 0;11-1;0 año. ....	225
Gráfico 6.21. Carta de formantes del niño 1;0-1;6 años. ....	226
Gráfico 6.22. Carta de formantes del niño 1;6-2;0 años. ....	227
Gráfico 6.23. Carta de formantes del niño 2;0-2;6 años. ....	228
Gráfico 6.24. Carta de formantes del niño 2;6-3;0 años. ....	229
Gráfico 7.1. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año). ....	250
Gráfico 7.2. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años). ....	251
Gráfico 7.3. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en el tercer año de vida (2;0-3;0 años). ....	251
Gráfico 7.4. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en los tres primeros años de vida (0;4-3;0 años). ....	252
Gráfico 7.5. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año). ....	254

Gráfico 7.6. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años). .....	255
Gráfico 7.7. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en los tres primeros años de vida (0;4-3;0 años). .....	256
Gráfico 7.8. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año).....	258
Gráfico 7.9. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años). .....	259
Gráfico 7.10. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años).....	260
Gráfico 7.11. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en los tres primeros años de vida (0;4-3;0 años).....	260
Gráfico 7.12. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año).....	263
Gráfico 7.13. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años). .....	264
Gráfico 7.14. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años).....	265
Gráfico 7.15. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en los tres primeros años de vida (0;4-3;0 años). .....	266
Gráfico 7.16. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año).....	268
Gráfico 7.17. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años). .....	269
Gráfico 7.18. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años).....	270
Gráfico 7.19. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en los tres primeros años de vida (0;4-3;0 años).....	271
Gráfico 7.20. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año). .....	273
Gráfico 7.21. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años). .....	274
Gráfico 7.22. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años). .....	274
Gráfico 7.23. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en los tres primeros años de vida (0;4-3;0 años). .....	275



Gráfico 7.24. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año).....	277
Gráfico 7.25. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años). .....	278
Gráfico 7.26. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años).....	279
Gráfico 7.27. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en los tres primeros años de vida (0;4-3;0 años).....	279
Gráfico 7.28. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año). .....	282
Gráfico 7.29. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años). .....	282
Gráfico 7.30. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años). .....	283
Gráfico 7.31. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en los tres primeros años de vida (0;4-3;0 años). .....	284
Gráfico 7.32. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en el primer año de vida (0;4-1;0 año).....	286
Gráfico 7.33. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en el segundo año de vida (1;0-2;0 años). .....	286
Gráfico 7.34. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en el tercer año de vida (2;0-3;0 años).....	287
Gráfico 7.35. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en los tres primeros años de vida (0;4-3;0 años).....	288
Gráfico 7.36. Evolución de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña en los tres primeros años de vida (0;4-3;0 años). .....	290
Gráfico 7.37. Área vocálica de la niña en el primer año de vida (0;4-1;0 año).....	291
Gráfico 7.38. Área vocálica de la niña en el segundo año de vida (1;0-2;0 años). .....	292
Gráfico 7.39. Área vocálica de la niña en el tercer año de vida (2;0-3;0 años).....	293
Gráfico 7.40. Evolución del área vocálica de la niña en los tres primeros años de vida (0;4-3;0 años). .....	294
Gráfico 7.41. Evolución de los promedios de los primeros formantes de la vocal [a] del niño en el primer año de vida (0;4-1;0 año). .....	295
Gráfico 7.42. Evolución de los promedios de los primeros formantes de la vocal [a] del niño en el segundo año de vida (1;0-2;0 años).....	296

Gráfico 7.43. Evolución de los promedios de los primeros formantes de la vocal [a] del niño en los dos primeros años de vida (0;4-2;0 años). .....	296
Gráfico 7.44. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] del niño en los dos primeros años de vida (0;4-2;0 años). .....	298
Gráfico 7.45. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año). .....	299
Gráfico 7.46. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años). .....	300
Gráfico 7.47. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años). .....	301
Gráfico 7.48. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en los tres primeros años de vida (0;4-3;0 años). .....	301
Gráfico 7.49. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año). .....	304
Gráfico 7.50. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años). .....	304
Gráfico 7.51. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años). .....	305
Gráfico 7.52. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en los tres primeros años de vida (0;4-3;0 años). .....	306
Gráfico 7.53. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año). .....	308
Gráfico 7.54. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años). .....	309
Gráfico 7.55. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años). .....	310
Gráfico 7.56. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en los tres primeros años de vida (0;4-3;0 años). .....	310
Gráfico 7.57. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año). .....	312
Gráfico 7.58. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años). .....	313
Gráfico 7.59. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años). .....	314
Gráfico 7.60. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en los tres primeros años de vida (0;4-3;0 años). .....	314

Gráfico 7.61. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en el primer año de vida (0;4-1;0 año). .....	316
Gráfico 7.62. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en el segundo año de vida (1;0-2;0 años). .....	317
Gráfico 7.63. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en el tercer año de vida (2;0-3;0 años). .....	318
Gráfico 7.64. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en los tres primeros años de vida (0;4-3;0 años). .....	319
Gráfico 7.65. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en los tres primeros años de vida (0;4-3;0 años). .....	321
Gráfico 7.66. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el primer año de vida (0;4-1;0 año). .....	323
Gráfico 7.67. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el segundo año de vida (1;0-2;0 años). .....	324
Gráfico 7.68. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años). .....	324
Gráfico 7.69. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en los tres primeros años de vida (0;4-3;0 años). .....	325
Gráfico 7.70. Evolución de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años). .....	328
Gráfico 7.71. Evolución de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en los dos últimos años de vida (1;0-3;0 años). .....	328
Gráfico 7.72. Área vocálica del niño en el primer año de vida (0;4-1;0 año). .....	330
Gráfico 7.73. Área vocálica del niño en el segundo año de vida (1;0-2;0 años). .....	331
Gráfico 7.74. Área vocálica del niño en el tercer año de vida (2;0-3;0 años). .....	332
Gráfico 7.75. Evolución del área vocálica del niño en los tres primeros años de vida (0;4-3;0 años). .....	333



## 1. INTRODUCCIÓN.

Llama la atención de todo investigador cuando este se acerca a la literatura sobre el estudio de los sonidos emitidos por bebés y niños pequeños desde un punto de vista fonético, el hecho de que la mayor parte de los trabajos se centre en la caracterización de los sonidos consonánticos cuando la vocal es el elemento crucial de la lengua humana (Maddieson y Ladefoged [1995]), la discriminación vocálica es fundamental en el procesamiento del habla (Kewley-Port y Watson [1994]), y al surgir en el habla temprana las vocales son importantes hitos para el desarrollo del lenguaje (Vorperian y Kent [2007]).

Muchos autores denuncian este abandono generalizado del análisis vocálico y coinciden en que entre las razones que motivan esta escasez se encuentran por un lado la falsa creencia de que las vocales son aprendidas con poco esfuerzo, y por otro el que muy pocas veces las vocales se articulan mal (Davis y MacNeilage [1990], Palethorpe et al. [1996], Stokes y Wong [2002] y Van der Stelt et al. [2005]).

Sin embargo, tal y como se refuta en el trabajo de Stokes y Wong (2002) la experiencia confirma lo contrario: en el estudio de Paschall (1983) se encontró que la precisión media con la que veinte niños de la edad de 1;4-1;6 años producían las vocales estaba por debajo del 60%; de la misma forma aunque con niños algo mayores, el estudio de Otomo y Stoel-Gammon (1992) mostró como el porcentaje medio de producción correcta de vocales no redondeadas para infantes de la edad de 2;2 años era del 63,5%. Resultados similares arrojaron los trabajos que analizaron el habla de niños que presentaban desórdenes comunicativos como el de Pollock y Keiser (1990) quienes al examinar quince infantes con desórdenes fonológicos descubrieron que la mitad de los sujetos produjo vocales con menos del 80 % de precisión. De la misma forma, el estudio de Stoel-Gammon y Herrington (1990) reveló errores vocálicos en dos niños con desorden fonológico, y el seguimiento llevado a cabo por Pollock y Hall (1991) a niños con apraxia del habla también mostró la incorrecta articulación vocálica. Estos trabajos invitan a pensar que las vocales presentan unas particularidades, ya sea tipo acústico o de origen anatómico, que hacen que estas no resulten tan fáciles de articular por los infantes como se dedujo del estudio de Waterson (1971), aunque esta creencia haya ocasionado una falta de interés teórico por estos sonidos en comparación con el de las consonantes (Davis y MacNeilage [1990], Palethorpe et al. [1996] y Stokes y Wong [2002]), y consecuentemente una escasez de investigaciones tal y como afirman Chung et al. (2008).

La relevancia que tiene el estudio de los sonidos vocálicos en edades tempranas ha sido subrayada por su papel en la comprensión de otros componentes fonético-fonológicos de una

lengua. De este modo, autores como Werker y Polka (1993) o Lee et al. (2009) señalan la importancia de una investigación vocálica en la temprana infancia para diagnosticar a qué edad puede verse la influencia de la lengua ambiente en las emisiones del bebé ya que piensan que dicha influencia se manifiesta más en las vocales que en las consonantes por la mayor estabilidad formántica de aquellas.

Además de esta estabilidad formántica, otras son las particularidades que diferencian los sonidos vocálicos de los consonánticos y puesto que nuestro trabajo se centrará en los primeros, creemos conveniente referirnos brevemente a estas. Desde el punto de vista articulatorio las vocales se caracterizan por facilitar la salida del aire procedente de los pulmones sin oponer una gran resistencia. Las consonantes por su parte, obstaculizan en mayor medida dicha salida siendo el grado y el modo de interrupción del flujo del aire los responsables de los distintos tipos de consonantes (consonantes aproximantes si la obstrucción a la salida del aire es ligeramente mayor que en el caso de las vocales; consonantes oclusivas si dicha salida se obstruye completamente, etc.).

Desde una perspectiva acústica, las vocales presentan una mayor intensidad que las consonantes siendo sus resonancias formánticas más intensas que las de las consonantes (interpretándose dicha intensidad según Martínez Celdrán [1998] como el correlato acústico de la amplitud de la onda creado articulatoriamente por la mayor o menor energía de aire espirado).

Finalmente, también desde un enfoque auditivo pueden distinguirse las vocales de las consonantes, al menos de las oclusivas, ya que el oído humano percibe estas últimas de forma categórica puesto que en el continuum de [p] a [b] el oyente puede reconocer en un punto determinado la [b], frente a lo que ocurre con las vocales las cuales son percibidas por el oído de forma continua (los trabajos y experimentos que muestran las diferencias en la percepción de estos sonidos serán explicados con mayor detenimiento en el apartado 2.2 del estado de la cuestión dedicado a las capacidades perceptivas del bebé y niño pequeño). Por tanto los sonidos vocálicos, además de percibirse de manera diferente a las consonantes, son más sonoros y estables que las consonantes y pueden ayudar además a su percepción.

Junto a estas características y a las ofrecidas en el trabajo de Werker y Polka (1993), Vihman (1996) recalca la importancia del estudio de las vocales en hablantes de corta edad por ser estas además las que transportan la información prosódica y los rasgos de identidad del hablante.

Una vez justificada la necesidad del estudio vocálico en hablantes infantiles, consideramos necesario que dicho análisis comprenda al menos los tres primeros años de vida

debido por un lado a que como exponen Selby et al. (2000), y tal y como veremos en la revisión fisiológica del estado de la cuestión, los mayores cambios anatómicos del tracto vocal ocurren entre el nacimiento y los tres años de edad, y debido por otro lado a que estudios como el de Templin (1957) han demostrado que la producción correcta de la vocal no mejora entre los 3;0 y los 8;0 años de vida sugiriendo que las características fundamentales de las vocales aparecen antes de los 3;0 años.

Durante estos tres primeros años de vida muchos son los autores que interesados en la adquisición fonológica de la lengua han investigado las emisiones de los bebés y niños pequeños estableciendo distintas etapas de adquisición las cuales según la mayoría de estudios pueden agruparse en dos grandes periodos: el periodo prelingüístico y el lingüístico. A continuación exponemos brevemente en qué consisten.

El periodo prelingüístico o etapa presemiótica, según Alarcos Llorach et al. (1976), es aquella etapa en la que el bebé experimenta los gestos articulatorios que le llevarán posteriormente a emitir la primera palabra en la etapa prelingüística. Se trata de la ejercitación del aparato fonador que se traduce según estos autores en “una intensa actividad fónica que sirve de preludio al futuro buen funcionamiento de los órganos destinados a materializar el lenguaje” (1976:12). Dichas emisiones son “indiferenciadas desde el punto de vista semántico” y aunque los adultos puedan atribuirles algún sentido, para el niño no son aún señales simbólicas sino simplemente emisiones para llamar la atención de quienes les rodean.

Según las características de las emisiones que produce el infante a lo largo de este periodo prelingüístico se suelen distinguir cinco etapas aunque el número de estas puede variar según el autor considerado (los distintos autores y las etapas diferenciadas quedan recogidos en el libro de adquisición fonológica de Vihman [1996]). Si seguimos el manual de adquisición del lenguaje de Serra et al. (2000) encontramos las siguientes:

- Etapa de fonación o de vocalizaciones reflejas. Abarca desde los 0 a los 0;2 años y en ella encontramos por un lado los sonidos que producen los bebés en relación a los cambios que tienen que ver con la nutrición, la respiración, el tragado, etc., y por otro los sonidos procedentes de las situaciones de dolor, placer o confort. Se trata por tanto de sonidos vegetativos.
- Etapa de gagueo y sonrisas. Tiene lugar entre los 0;2 y 0;4 años y su nombre se debe al aumento de las producciones velares (gagueo) producidas normalmente por las sonrisas y el habla que el interlocutor dirige al bebé.

- Etapa de expansión fonética o juego vocal. Es el periodo comprendido entre los 0;4 y los 0;6 años y a lo largo de esta el bebé aumenta el control articulatorio con lo que es probable que este produzca sonidos de manera voluntaria. En esta etapa aparecen por primera vez los sonidos vocálicos con resonancia completa y el bebé es además capaz de producir sonidos consonánticos acompañados de ruidos, gritos y otras manifestaciones expresivas.
- Etapa del balbuceo canónico. Se inicia en el sexto mes de vida y en ella aparece por primera vez la unidad silábica constituida por un cierre (consonante) y una abertura (vocal) similar a la del adulto. El balbuceo canónico puede dividirse en dos: el balbuceo reduplicativo que no es sino la repetición de gestos articulatorios de manera constante y con diferente entonación, y el balbuceo melódico que se da cuando el bebé modifica la vocal y la consonante e introduce diferentes contornos melódicos.
- Etapa del balbuceo variado o conversacional. Surge a los 0;10 años y se caracteriza por el aumento de la producción de consonantes y de sus tipos y por el incremento de los registros tonales y melódicos. El término *conversacional* alude a la capacidad del niño de seguir turnos alternativos de habla.

A partir de los 0;10 años el niño comienza a hacer uso de determinados gestos acompañándolos o no de vocalizaciones para pedir objetos (señalándolos con alguna parte del cuerpo), para mostrarlos o para entregarlos. La presencia de estos gestos y la tendencia a acompañarlos con vocalizaciones aumentará con el paso del tiempo. Se trata en palabras de Vihman (1996) de formas de transición (entre el balbuceo y la primera palabra) que sirven para facilitar al infante la inmersión en la función simbólica del lenguaje.

Pasado el primer año de vida el niño comienza a atribuir a los sonidos que emite una cierta intención simbólica, la misma que caracteriza a la palabra, intención que no es otra que la de utilizar dichos sonidos para anunciar un acto antes de que el acto se produzca. Estamos ya en el periodo propiamente lingüístico, en concreto en el de la adquisición de la primera palabra (si bien Miquel Serra et al. [2000] nos advierten de la necesidad de considerar estas etapas como transformaciones hacia otros tipos de comportamientos sin que por ello desaparezcan los anteriores). Importante es aquí el input puesto que tal y como nos dicen Alarcos Llorach et al. (1976) tras observar que en una situación determinada los adultos repiten una misma expresión, el niño asocia una a la otra y si dicha expresión se incluye dentro de la sucesión de elementos fónicos que él ya sabe reproducir, el sujeto podrá utilizar dicho signo recién adquirido no ya



para acompañar a una situación u objeto sino para reclamar dicho objeto o provocar aquella situación.

En cuanto a la forma que presentan las primeras palabras, en un primer momento estas son muy simplificadas con relación a las palabras equivalentes en el modelo adulto y esta simplificación viene dada porque muchos de los fonemas emitidos por el niño son imitados de una manera imperfecta, pueden ser omitidos o cambiados por otros. Estamos ante los procesos fonológicos de simplificación del habla.

Y en relación al número de palabras que emite el niño, en un primer momento el niño emite una única palabra y ello ha llevado a muchos autores a denominar esta etapa como la etapa de la palabra-frase, o etapa holofrástica, que suele aparecer a partir de primer año de vida y que se denomina de esta forma porque en ella el niño utiliza una sola palabra a modo de frase. A partir del año y medio aproximadamente, el niño comenzará a combinar varias palabras surgiendo entonces una sintaxis primitiva que se produce según Alarcos Llorach et al. (1976) mediante la yuxtaposición de dos signos (o signos-frase). Estamos ante la etapa telegráfica. Pasados los dos años de edad, el niño irá reproduciendo los marcadores gramaticales como la flexión de género y número, la conjugación verbal, los morfemas de diminutivo, etc. y a partir de los tres años, el niño que aún conservará el lenguaje teleográfico, comenzará a construir enunciados de más de dos palabras añadiendo por tanto complejidad a la frase.

Después de haber expuesto las características que rodean las emisiones de los infantes desde el punto de vista de la adquisición de una lengua con el objetivo de presentar el marco en el que se mueven la mayor parte de los trabajos centrados en el habla infantil en el periodo que comprende nuestro estudio, esto es entre los 0;4 y los 3;0 años, nos referimos a continuación a las perspectivas desde las que un estudio como el nuestro, cuyo interés está en la descripción física de los sonidos vocálicos emitidos por bebés y niños pequeños, puede abordarse. Entre ellas encontramos el punto de vista anatómico y de control motor (incluyendo el desarrollo de las habilidades articulatorias y el control neuromuscular), el acústico y el perceptivo (y en relación directa con este último encontramos la influencia de la lengua ambiente).

Como el lector habrá observado, se trata de las tres ramas tradicionales desde las que puede acometerse el estudio de los sonidos de una lengua como materia física, no lingüística, estudio del que tradicionalmente se ha ocupado la Fonética. No obstante, este perspectivismo cobra importancia al tratarse de los sonidos del habla infantil, y aún más al considerarse edades tan tempranas como las que comprende nuestro trabajo, puesto que ayuda a determinar el origen lingüístico en el ser humano al averiguar los requisitos anatómicos y de control motor necesarios para producir los sonidos del habla, al describir las características acústicas iniciales

de las primeras emisiones del humano y su evolución hasta acercarse a las del adulto, y por supuesto, al tratar de determinar las capacidades perceptivas con las que nace el bebé y su desarrollo.

La importancia de la consideración anatómica y perceptiva a la hora de estudiar los sonidos del infante ha sido señalada por distintos autores como Vorperian y Kent (2007:1511): “[...] An ultimate goal is the integration of acoustic data with anatomic, physiologic and perceptual data, to produce a comprehensive account of patterns in the development of speech”, o por Chen y Kent (2010) quienes en su artículo sobre la producción segmental de infantes que hablaban mandarín expusieron que el comportamiento fonético temprano parece ser determinado por tres factores: la competencia articulatoria, la influencia de la lengua ambiente y la capacidad neural. De la misma opinión son los estudiosos Oller y MacNeilage (1983) que coinciden con Chen y Kent (2010) en la creencia de que entre los factores que determinan las características de los sonidos de los niños se encuentran los factores motores, cognitivos y perceptivos, o Serkhane et al. (2002) quienes piensan que las vocalizaciones del bebé y del niño pequeño son el resultado de una adaptación a los cambios del tracto vocal así como de un desarrollo guiado por la percepción. Estos últimos sugieren que son dos los mecanismos que intervienen la producción de habla de los infantes: la exploración de las habilidades sensoriales-motoras del tracto vocal, y la imitación de los sonidos percibidos de la lengua de los cuidadores.

También Kent (2004) cree que las particularidades del habla de los infantes pueden explicarse como una combinación de factores relacionados con la percepción y las habilidades motoras.

Por su parte Rvachew et al. (2006:2251) consideran que estas perspectivas nos ayudan a explicar los cambios acústicos que experimenta la vocal en la infancia: “[...] Several factors can explain the observed acoustic shifts in vowel production across the infancy period: anatomical growth, motor control development, auditory (peripheral and central) abilities, and other cognitive factors”.

En cuanto al primero de los enfoques, los múltiples trabajos que han analizado la producción de sonidos vocálicos o consonánticos por el infante desde el punto de vista fisiológico coinciden en que la configuración anatómica del niño (incluyendo esta la morfología y el desarrollo del tracto vocal, de todas sus estructuras y sus órganos) y el desarrollo del control motor son componentes esenciales en la producción vocal del bebé y del niño pequeño.

Según Bosma (1975, 1976), el surgimiento y el desarrollo del habla está determinado, al menos en parte, por los cambios físicos que sufren el tracto vocal, sus estructuras y sus órganos,

durante el desarrollo y el crecimiento del bebé. En “Anatomical and neuromuscular maturation of the speech” Kent (1976) afirmó que son tres las áreas implicadas en el desarrollo de las estructuras y de las funciones para el habla: los cambios del aparato fonador, la forma del trato vocal y el *timing*, y la coordinación de la articulación. En un trabajo posterior acerca de la psicobiología del desarrollo del lenguaje, Kent (1984) manifestó que para la emisión de los sonidos de una lengua es imprescindible el control motor de los órganos por parte del niño (capacidad del bebé para configurar su tracto vocal y dirigir los músculos para producir los sonidos). Más tarde y junto con Murray, los autores subrayaron que “[...] the anatomic determinants of sound patterns, such critical anatomic changes... should be carefully considered in explaining patterns of change in infant vocalizations” (1992:353).

Por otro lado, Clement y Wijnen (1994:83) afirmaron que:

“[...] The development of the ability to produce the speech sounds of a language depends on two processes: (a) articulatory learning (i.e., shaping and refining articulatory motor control processes, including the perceptually based skills of monitoring and feedback) and (b) the development of a vocal tract allowing for the implementation of articulatory configurations yielding recognizable acoustic tokens”.

Por su parte, Vorperian et al. (1999) en un trabajo en el que usaron resonancias magnéticas para estudiar el desarrollo de las diferentes estructuras del tracto vocal, afirmaron que la información sobre los cambios anatómicos que sufre el tracto vocal contribuye al mejor entendimiento de las características de los sonidos que emite el niño. También con ayuda de un estudio que se valió de imágenes de resonancia magnética, en este caso para intentar establecer el desarrollo anatómico de las porciones oral y faríngea del tracto vocal, Vorperian et al. (2008) indicaron que las propiedades del habla de los niños cambian con la maduración de la anatomía del tracto vocal y el aumento de tamaño y longitud de este.

El ya citado Hopkins (2005) señaló la importancia no solo del desarrollo del tracto vocal sino de la maduración del sistema nervioso cuando al hablar de los constituyentes que intervienen en la producción del lenguaje especificó que estos dependen de la formación y desarrollo del sistema nervioso y de los sistemas de producción de habla incluyendo el sistema respiratorio, la laringe y el tracto vocal. Otros autores como Guenther (1994), Callan et al. (2000) y Guenther y Perkell (2004) crean incluso modelos que imitan redes neuronales para mostrar la importancia que la configuración anatómica y el sistema nervioso tienen en la producción del habla (tal y como explicaremos en el apartado del estado de la cuestión).

También Smith (2010:1) subrayó la relevancia del sistema cognitivo y nervioso al afirmar que para que pueda producirse el habla el sistema nervioso debe generar una serie de órdenes a los músculos para alcanzar las metas lingüísticas:

“[...] These command signals must be coordinated in time and space for the appropriate sequences of muscle activation to occur. Therefore, we can attempt to understand speech motor control processes by investigating where in the nervous system these neural commands to muscles are generated and how they are modified to achieve a variety of linguistic and metalinguistic goals”.

La importancia de los factores cognitivos está presente ya en el feto tal y como describe el trabajo de Ramón y Cajal (1996) en el que a partir de la observación de movimientos laríngeos fonatorios en el feto se sugirió que la producción de un sonido refleja una actividad cerebral.

Respecto al segundo enfoque desde el que puede acometerse el estudio vocálico, estudios como los Oller et al. (1985), Oller y Eilers (1988), Koopmans-van Beinum et al. (2001) o Van der Stelt et al. (2008), entre otros, revelaron la influencia de la percepción en la producción y en las características de los sonidos de una lengua considerando como prueba el hecho de que los niños que han nacido sordos o que poseen problemas de audición producen de manera deficiente los sonidos propios de las distintas etapas en comparación con los bebés de audición normal. Un ejemplo de como la percepción normal afecta también al rango de frecuencias formánticas producidas se encuentra en el estudio de Rvachew y Slawinski (1995) en el que después de analizar acústicamente las vocales del balbuceo de nueve niños con otitis media a los 0;6, 0;9, 1;0, 1;3 y 1;6 años de edad, se observó que las vocales producidas por algunos de estos niños con otitis media mostraron un rango restringido de valores del segundo formante.

De la misma manera, el trabajo de Dawn Warner-Czyz (2005), en el que se compararon las emisiones de niños con desarrollo normal con las de niños con implantes cocleares, mostró los efectos de la percepción en la producción. En dicho estudio la frecuencia de aparición vocálica y la precisión a la hora de producir las vocales fueron menores en el caso de los niños de entre 0;11 y 1;1 años que habían sido sometidos a un implante coclear a la edad de 0;6 años, que en el caso de los niños con un desarrollo normal de la audición.

También el estudio de McGowan et al. (2008) reveló la importancia que la correcta audición tiene en la producción del habla. En este trabajo se analizaron las muestras de habla espontánea de diez niños con percepción normal y de diez con pérdida de audición a la edad de 1;0 año. Después de realizar análisis acústicos en los que se midió el F1 y el F2 en vocales

acompañadas de consonantes, se observó que los niños de escucha normal mostraron más variabilidad en el F2 que los niños diagnosticados con pérdida de audición. Estos datos indicaron que los niños con percepción normal usan la dimensión anterior-posterior de la cavidad oral más de lo que la utilizan los niños con pérdida de audición.

La importancia de la audición es también demostrada en dos trabajos citados en McGowan et al. (2008) como son el de Kent et al. (1987) y el de Rvachew et al. (1996). En el primero se examinó el desarrollo de la vocal en gemelos entre los 0;8 y los 1;3 años de edad en los que uno de los gemelos presentó pérdida en la audición y otro un desarrollo normal. A lo largo del periodo de estudio se observó una expansión en el espacio vocálico del F1-F2 para el gemelo con percepción normal pero no para su hermano. En el segundo, el artículo de Rvachew et al. (1996), se comparó el habla de niños hablantes de inglés canadiense con audición normal con las emisiones de niños procedentes de la misma comunidad lingüística pero con un comienzo temprano y tardío de otitis media entre los 0;6 y los 1;6 años. Los autores encontraron que los niños pertenecientes al grupo de comienzo temprano de otitis media mostraron retrasos a la hora de dominar el movimiento de la lengua en la dimensión anterior-posterior, y que los sujetos con comienzo tardío de otitis media presentaron una expansión significativa en los rangos del F2.

Otros experimentos en los que se demostró que la percepción condiciona la producción son los de Bloom et al. (1987) y Bloom (1988) en los que tras modificar el input del habla a niños de 0;3 años se observó que estos producían una mayor proporción de emisiones parecidas a las del habla, con una completa resonancia espectral, cuando ellos recibían input de habla. Cuando los infantes escuchaban input que no era del habla, dicha proporción se reducía.

En relación con el espectro acústico, es mediante la percepción que el bebé (al igual que otros animales como los monos según Sommers et al. [1992]) puede detectar la forma espectral de los sonidos que escucha (Bailey y Snowling [2002]), y es la percepción del F1 y F2 de los sonidos vocálicos de una lengua particular la que condiciona dentro de un patrón universal los pequeños descensos o aumentos de los valores de los formantes de los bebés y niños pequeños (Boysson-Bardies et al. [1989], Rvachew et al. [2006] y Kern y Davis [2009]).

Por otro lado, es importante examinar las capacidades perceptivas del infante durante los primeros años de vida debido a que es la percepción de los sonidos de una determinada lengua y no los de otra la que condiciona el tipo de vocal y de patrón silábico que más frecuentemente producirá el bebé o niño. Así lo afirman los trabajos de Kern et al. (en prensa), Boysson-Bardies et al. (1980), Boysson-Bardies et al. (1984), Boysson-Bardies et al. (1992),

Kuhl y Meltzoff (1996), Menard (2005), Lee et al. (2009) y Chen y Kent (2010) cuando dicen que las vocales más comunes pueden ser explicadas por esquemas perceptivos.

De especial relevancia resulta el estudio de las habilidades perceptivas del infante en los trabajos que analizan los primeros sonidos producidos por este debido a la reorganización perceptiva que tiene lugar durante el primer año de vida la cual condicionará la percepción de dicho infante ya que este será sensible a determinados sonidos y no a otros según autores como Swoboda et al. (1976), Werker et al. (1981), Werker (1982), Werker y Tees (1984a), Werker y Tees (1984b), Werker y Logan (1985), Burnham (1986), Studdert-Kennedy (1986), Werker y Lalonde (1988), Best y McRoberts (1989), Werker y Pegg (1992), Kuhl et al. (1992), Werker y Polka (1993), y Moon (2009).

Otras razones que pueden esgrimirse a favor del estudio de las capacidades perceptivas del bebé en el análisis de las vocales de su lengua materna las encontramos en la existencia de un modo de percepción diferente según sea el sonido vocálico o consonántico (Swoboda et al. [1976]), y en la importancia de la percepción en lo que respecta al *feedback* auditivo del niño (Kent y Miolo [1995] y Callan et al. [2000]).

Vinculada a la percepción se encuentra la interacción del bebé o niño con la lengua ambiente que según Clement (2004) influye de manera determinante en el habla del infante. De la misma manera Chen y Kent (2010:342) subrayaron la importancia que tiene la lengua ambiente en las producciones tempranas de los bebés, su dependencia de los procesos generales de percepción y la relación con el desarrollo articulatorio:

“[...] Early phonetic behavior appears to be determined by three primary factors: (1) articulatory competence, related especially to developing vocal tract anatomy and motor function; (2) influence of the ambient language, mediated by auditory *perceptive* processes; and (3) neural capability, both as it relates to the first two factors and as it pertains to other aspects of psychobiological development underlying communication [...] Furthermore, because these factors are closely interwoven, is difficult to identify their separate contributions, except through cross-linguistic studies or studies of clinical populations”.

Años antes, Kent (1984) sostenía que el infante es sensible a algunos aspectos del ambiente lingüístico general y que esta sensibilidad se ve en los sonidos que produce el niño.

Una conclusión subyacente a todos estos trabajos es que el estudio de la adquisición de las vocales de la lengua materna por parte del bebé puede abordarse también desde el punto de

vista de la percepción intentando determinar cuáles son los mecanismos perceptivos con los que nace el bebé, cómo evolucionan y si estos están condicionados o no por la lengua ambiente.

Dentro de esta contemplación multifocal (anatómica, perceptiva y acústica) y en relación al tercer enfoque, el acústico, muchos son los autores que otorgan al análisis acústico un papel especialmente relevante en el estudio de las producciones tempranas del bebé. Así Eguchi y Hirsh (1969) y Kent y Forner (1979) ponen de manifiesto que debido a los cambios que sufre el tracto vocal en cuanto a tamaño y forma en el desarrollo del bebé, es consecuentemente importante estudiar los aspectos acústicos del habla en diferentes edades para alcanzar un conocimiento completo de la producción de habla del bebé o niño pequeño.

En un trabajo sobre el desarrollo anatómico y neuromuscular, Kent (1976:421) declaró que la acústica del habla es un indicativo importante del control motor del bebé y del niño “[...] The acoustic data on various aspects of speech production indicate that the accuracy of motor control improves with age”.

Por otro lado, según Maxwell (1979) y Kornfeld (1971) la importancia de la consideración acústica a la hora de estudiar los sonidos de los infantes se manifiesta en el hecho de que los niños dependen de distinciones acústico-fonéticas que los adultos no perciben para distinguir las distintas categorías de la lengua. Y en cuanto a estas distinciones fonéticas, autores como Delgutte (1997) creen que el oyente identifica la estructura fonética de las distintas emisiones del hablante en base a la onda de la señal acústica. Por su parte, Studdert-Kennedy (1986) piensa que el bebé aprende a hablar gracias a que cada contraste fonético está marcado por un contraste acústico.

Otros estudios se valen de la señal acústica para inferir procesos fisiológicos. Algunos de estos son los de Stathopoulos y Sapienza (1993), Boliek (1997), Huber et al. (1999), Moore et al. (2001) y el de Connaghan et al. (2004) y en ellos se examinó el desarrollo respiratorio y laríngeo a partir de diferentes elementos de la acústica del habla. También Huber et al. (1999) subrayaron la importancia que poseen las medidas acústicas a la hora de proveer de información sobre las diferencias que pueden darse en la articulación y en la configuración glotal entre hombres y mujeres.

Pero son Vorperian y Kent (2007:1511) los que resaltan el valor del componente acústico particularmente en los sonidos vocálicos por el papel decisivo que tiene el análisis acústico de la vocal para el estudio de otros elementos:

“[...] Vowels are important in their own right, but acoustic data on vowels also inform several other topics, including the acoustic cues for consonants (e. j. formant transitions

for consonant-vowel or vowel-consonant sequences), speaker normalization (which is usually based on formant frequencies), and prosodic patterns of speech (given that vowels carry a substantial part of prosodic information)”.

Sin embargo, entre los múltiples índices acústicos que pueden tenerse en cuenta la hora de estudiar los sonidos vocálicos de los bebés y niños pequeños (como la duración, la  $f_0$ , o la amplitud), los formantes parecen prevalecer sobre cualquier otro componente por la función que estos tienen en el reconocimiento de la vocal. Entre los primeros trabajos que manifestaron la importancia que dichos formantes poseían a la hora de establecer la cualidad vocálica encontramos los de Delattre et al. (1951), Cooper et al. (1952), Delattre et al. (1952), Stevens y House (1955), Flanagan (1955), Fant (1956, 1960, 1965, 1968, 1973, 2004), y Fry et al. (1962), y en ellos se demostró que los formantes eran los índices acústicos capaces de diferenciar unas vocales de otras. De esta forma y debido a que nuestro estudio pretende analizar las distintas vocales que producen el bebé y el niño pequeño durante los tres primeros años de vida, parece que el análisis de estas resonancias es lo más conveniente para una investigación de semejantes características. Por otro lado, tal y como se verá en la revisión bibliográfica sobre las capacidades perceptivas del bebé, estudios como los de Hoemeke y Diehl (1994) muestran que la  $f_0$  no es un factor determinante a la hora de diferenciar el timbre vocálico, como tampoco lo es el ancho de banda formántico según los experimentos de Fant (1960) y Flanagan (1972).

Dentro de esta primacía espectral y gracias en parte a los experimentos de los múltiples autores que acabamos de mencionar (y a sus pruebas con síntesis de habla) y en parte a otros estudios de percepción y de reconocimiento de habla (tanto natural como sintética) en los que se ha demostrado que el oído humano es capaz de identificar los sonidos vocálicos simplemente escuchando la información acústica recogida en la zona de bajas frecuencias, existe el acuerdo generalizado de que son los dos primeros formantes, esto es el F1 y el F2, los que mayor información aportan para la identificación de la vocal. Esta importancia del F1 y del F2 en el reconocimiento de la vocal es la que ha conducido a distintos investigadores interesados en el estudio acústico de las vocales tempranas a analizar los dos primeros formantes como índices acústicos más importantes y es también la que nos ha llevado a nosotros a seguir los mismos pasos (entre estos investigadores encontramos a estudiosos como Eguchi y Hirsh [1969], Buhr [1980], Kent y Murray [1982], Boysson-Bardies et al. [1989], Lee et al. [1999], Robb et al. [1997], Gilbert et al. [1997], Huber et al. [1999], Mattock et al. [2005], Lyakso et al. [2005], Van der Stelt et al. [2005], Van der Stelt et al. [2006], Rvachew et al. [2006], Chen [2006] e Ishizuka et al. [2007]).



No obstante, y a pesar de haber considerado por separado la importancia que cada uno de estos factores posee en la caracterización del habla del infante, y concretamente en el análisis vocálico, todos ellos están íntimamente relacionados. Esta correlación ha sido subrayada por múltiples autores en estudios variados y aunque la explicaremos con detalle en el capítulo cuarto (que trata sobre la caracterización articulatoria y acústica de las vocales del español) exponemos a continuación algunos ejemplos de dicha dependencia.

Para Stevens y House (1955) el espectro vocálico está basado en la longitud del tracto vocal y en la posición y grado de la constricción creada. Por otro lado, de los estudios de Peterson y Barney (1952) y Vorperian y Kent (2007) se desprende que la longitud del tracto vocal determina el patrón general de las frecuencias de los formantes.

En un trabajo sobre el desarrollo de la morfología del tracto vocal a partir de imágenes con resonancias magnéticas, Fitch y Giedd (1999:1511) atribuyeron a Chiba y Kajiyama (1941), Fant (1960), Lieberman y Blumstein (1988) y a Titze (1994) la opinión de que “[...] the shape of the vocal tract determines the articulatory possibilities, and thus possible formant patterns in speech”. También con ayuda de imágenes de resonancia magnética para investigar el desarrollo anatómico de las porciones faríngea y oral del tracto vocal, Vorperian et al. (2008) expusieron que los patrones del crecimiento anatómico del bebé y niño pequeño están muy relacionados con los cambios que presenta la acústica del habla del sujeto. Vorperian et al. (2005:348) después de medir el crecimiento desigual de las distintas estructuras del tracto vocal declararon que “[...] The current data on anatomic growth of the vocal tract structures provide an important first step towards exploring anatomic-acoustic interdependences, that is, relating anatomic growth patterns to developmental changes in speech acoustics”.

Por su parte, Perrier (2005:111) defiende que los movimientos articulatorios y la señal acústica están fuertemente relacionados al afirmar que “[...] In natural speech, articulatory movements and the acoustic signal are obviously strongly coupled, since articulatory movements are the source of the acoustics and determine its spectro-temporal characteristics”.

Finalmente y en opinión de Guenther (1994:2), la función primordial del sistema de producción de habla es la de la creación de la onda sonora “[...] the true job of the speech production mechanism is the creation of an appropriate set of acoustic signals to convey linguistic units from the speaker to listeners”.

De igual manera el estudio acústico de las producciones del bebé, y concretamente el análisis formántico, puede aportar al investigador un conocimiento bastante aproximado de las características fisiológicas y de las habilidades articulatorias que el bebé posee a lo largo del

desarrollo. Así lo confirman Lee et al. (1999) quienes citando a Goldstein (1980) en un artículo sobre los cambios espectrales y temporales de la acústica del habla que experimentan los niños según se van desarrollando, afirman que el conocimiento de dicha acústica podría facilitar el conocimiento de la comprensión del desarrollo de los órganos del habla y el del control motor del habla en los niños, y permitiría conocer de forma más precisa el desarrollo del tracto vocal.

También lo hacen autores como Ishizuka et al. (2007) cuando al estudiar los cambios producidos por el desarrollo en los picos espectrales de las vocales producidas por infantes japoneses manifiestan que el cambio en los formantes refleja el dominio de las distintas formas de articulación y la madurez de diferentes órganos articulatorios como el tamaño de la faringe, la longitud del tracto vocal completo o el descenso de la laringe. Asimismo y en un trabajo precisamente titulado “Evidence from acoustic studies”, Kent (1976) cree que los análisis acústicos son convenientes para comprobar algunas hipótesis sobre los cambios en la anatomía y el control motor. En el artículo de Vorperian y Kent (2007:1511) sobre el desarrollo del espacio acústico vocálico en niños, ambos sostienen que los datos acústicos del habla de los niños reflejan las diferentes transformaciones que sufre el tracto vocal:

“[...] Acoustic measures of children’s speech potentially reflect several developmental processes, including the growth of vocal tract structures (and sex differences in these growth patterns), changes in the relative geometry of the components of the vocal tract, maturation of speech motor control...”.

Por otro lado, Harshman et al. (1997) y Ladefoged et al. (1978) piensan que el F1 y el F2 pueden reflejar dos rasgos articulatorios: la altura vocálica y la posterioridad de los sonidos. En cuanto a otros índices acústicos y a su relación con la configuración anatómica del bebé y su control motor, Huber et al. (1999) opinan que las amplitudes de los formantes ofrecen información sobre los niveles de intensidad vocal y sobre el emparejamiento entre las cavidades supraglotales y subglotales. Estos también creen que el análisis acústico puede ayudarnos a establecer las primeras diferencias entre la anatomía del varón y de la mujer, concretamente en la configuración glotal de ambos.

En lo que respecta a la percepción de los sonidos, encontramos que esta se encuentra estrechamente relacionada con el espectro acústico porque como ya hemos mencionado, es mediante aquella que el bebé puede identificar las frecuencias de los formantes de las vocales que tendrá que producir. Según Fitch (2002) esta es una habilidad que compartimos con otros vertebrados que se valen de la percepción para reconocer los formantes y con ellos obtener información acerca del tamaño del vocalizador (algo útil en muchos contextos como en el caso de tener que averiguar el tamaño de otro animal en la oscuridad o cuando este está cubierto por

un espeso follaje). Añade Fitch (2002) que al igual que muchos aspectos de la producción vocal de los humanos son compartidos con otros mamíferos (como la musculatura, el control del tracto vocal, etc.) y debido a la larga historia evolutiva de los formantes entre los mamíferos, podríamos esperar que la percepción humana del habla estuviera establecida sobre esta base compartida y primitiva. De la misma manera Bailey y Snowling (2002) aseguran que la percepción del habla requiere una capacidad para determinar la forma espectral y discriminar la amplitud y la frecuencia fundamental de los sonidos de un lengua (para ellos algunas de estas capacidades están presentes desde el útero). De igual opinión son Lieberman y Blumstein (1988) quienes afirman que entre los múltiples mecanismos que envuelven el proceso de percepción del habla humana se encuentra el de extraer los formantes y la frecuencia fundamental de las señales del habla. En el mismo artículo se cita el trabajo de Halle y Stevens (1959) para añadir que además de esta capacidad, el humano debe filtrar las características del conducto de aire supra-laríngeo y formarse una estimación de su longitud para asignar un determinado patrón de frecuencia formántica a una meta fonética intencionada.

Asimismo es incuestionable la relación entre la percepción, el desarrollo anatómico y las habilidades de control motor. Dicha dependencia quedó establecida en las primeras décadas del siglo veinte cuando Bühler (1930) en un trabajo sobre el desarrollo mental del niño defendió la asociación entre la impresión auditiva y los movimientos que la producían ya que según el filósofo esto constituía la base para la imitación de los sonidos que escucha el niño en la que tiene que convertir lo que ha oído en movimientos vocales concretos.

A mediados de la década de los ochenta Bosma (1985) dejó patente este vínculo anatómico-perceptivo cuando expuso que dentro de la boca la capacidad para la discriminación perceptiva está centrada en la porción apical de la lengua y en los márgenes del labio. También Boysson-Bardies et al. (1989) señalaron esta dependencia cuando comentaron que los movimientos articulatorios son guiados por configuraciones auditivas. Años más tarde en un estudio sobre la voz *creaky* en infantes que aún no balbuceaban, Giesbrecht (2002) manifestó que el bebé se valía de la audición para guiar el desarrollo de la producción. En un artículo sobre el desarrollo fisiológico del control motor del habla Green et al. (2000) expusieron que en el incremento de dicho control motor intervienen varias fuerzas relacionadas entre sí: fuerzas intrínsecas (como la maduración cognitiva y sensorial-motora) y fuerzas extrínsecas (como la estimulación auditiva). Por su parte Menard (2005) expuso que las vocales más frecuentes en el repertorio del bebé podían ser explicadas por esquemas perceptivos. Destaca la coincidencia entre Kent et al. (1991), Guenther (1995) y Smith (2010) sobre el hecho de que el *feedback* auditivo juegue un papel crítico en el desarrollo de los procesos del control motor del habla, y el hecho de que la producción normal del habla no puede ser establecida sin la información

auditiva. De la misma manera piensa Pierrehumbert (2011:126) que al tratar de los modelos de procesamiento del habla que poseen distintos módulos para la percepción y la producción expuso que “[...] in early language acquisition, perception leads production” teniéndose que examinar las conexiones entre estas.

Por otro lado y según Kent (1984), aunque ambas capacidades (la de percepción y la de producción) están inicialmente separadas, estas comienzan a integrarse durante los primeros meses de vida. Para este autor, la integración de los dos sistemas también interactúa con la experiencia lingüística del niño pues como ya hemos mencionado son los sonidos que le rodean los que influyen en su patrón de vocalización. Nishitani et al. (2005:64) subrayaron la importancia de esta relación en el establecimiento de las bases del habla del infante cuando comentaron que

“[...] when the child listens to a new word, s/he automatically tries to imitate it, thereby forming a close temporal link between sensing (hearing) and acting (articulating)”.

La relación tridimensional (anatómica, acústica y perceptiva) es defendida por Rvachew et al. (2006) según los cuales son los factores anatómicos y de control motor así como las habilidades auditivas (y especifica tanto las periféricas como las centrales), los que pueden explicar los cambios acústicos observados en la producción de la vocal. Se trata pues de tres factores íntimamente relacionados en la producción del habla temprana. Es tal la dependencia entre la percepción, la acústica y la anatomía y el control motor, que según expone Remez (1995) las distintas corrientes encargadas de la percepción del habla aún no se ponen de acuerdo en si la meta de la percepción es recuperar los gestos articulatorios que crean la señal acústica (como defienden Liberman y Mattingly [1985] a partir de la teoría motora del habla según la cual el objeto de la percepción del habla no es sino recuperar los movimientos articulatorios y las órdenes motoras que el cerebro envía a los distintos órganos para producir el habla), o por el contrario lo es la categorización acústica de la señal emitida (tal y como defiende Ohala [1996] y según el cual la percepción del habla no puede ser motora ya que los niños y algunos animales son capaces de percibir sonidos sin poder producirlos, y los humanos pueden diferenciar sonidos que no son del habla y que por tanto no tienen una base articulatoria. Citado en O’Shaughnessy [1996]).

Por todo lo anterior, y a pesar de que nuestro estudio se centra en el componente acústico del habla, concretamente en el análisis del F1 y del F2 de los sonidos vocálicos aislados producidos por infantes entre los 0;4 y los 3;0 años de vida, hemos considerado necesaria la revisión bibliográfica de todos aquellos trabajos que han ahondado en la configuración

anatómica y el desarrollo del control motor del bebé, así como la de aquellos otros que han analizado las capacidades perceptivas del niño y las influencias que la lengua ambiente ejerce en las emisiones de este último. Dicha revisión, además de estar justificada por la ya explicada relación entre estos tres componentes, resulta pertinente por su contribución a la mejor comprensión de los resultados del presente estudio. De esta forma, si se da la circunstancia de que los sujetos del estudio son capaces de producir al poco de nacer determinadas vocales pero no otras, serán la perspectiva anatómica y de control motor así como la perceptiva las que arrojen las explicaciones oportunas. De la misma manera, estas perspectivas pueden ser útiles a la hora de explicar la mayor presencia de algunos sonidos en los primeros meses de vida del mismo modo que la influencia de la lengua ambiente podría aclarar la persistencia de ciertos sonidos frente a otros reflejando tal vez el patrón vocálico de la lengua materna. Queda así justificada la necesidad de una consideración anatómica, perceptiva y de la lengua ambiente en nuestro estudio acústico.

En cuanto a la distribución de los contenidos, creemos conveniente comenzar con una breve revisión de los primeros trabajos que se acercaron al estudio de la lengua del bebé, aun situándose estos fuera del campo de la Fonética y la Fonología, para pasar después al tratamiento de las investigaciones sobre el desarrollo anatómico y las habilidades perceptivas de los bebés, y dejar para el final aquellas publicaciones que profundizan como esta tesis en la descripción y el desarrollo del espacio acústico vocálico (capítulo segundo o del estado de la cuestión).

A continuación expondremos los objetivos que pretendemos conseguir con esta investigación y las hipótesis de trabajo desde las que partimos (capítulo tercero). No obstante, para una buena comprensión de los resultados es imprescindible que antes de llevar a cabo nuestro cometido aclaremos algunos conceptos acerca de la producción de las vocales del español desde el punto de vista fisiológico (cuestiones relacionadas con los procesos de respiración, fonación, articulación, etc.), y que ofrezcamos una breve explicación sobre las características acústicas que estas vocales presentan al transmitirse como ondas sonoras (definición de formantes, de sus correlatos articulatorios, de la función de transferencia del tracto vocal, etc.). Esta caracterización de los sonidos vocálicos del español desde el punto de vista articulatorio y acústico se verá en el capítulo cuarto.

Para lograr nuestros objetivos y ver si se cumplen las hipótesis de partida, es necesario seguir un buen método de trabajo que nos conduzca al producto final del estudio. Así pues en la metodología, que veremos en el capítulo quinto, se describirán tanto los procedimientos de

observación, como las herramientas de las que nos hemos valido y los distintos tipos de análisis acústicos y de percepción que fundamentan nuestro trabajo.

Los resultados de dichos análisis se revelarán en el capítulo sexto que recogerá por un lado los datos del examen acústico de las emisiones vocálicas de los bebés a lo largo de los tres primeros años de vida, y por otro los resultados de un test de percepción en el que cinco oyentes intentarán identificar el timbre vocálico de esas emisiones.

El capítulo séptimo continuará analizando los datos obtenidos pero desde un punto de vista estadístico.

La equiparación de todos estos resultados (acústicos, perceptivos y estadísticos) con los obtenidos por otros autores en trabajos similares tendrá lugar en el capítulo octavo, dedicado a la discusión.

Los capítulos anteriores nos permitirán llegar a unas conclusiones que expondremos en el capítulo noveno dejando para el último lugar un epílogo que recogerá la contribución de esta tesis en la literatura sobre el análisis acústico de los sonidos vocálicos de los bebés, así como las posibles aplicaciones de la misma.

## 2. ESTADO DE LA CUESTIÓN.

Los primeros trabajos interesados en la lengua del bebé y del niño pequeño abordaban su estudio desde otras disciplinas tales como las Ciencias Naturales, la Medicina, la Sociología o la Psicolingüística. Entre los biógrafos de bebés encontramos a Charles Darwin (1877), que aunque experto en las Ciencias Naturales publica en 1877 un artículo en la revista *Mind* titulado “A biographical sketch of an infant”, y a Hippolyte Taine (1877) quien siendo filósofo y uno de los fundadores de la Sociología, se interesa por la adquisición del lenguaje del niño y publica también en *Mind* un trabajo que lleva por nombre “The acquisition of language by childrens”. Uno año después la misma revista divulgó el trabajo de Pollock (1878) “An infant’s progress in language” y tres años más tarde acogió entre sus páginas la publicación del doctor Francis H. Champneys (1881) titulada “Notes on an infant”. En estos últimos así como en los de Humphreys (1880), Perez (1885), Tracy (1895), Moore (1895-1897), Mickens (1897-1898), Major (1906), Dearborn (1910), Bohn (1914), Hills (1914), Bateman (1915), Drummond (1916), Watson (1925), Kenyères (1927) y Foulke y Stinchfield (1929), el estudio del lenguaje se incluía en un estudio más general sobre el desarrollo de la mente en el niño. No obstante, los trabajos de Hills (1914), Bateman (1915), Watson (1925) y Tracy (1895) están más especializados en la adquisición de sonidos por parte del niño (una explicación más detallada de los trabajos citados aparece en el trabajo de Wellman et al. [1931]).

A finales de la década de los treinta ya es un fonetista de origen alemán, W. Leopold, quien publica uno de los primeros estudios sobre el bilingüismo en niños pequeños a partir del análisis y la observación de su propio hijo (*Speech Development in a Bilingual Child: A Linguist’s Record* [1939]). Este autor seguirá investigando el lenguaje del niño en los años siguientes (1947, 1949a, 1949b). En 1941 el lingüista ruso Roman Jakobson publica *Lenguaje infantil y afasia* donde explicará la evolución fónica del lenguaje infantil a partir de leyes fónicas universales (principios del menor esfuerzo, excepciones supuestas en el orden de aparición de los sonidos, etc.).

Posteriormente y desde el campo de la Psicolingüística, nos llegan estudios como el de Piaget (1955) con su *The Language and Thought of the Child*, o como el de Louis Bloom (1970) con *Language development: Form and function in emerging grammars* que indagan en la relación entre pensamiento y lenguaje en el niño.

Después de habernos acercado a los pilares de todo estudio longitudinal del habla infantil, continuamos con los trabajos que intentan aproximarnos a los condicionamientos fisiológicos con los que nace el bebé y a su desarrollo a lo largo del tiempo.

## **2.1. Configuración anatómica del bebé y del niño pequeño.**

### **2.1.1. Características del tracto vocal.**

Comenzamos con el estudio de Ramón y Cajal (1996) que tras describir la fisiología y las funciones laríngeas del infante cuando este es aún feto (observaciones realizadas sobre fetos de desarrollo normal estudiados en la propia consulta del hospital), concluyó que el primer tragado se observa a las 13 semanas, la función respiratoria a las 15 y la función fonatoria a las 18 (observó además que esta última función se produce durante el cierre de la glotis). Todo ello refleja según Ramón y Cajal una incipiente actividad cerebral ya que toda producción de un sonido es un proceso cortical. Sabemos ahora por tanto que la primera función fonatoria necesaria para la producción del sonido se da a las 18 semanas del periodo de gestación. Subrayamos el hecho de que se trate de fetos de desarrollo normal ya que alguna alteración en el desarrollo o el nacimiento prematuro de estos puede según Feld (1999) alterar las funciones laríngeas de los bebés, el juego vocal exploratorio de estos o incluso la energía con la que se transmiten las vocales (entre las causas que se recogen en su estudio con niños prematuros cuyo peso no excedía los 1500 g está la de la influencia de los tubos endotraqueales o las sondas nasográsticas que están ligados a su condición de prematuros).

Siguiendo las funciones de Ramón y Cajal y alegando que la producción del habla no puede concebirse sin tener en cuenta el contexto anatómico y fisiológico del niño (concretamente el grado de desarrollo de las estructuras, su forma y dimensión y su disposición espacial dentro de la región), Bosma (1975) confirmó que entre las funciones del infante se encuentran la posicional, la respiratoria-fonatoria y la de alimentación (funciones de las siguientes estructuras: faringe, laringe y boca del infante). Según el estudio estas funciones son generalmente independientes en su desarrollo pero interactúan particularmente en el infante. En relación a la función respiratoria y a la fonatoria, Lieberman (1986) cree que debe haber un periodo crítico durante el cual el niño debe ejercitar la presión subglotal de aire y el control laríngeo ya que si no lo hace durante este periodo, puede haber un control fonatorio deficiente.

Continuando el curso de desarrollo normal, el feto llega a su término y el bebé viene a este mundo con una configuración anatómica particular que presenta según Lieberman et al. (1972), Fletcher (1973), Bosma (1975), Sasaki et al. (1977), Goldstein (1979), Kent (1981), Kent y Murray (1982) y Kent (1992) las siguientes características: un tracto vocal más corto (que mide según Ménard y Boë [2004] unos 8 cm en el recién nacido frente a los 17 cm del adulto de sexo masculino y los 15 cm del adulto de sexo femenino); una pobre diferenciación de la cavidad oral y faríngea; una cavidad faríngea más corta; una masa de la lengua situada en la



parte anterior, no posterior (además, según Clement [2004], la lengua es relativamente grande y ocupa casi por completo la cavidad bucal); un ángulo gradual, más que un ángulo orientado hacia la derecha, del inclinado canal oro-faríngeo, es decir, un solo tubo resonador frente al ángulo desviado hacia la derecha que crea dos tubos resonadores en el adulto humano siendo esta característica lo que según MacNeilage (1998) y Carré et al. (1995) aumenta la capacidad para producir sonidos diferentes. Para Lieberman (1975, 1977, 1984) un ángulo hacia la derecha en el tracto vocal del adulto humano permite las discontinuidades necesarias para la producción de determinados sonidos como para la articulación de la [i] y de la [u]. El tracto vocal del infante y el del primate no humano son estructuralmente incapaces de producir dichas discontinuidades sugiriendo que dicha curvatura hacia la derecha es de vital importancia.

Por otro lado, el tracto vocal del bebé presenta una alta posición de la laringe, un acoplamiento entre esta y la naso-faringe y una aproximación de la velo-faringe y la epiglotis (situándose esta última a la altura de la primera vértebra cervical según Sasaki et al. [1960]). El hecho de que la laringe se encuentre en una posición más alta (justo detrás del velo del paladar) por un lado, y la aproximación del velo del paladar y la epiglotis por otro, permiten al infante respirar y tragar al mismo tiempo ya que al estar situada la laringe a la misma altura del velo en la cara posterior, el alimento pasa de la cavidad bucal al esófago sin que haya riesgo de que caiga en aquella.

Esto condiciona al bebé como un respirador y vocalizador nasal (Sasaki et al. [1960], Oller [1978] y Sapienza et al. [2004]). Según Crelin (1972), cuyas palabras recogen Robb et al. (1997), la presencia de la nasalización es el resultado de la anatomía del tracto vocal del bebé en la que la competencia velo-faríngea tiene que ser adquirida. De esta forma, la configuración anatómica del tracto vocal del infante presenta una laringe que está esencialmente *locked* o trabada a la naso-faringe, lo que obliga al infante a respirar y a vocalizar por la nariz. Clement (2004) añade a estos datos otros como el hecho de que los músculos extrínsecos del velo no puedan elevarlo de forma activa ya que este está localizado en una posición más hacia arriba que la de estos músculos (Fletcher [1973]). Esto hace que la corriente de aire vaya principalmente por la nariz y no por la boca en los recién nacidos explicando así la nasalidad de sus emisiones. Esta peculiar disposición de la laringe y del emparejamiento epiglótico asemejan el tracto vocal del bebé al de otros mamíferos no humanos como el del chimpancé según Sasaki et al. (1960), Sloan (1967), Wind (1970), Lieberman et al., (1971), Lieberman (1972), Fletcher (1973), Bosma (1975), Laitman y Crelin (1976), Laitman et al. (1977), Dubrul (1977), Ménard y Boë (2004) y Hopkins (2005).

Para Bosma (1986) y para Margriples y Laitman (1987), citados en el trabajo de Lieberman et al. (2001), entre la semana veintitrés y veinticinco en el útero el hioides y la laringe del feto están en una posición alta en relación a las vértebras cervicales, y la laringe se extiende desde la base occipital al nivel de la C3 o C4. En el neonato y según King (1952), Laitman y Crelin (1976), Senecail (1979) y Bosma (1986), el hioides se encuentra en una posición opuesta a la unión entre la C2 y la C3 pero a la edad de dos años tanto la laringe como el hioides han alcanzado prácticamente su posición adulta en relación a las vértebras cervicales (Carlsöö y Leijon [1960], Roche y Barkla [1965] y Westhorpe [1987]).

Además de esta particular disposición, las cuerdas vocales del recién nacido miden entre 2,5 y 3,0 mm según Sapienza et al. (2004), apareciendo las primeras diferencias de sexo entre los 10 y los 14 años de edad, frente a los aproximadamente 17-21 mm que miden en los adultos masculinos y los 11-15 mm que miden en los femeninos según Hirano et al. (1983) (los datos dados por Tizte [1989] son similares ya que las cuerdas vocales miden según este autor unos 4 mm en el recién nacido, unos 12 mm en la mujer y unos 18 mm en el hombre). También de manera distinta a la configuración adulta, las cuerdas vocales del bebé tienen una inclinación más hacia abajo desde la parte posterior a la anterior. Según Sapienza et al. (2004) otra distinción entre la laringe pediátrica y la adulta está en la proporción de membrana de las cuerdas vocales que supone menos del total de la longitud de la cuerda en el recién nacido y en el niño pequeño en comparación con la del adulto. Además la mucosa de la cuerda vocal es más delgada en aquellos que en este último. Por otro lado, la lámina propia de la laringe del infante es muy uniforme frente a la clara diferenciación en capas superficiales, intermedia y profunda de esta lámina en los adultos. Los recién nacidos carecen de estructura ligamentosa apareciendo esta de manera inmadura entre el año y los cuatro años de edad, edad a partir de la cual esta comienza a observarse según Hirano et al. (1983).

En cuanto a la posición de la laringe según Fried (1983), esta se encuentra entre la primera y la tercera vértebra cervical y en su desarrollo va pasando de una forma cónica a una más cilíndrica (Sapienza et al. [2004]). Por otro lado, los cartílagos laríngeos al estar poco calcificados son mucho más blandos que los de los adultos. En lo que respecta a la epiglotis y además de estar en contacto directo con el paladar blando, esta se encuentra muy aproximada a la base de la lengua. El hioides, al estar situado en una posición mucho más baja, se superpone al cartílago tiroides. Los aritenoides son mayores en el bebé y niño pequeño que en el adulto. Por todo ello las funciones respiratorias y de protección tienen un mayor papel en el infante que la función de fonación (Hirano et al. [1983]).

Así, aunque en el instante de nacer, tal y como reivindican Ruch y Zimbardo (1971) (ambos citados en Mowrer [1980]), el bebé es capaz de vocalizar, digerir comida y flexionar algunos músculos como los de los labios que permiten al recién nacido apretar el pezón de la madre para succionar la leche, tal vez muchos de estos movimientos sean puramente reflejos. Es lo que piensa Olmstead (1971) que ocurre con la acción de tragar ya que según este autor el proceso de tragado es una acción reflexiva que tiene lugar cuando los músculos laríngeos fuerzan el paso del líquido a través del esófago. Esta es también la opinión de Lieberman (1968) que cree que el niño recién nacido solo puede emitir sonidos cuando se combina la actividad de vocalizar con la de respirar o con la de la ingestión de comida, y que más allá de estos comportamientos ningún sonido del habla podrá ser producido.

Todas estas características hacen según Kent y Bauer (1985) y Kent y Vorperian (2006) que el tracto vocal del bebé no sea simplemente una miniatura del tracto vocal del adulto sino que sea además, funcionalmente diferente. De esta manera los primeros afirman que (1985:492):

“[...] Although data on vocal tract development are still sparse the data that do exist indicate that the infant’s vocal tract differs substantially from the vocal tract of the adult. The infant’s vocal tract is not only shorter, thereby making for absolute differences in dimensions and configuration, but it differs also in the relative or proportionate size of its subdivisions”.

Esta es también la opinión de Polgar y Weng (1979) quienes en una revisión sobre la fisiología de la respiración de los niños demostraron que las vías respiratorias de estos no son solo más pequeñas que las del adulto sino que presentan además diferencias funcionales. A la misma conclusión llegaron Stathopoulos y Sapienza (1993:2531) tras estudiar el movimiento de las cuerdas vocales, la presión traqueal, la corriente de aire glotal, el volumen pulmonar, el desplazamiento de la caja torácica y el desplazamiento abdominal entre otros, en niños de 4;0 y 8;0 años y en adultos de 20 años: “[...] Laryngeal and respiratory results indicate that speech production differences between the children and adults are based both on size and function”. Y entre las diferencias de tipo funcional o cualitativo (no solo cuantitativo o de tamaño) entre el tracto vocal del adulto y el del infante, Stathopoulos y Sapienza (1993:2540) encontraron las siguientes:

“[...] On the other hand, several qualitative/functional differences were uncovered. For example, the amount of time the vocal folds were closed, the percentage of vital capacity used during speech, the rib cage excursion range, and the abdominal displacement at utterance termination were functionally different for children compared to the adults”.

La misma postura es defendida por Callan et al. (2000:722) quienes afirmaron que “[...] The infant vocal tract is not just a miniaturized version of the adult vocal tract. There are significant differences in the shape, size, innervation, and functional aspects of the structures involved with speech production in the infant and adult”. Asimismo y con respecto a los cambios ontogénicos en la laringe, Eckel et al. (2000:501) subrayaron que “[...] The infant larynx is not just a miniature of the adult organ. It shows differences in its position relative to the vertebral column, in the composition of cartilages and soft tissues, and in environmental adaptation”.

Esta disposición fisiológica se irá abandonando mediante un proceso conocido como *reestructuración anatómica* que se produce entre los cuatro y los seis meses de vida y que trae consigo la formación de un ángulo hacia la derecha en la región nasofaríngea y por tanto la diferenciación en dos tubos resonadores, la separación del velo y la epiglotis, el descenso de la laringe, la epiglotis, el hioides y el descenso de la lengua para formar la pared anterior de la faringe (Bosma [1975a], [1975b], [1976], [1985], Crelin [1972], [1976], Laitman y Crelin [1976], Lieberman [1977], [1984], Sasaki et al. [1977], Kent [1981], Fried et al. [1982], Westhorpe [1987] y Kent y Vorperian [1995]).

El desaparejamiento de la laringe y la naso-faringe entre los cuatro y los seis meses de edad permite al bebé respirar por la nariz y disminuir el número de vocalizaciones nasales (Sasaki et al. [1977] y Crelin [1987]). Clement (2004) añade que entre los cuatro y los seis meses de vida el crecimiento de la mandíbula hacia abajo dota a la lengua de un mayor espacio para moverse y da a la corriente de aire la oportunidad de ir hacia la cavidad oral. A partir del sexto mes de vida según Clement (2004) el bebé puede mover la mandíbula libremente hacia arriba y hacia abajo y si el movimiento es repetido, puede resultar en el balbuceo. Todo ello interviene en el alargamiento del tracto vocal que se produce a lo largo de estos meses según Vorperian et al. (1999).

Con respecto a esta reestructuración anatómica y en cuanto a la laringe, según esta aumenta su longitud el tamaño de los aritenoides y el de las cuerdas ariepiglóticas desciende (Sapienza et al. [2004]). Eckel et al. (2000) opinan que la laringe refleja varias adaptaciones según el niño se desarrolla y que entre ellas se encuentran su descenso, los ajustes de tensión, longitud y forma de las cuerdas vocales y el aumento progresivo de la porción cartilaginosa frente a la prominencia de la parte membranosa en las primeras semanas de vida. En cuanto al descenso de la laringe, Nishimura (2003) asegura que el descendimiento ocurre en dos pasos: en el primero el tiroides comienza a descender en relación con el hioides y en el segundo se produce el descenso del hioides en el cuello. Este descenso de la laringe hacia una posición

adulta conlleva el riesgo de ahogo para el infante. Por otro lado, Clement (2004) piensa que el descenso de la laringe puede estar asociado a cambios en el desarrollo del mecanismo de tragar y cree que de acuerdo con Nishimura et al. (2003) este proceso contribuye físicamente a una mayor independencia entre los procesos de fonación y articulación para la vocalización.

Unos años antes del ya mencionado trabajo de Clement (2004), Lieberman et al. (2001:124) defendieron que es el proceso de deglución el que más contribuye, frente al proceso vocalizador, no solo al descenso de la laringe sino al del hioides (durante el proceso de tragado son los músculos que hay por encima del hioides los que elevan el complejo hioides-faríngeo abriendo el esfínter superior del esófago y alejando a la laringe de la ruta por la que ha de discurrir el bolo alimenticio). Así lo exponen Lieberman et al. (2001:124) en el apartado de la discusión después de haber analizado las radiografías tomadas a veintiocho niños a la edad de 0;1, 0;3 y 0;9 años y a partir de ahí cada doce meses hasta la edad adulta: “[...] the rate of hyo-laryngeal descent appears to be constrained by the need to maintain functional equivalence among the hard tissues structures that interact with soft tissues structures involved in deglutition”. Otra conclusión a la que llegan es la de que el crecimiento de la faringe está también constreñido por la acción muscular relacionada con la deglución.

Bloom (1998) piensa que el hecho de que la lengua disponga ahora de más espacio para moverse debido a este cambio anatómico contribuye a un mayor control de los articuladores resultando en diferentes tipos de vocalización (aumentando por tanto el número de vocales y el rango de las frecuencias de los formantes). Según Bosma (1963) con la transición postnatal la faringe, cuya función hasta ahora era la de llevar la comida o el aire hacia el esófago y participar en la respiración, toma un papel activo. El desarrollo del conducto de aire faríngeo se estabiliza progresivamente unos meses después del nacimiento de forma que su diámetro varía menos con la respiración (este proceso de estabilización de la faringe es paralelo y está muy relacionado con el desarrollo de la cabeza y la postura del cuello). Por otra parte, según Polger y Weng (1979), Hirano et al. (1983), Kent (1984) y Kent y Miolo (1995), todos citados en el trabajo de Ménard y Boë (2004), las modificaciones de los tejidos óseos, de la lengua, de los cartílagos y las dimensiones del conducto continúan acercándose a las del adulto durante el primer decenio de vida aunque el tracto vocal del bebé seguirá desarrollándose hasta alcanzar el tamaño adulto en un rango de edad de entre los 7 y los 18 años según Kent y Vorperian (1995) y Lieberman et al. (2001).

Además de estar de acuerdo con esta última observación, Fitch y Giedd (1999) creen que es durante el periodo de la infancia y pubertad (considerando la plena pubertad entre los 10;3 y los 14;5 años) cuando se produce un aumento significativo de la longitud del tracto vocal

siendo el velo y la faringe los que más aumentan durante la temprana madurez (entre la pubertad y la madurez la faringe aumenta un 25% mientras que la longitud del tracto vocal en este mismo periodo solo crece el 5%). En cuanto al dimorfismo sexual, el estudio de Fitch y Giedd (1999) reveló que aunque no hay diferencia en la longitud del tracto vocal de los niños y las niñas antes de la pubertad (hasta los 10;3 años), pequeñas disimilitudes aparecen en la plena adolescencia (de 10;3 a 14;5) y se acentúan pasada la pubertad haciéndose más significativas. Los mismos estudiosos demostraron con ayuda de un análisis de covarianza (*Analysis of Co-Variance* o ANCOVA) realizado a ciento veintinueve niños y a adultos de entre 2;0 y 25 años y para el que se tomaron como variables la altura del cuerpo, la edad y el sexo, que existe una correlación entre el tamaño del cuerpo y la varianza de la longitud del tracto vocal siendo la altura del cuerpo la que explica una mayor varianza. Dicha correlación anatómica entre la longitud del tracto vocal y el tamaño del cuerpo sugiere que los formantes podrían ser un índice acústico del tamaño del cuerpo teniendo esto implicaciones para la fonética forense o el reconocimiento automático de la voz.

Esta reestructuración anatómica queda también explicada en los trabajos de Vorperian et al. (1999), Vorperian et al. (2005) y Vorperian et al. (2008) en los que con ayuda de imágenes de resonancia magnética (un método muy apropiado ya que permite visualizar los tejidos blandos en la región oro-faríngea así como las estructuras cartilagosas y los huesos), se lleva a cabo una descripción minuciosa del desigual crecimiento de las distintas estructuras del tracto vocal que forman parte de esta reestructuración, información que contribuye a entender cómo surge y se desarrolla el habla. En estos estudios las estructuras del tracto vocal quedan divididas según la zona y la orientación en la que se encuentran. De este modo encontramos las estructuras anteriores u orales situadas en un plano horizontal (VT-H) que comprende la distancia desde una línea tangencial de los labios a la pared posterior de la faringe y que incluyen el grosor del maxilar y del labio de la mandíbula, y la longitud del paladar duro y la de la mandíbula; y las estructuras posteriores (VT-V) repartidas en la distancia vertical desde la glotis hasta el plano del paladar y que incluyen la longitud del paladar blando, la longitud de la faringe, el hioides y el descenso de la laringe (Vorperian et al. [2005]). No obstante, cada trabajo midió una serie de variables y aunque los tres compartieron la descripción de la longitud del tracto vocal (definido este como la distancia curvilínea que comienza en la glotis hasta la intersección con una línea dibujada tangencialmente a los labios), presentaron características particulares.

En el estudio de Vorperian et al. (1999: 201-202) se tomaron imágenes de resonancia magnética a dos sujetos entre el nacimiento y los 3;9 años en las que se midieron las siguientes estructuras: longitud de la cabeza: “distancia máxima desde la glabella hasta el opisthocranium”;

altura de la cara: “la suma de lo que pesa la parte alta y baja de la cara o altura de la cara”; longitud del paladar duro: “la distancia curvilínea entre el paladar duro hasta el comienzo del paladar blando”; longitud del paladar blando: “la distancia curvilínea desde el borde posterior del paladar duro hasta el borde inferior de la úvula”; anchura del arco maxilar: “la anchura de arco en la cúspide de la superficie bucal”; longitud del arco maxilar: “la distancia de la media sutura palatal del paladar duro, medida desde el borde anterior de la cúspide de los incisivos hacia una línea dibujada entre el tubo maxilar”; anchura mandibular: “la distancia entre los bordes temporales de los dos procesos condilares”; longitud y profundidad mandibular: “la distancia horizontal y vertical en el plano medio-sagital”; longitud del tracto vocal: “la distancia curvilínea a lo largo del comienzo del tracto vocal en el borde superior del cartílago tiroides hasta la intersección con una línea dibujada tangencialmente hasta los labios”; longitud de la lengua: “la distancia curvilínea a lo largo del contorno superior dorsal de la lengua, desde la punta de la lengua, hasta la depresión o grieta de la misma”; área de la lengua: “toda la región de la lengua, anterior y posterior, superior e inferior”; nivel del hueso hioides: “hueso hoides”; nivel laríngeo: “distancia vertical de una línea dibujada desde el tiroides hasta la ANS-PNS (*Anterior-Posterior Nasal Spine*) línea de referencia”; longitud de la oro-naso faringe: “la distancia curvilínea a lo largo de la pared posterior de la faringe sobre el paladar blando, extendiéndose desde los canales posteriores de la nariz hasta el cartílago del tiroides”; grosor del labio maxilar: “la distancia horizontal desde la zona subnasal hasta la espina nasal anterior”; grosor del labio mandibular: “la distancia horizontal y anteroposterior desde el supramentale hasta la línea de los tejidos blandos”; longitud del labio maxilar: “la distancia vertical desde la zona subnasal hasta el borde inferior del labio maxilar”; longitud del labio mandibular: “la distancia vertical desde el borde superior del labio mandibular hasta el surco inferior”; área del labio maxilar: “la región señalada anteriormente por la línea de los tejidos blandos, en la parte posterior por la línea de los tejidos duros y la línea del contorno de los labios, en la parte inferior por el borde del labio superior y en la parte superior por una línea que conecta las marcas o hitos subnasales y la espina nasal anterior”; y área del labio mandibular: “la región rodeada en la parte anterior y superior por el labio inferior y que está rodeada en la parte posterior por la línea de los tejidos duros y por la parte inferior por una línea horizontal que conecta el *sulcus* inferior con la línea de los tejidos duros”.

Los resultados hallados en este estudio fueron: un crecimiento coordinado de determinadas estructuras como la de la oro-nasofaringe con la altura superior e inferior de la cara; los periodos en los que coincide el crecimiento de bastantes estructuras (aunque no todas las hacen a la misma velocidad y alcanzan el tamaño adulto a la misma edad) son de 1-4 meses, de 13-15 meses y de 36-45 meses; las estructuras posteriores (como el descendimiento de la

lengua y el de la laringe) parecen contribuir más al alargamiento del tracto vocal que las estructuras anteriores (como el grosor de los labios o el crecimiento del paladar duro). Por otro lado, las resonancias magnéticas mostraron que el tracto vocal aumentó entre 1,5 y 2 centímetros en los dos primeros años de vida, y otro centímetro entre las edades de 2;1 y 3;0 años.

El análisis Vorperian et al. (2005:340-341) incluyó doce sujetos adultos (seis de sexo masculino y seis de sexo femenino), y veinticinco sujetos pediátricos (dieciséis niños y nueve niñas) desde las 2 semanas a los 6;9 años de edad, en los que se tomaron las siguientes medidas: longitud del paladar duro: “la distancia curvilínea a lo largo del contorno del paladar duro desde el punto anterior del diente incisivo al comienzo del paladar blando”; longitud del paladar blando: “la distancia curvilínea desde el borde posterior del paladar duro al borde inferior de la úvula”; longitud y profundidad de la mandíbula: “distancia horizontal y vertical en el plano mediosagital desde la protuberancia mental a la proyección ortogonal del proceso condilar en el plano mediosagital”; longitud de la lengua: “distancia curvilínea a lo largo del contorno superior dorsal de la lengua desde la punta de la lengua al *valleculae*”; nivel del hueso del hioides o nivel de la lengua: “distancia vertical desde el PNS (*Posterior Nasal Spine*) al nivel del punto anterior inferior del hueso del hioides; nivel laríngeo: “distancia vertical de una línea trazada desde el corte del cartílago tiroideo al PNS”; longitud naso-oro-faríngea: “distancia curvilínea a lo largo de la pared faríngea posterior encima del paladar blando extendiéndose desde la *naris* posterior (apertura externa de la cavidad nasal) al final del conducto de aire superior. Una línea trazada horizontalmente desde el borde superior del hueso del hioides a la pared faríngea posterior fue tomada para dividir la línea entre la nasofaringe y la orofaringe”; grosor del labio maxilar: “la distancia anteroposterior desde el subnasal (sn) a la ANS (*Anterior Nasal Spine*); espesor del labio mandibular: “la distancia horizontal anteroposterior desde el *sulcus* inferior (Si) o supra-mental a la línea del tejido duro”.

Después de medir todas estas variables los resultados del estudio de Vorperian et al. (2005) fueron los siguientes:

- Longitud del tracto vocal: entre las 2 semanas y el primer año de vida la longitud del tracto vocal en los niños está entre los 7 cm (45% del tamaño adulto) y los 8,2 cm (55%) aproximadamente, y en las niñas entre los 7,5 cm y los casi 9 cm, es decir entre el 50% y el 58% del tamaño adulto. Entre el 1;0 y los 2;0 años de edad dicha longitud está en los infantes varones entre los 9 cm y los 10 cm, suponiendo por tanto un 58%-65% del tamaño adulto, mientras que en las niñas está entre los 9 cm y los 10,5 cm, constituyendo casi el 60% y el



70% del tamaño adulto. Entre los 2;0 y los 3;0 años el tracto vocal mide en los niños entre 9 cm y algo más de 10,5 cm, lo que supone un 60%-70% del tamaño adulto, mientras que en las niñas las medidas se encuentran entre los 9 cm y los 9,5 cm, lo que se traduce en un 60-63% del tamaño adulto (entre los 3;0 y los 4;0 años de vida, el tracto vocal de la niña alcanza los 11,5 cm correspondiéndose prácticamente con el tamaño del tracto vocal del niño, que mide unos 11,7 cm). A los seis años de edad ambos alcanzan un 75% de la longitud del tamaño adulto. Por lo que respecta al desarrollo de cada una de sus partes, la mayoría de las estructuras tienen un crecimiento más rápido durante los dieciocho primeros meses de vida, adquiriendo a esta edad entre el 55% y el 80% del tamaño adulto, y a los seis años entre el 65-80% del tamaño adulto.

- Longitud de la naso-oro-faringe: entre las 2 semanas y el 1;0 año la longitud se mueve en los niños entre los 4,3 cm (los seis primeros meses) y los 6 cm (al final del periodo), es decir, supone entre el 45% y el 60% de la longitud del adulto; en las niñas la longitud es de casi 5 cm al principio y de casi 7 cm al final del periodo, lo que se traduce en un 50% y 70% del tamaño adulto. Entre el 1;0 y los 2;0 años la longitud de la naso-oro-faringe se mueve en los niños entre los casi 6 cm (5,8 cm) y los 7,5 cm, alcanzando el 55% y el 75% del tamaño adulto, mientras que en las niñas se encuentra entre los 6 cm y los casi 7 cm (6,8 cm), suponiendo entre el 60% y casi el 70% del tamaño adulto. Entre los 2;0 y los 3;0 años de edad la longitud se sitúa en los niños entre los 5,7 cm y los 7,8 cm, suponiendo un porcentaje de entre el 60% y hasta algo más del 75% del tamaño adulto, mientras que dicha longitud en los infantes femeninos está entre los 7,5 cm y los 8 cm, lo que supone un 75% y casi 80% del adulto.
- Longitud de la lengua: entre las 2 semanas y el 1;0 año la longitud de la lengua en los niños varones se sitúa entre los 5,5 cm y aproximadamente los 5,8 cm, lo que se traduce en un 45% y el casi 60% del tamaño adulto; en las niñas está entre los 6 cm (50%) y los 6,3 cm (53%) aproximadamente. Del 1;0 a los 2;0 años la lengua mide en los sujetos masculinos entre 6,8 cm, algo más del 55%, y 8 cm, un 65 %, siendo estas medidas muy similares a las de la longitud de la lengua en las niñas: 6,5 cm (algo más del 50%) y casi 8 cm (un 65% del tamaño adulto). Entre los 2;0 y los 3;0 años de edad la longitud de la lengua es en los niños de entre 6,8 cm, algo más del 55%, y 8 cm, casi un 65% (se mantiene la longitud de los 12 meses anteriores); en las niñas la lengua ocupa unos 7cm,

casi un 60% del tamaño adulto. Entre los 3;0 y los 4;0 años la longitud de la lengua en la niña alcanza a la del niño y en ambos está alrededor de los 11 cm, un 75% del tamaño adulto.

- El descenso de la laringe: entre las 2 semanas y el 1;0 año la longitud de la laringe en los infantes varones es de entre 3 cm y 4 cm, representando entre el 35% y el casi 45% del tamaño adulto; en las niñas, está entre los 3,3 cm y los 3,7 cm aproximadamente, correspondiéndose con el 40% y el 45% del tamaño adulto. Del 1;0 año a los 2;0 años la longitud de la laringe en el niño está alrededor de los 4,5 cm y 5,5 cm, esto es, entre el 55% y el 65% del adulto, mientras que en la niña está entre los 4 cm y los 5,2 cm, suponiendo el 45% y el 60% del adulto. De los 2;0 a los 3;0 años la laringe en el niño mide entre 4 cm, el 55% del tamaño adulto, y algo más de 5,5 cm, el 65% de dicho tamaño; en la niña y a la misma edad, las medidas son de entre 4 cm y 5 cm, representando el 45% y casi el 60% del tamaño adulto.
- Longitud del paladar duro: entre las 2 semanas y el 1;0 año la longitud del paladar duro está en los niños entre los 2,5 cm y los 3,2 aproximadamente, representando entre algo más del 55% y algo más del 70% del tamaño adulto; en las niñas dicha longitud es de entre 2,8 cm, algo menos del 55%, y 3,4 cm, algo menos del 89 %. Del 1;0 año a los 2;0 años el paladar duro en los infantes varones mide entre 3,3 cm, suponiendo algo más del 75% del tamaño adulto, y 3,8 cm, correspondiéndose con algo menos del 90% de dicho tamaño; en las niñas, el paladar duro está entre los 3,2 cm y los casi 4 cm, lo que se traduce en algo más del 70% y el 90% del adulto. Entre los 2;0 y los 3;0 años el paladar duro está en los niños entre los 3,6 cm, algo más del 80%, y los 4,4 cm, es decir el 95% del adulto. En las niñas, el paladar está entre los 3,7 cm y los 4 cm, lo que se traduce en un 45% y 90% del tamaño adulto. Entre los 3;0 y los 4;0 años el paladar duro de la niña y el del niño se van igualando hasta llegar a los 6;0 años, edad en que alcanzan el 100% del tamaño adulto.
- Longitud del paladar blando: entre las 2 semanas y el 1;0 año la longitud del paladar blando se sitúa en los niños entre los 2 cm (1,8 aprox.) y los 3,2 cm, es decir entre el 47% y el 80% del tamaño adulto; en las niñas está entre los 2 cm y los 3 cm, es decir entre el 50% y el 75% del adulto. Entre el 1;0 año y los 2;0 años, el paladar blando mide en los niños entre 2,4 cm (algo menos del 60%) y 3,3 cm (algo menos del 85%); en las niñas este se sitúa entre los 3,2 cm y los

casi 4 cm, correspondiéndose con algo más del 70% y algo menos del 90% del tamaño adulto. Entre los 2;0 y los 3;0 años la longitud del paladar blando en los niños es de entre 3,6 cm y 4,4 cm aproximadamente, suponiendo algo más del 80% y algo menos del 90% del tamaño adulto; en las niñas, el paladar blando mide entre 3,7 cm y 4,1 cm, es decir supone entre el 85 % y algo más del 90 % del tamaño adulto.

- Longitud de la combinación del paladar blando y del paladar duro: entre las 2 semanas y el 1;0 año la longitud de esta medida en los niños se sitúa entre los 4,6 cm y los 6,4 cm, es decir representa entre el 55% y el 75% del tamaño adulto; en las niñas, la combinación de ambas oscila entre los 5,2 cm aproximadamente y los 6 cm, moviéndose entre algo más del 60% y el 70% del tamaño adulto. Del 1;0 año a los 2;0 años la combinación de ambas longitudes está en los niños entre los 5,8 cm y los 6,8 cm aproximadamente, suponiendo entre el 70% y el 80% del tamaño adulto; en las niñas, la unión de ambas longitudes se encuentra a esta edad entre los 5,5 cm y los 6,7 cm aproximadamente, es decir representa entre el 65% y algo menos del 80% del tamaño adulto. De los 2;0 a los 3;0 años, la longitud combinada en los niños está entre los 5,8 cm y los 6,9 cm aproximadamente, lo que equivale al 70% y a algo más del 80% del tamaño adulto; en las niñas, dicha longitud se encuentra entre los 6 cm y 6,7 cm aproximadamente, suponiendo algo más del 70% y algo más del 80% del tamaño adulto.

Finalmente, en el trabajo de Vorperian et al. (2008:1669) se hacen trescientas siete resonancias magnéticas a trescientos veintisiete varones y doscientas setenta y ocho mujeres desde el nacimiento hasta los 19 años para medir nueve variables. En el plano horizontal (VT-H) encontramos las siguientes: la LTh o grosor de los labios que es la “distancia en el nivel de *stomion* entre dos líneas, la primera es una tangencial dibujada desde la parte anterior y la segunda hacia la parte posterior o parte bucal del maxilar y los labios”; la ACL como “la distancia horizontal de una línea dibujada desde el incisivo lingual a la intersección con la línea vertical dibujada desde la glotis al plano del paladar A a B; la OPhW o ancho de la oro-faringe: “VT-H menos LTh menos ACL” y la VT-O que es “VT-H menos LTh”. En el plano vertical están presentes: la PCL o longitud de la cavidad posterior siendo esta la “distancia vertical de una línea dibujada desde la glotis a la intersección con el final de la longitud de la cavidad anterior u oral”; la ACL o longitud de la cavidad anterior como la “distancia vertical desde la glotis hasta el comienzo de la cavidad oral; y la NphL o longitud nasofaríngea que constituye

“el tracto vocal vertical menos PCL” (consideremos importante incluir las siglas y la definición de cada una de estas variables para una mejor comprensión de las conclusiones).

Por lo que respecta a los resultados obtenidos tras realizar las mediciones de las nueve variables en el trabajo de Vorperian et al. (2008), y debido por un lado a que estos coincidieron con los de Vorperian et al. (2005) en cuanto al crecimiento de las estructuras anteriores y posteriores, y por otro a que dichos resultados están orientados a las posibles diferencias de sexo, hemos optado por presentar directamente las conclusiones y su relación con los resultados obtenidos en otros trabajos:

La tendencia de crecimiento para todas las variables es en general no lineal a lo largo de los primeros 18 años de vida tanto para hombres como para mujeres (Vorperian et al. [2005] y Vorperian et al. [2008]). Ménard et al. (2004) también hablan de un crecimiento no uniforme del tracto vocal debido a la presencia de periodos de crecimiento rápido frente a otros de crecimiento más lento de las distintas estructuras del tracto vocal. La hipótesis del crecimiento no lineal es también defendida por Goldstein (1980), Gollin (1981), Fitch y Giedd (1999) y Callan et al. (2000) que frente a la hipótesis del crecimiento uniforme de Mol (1963) (a cuyo estudio no hemos podido acceder) defienden un crecimiento desigual del trato vocal.

Hay un periodo de mayor crecimiento (denominado *growthspurt*) para la mayoría de las estructuras analizadas entre el nacimiento y los dieciocho meses de edad (Vorperian et al. [2005]).

El tipo de crecimiento de las diferentes estructuras del tracto vocal varía en relación a su región u orientación. De este modo, las estructuras anteriores tienen en general una curva de crecimiento neuronal, esto es, presentan un crecimiento muy rápido desde el nacimiento hasta un tiempo variable en la infancia temprana, momento en el que la estructura es dos tercios de su tamaño adulto, y a este le sigue un periodo de crecimiento muy lento hasta la madurez (un ejemplo de estas estructuras sería el paladar duro). Por otro lado, las estructuras posteriores o verticales poseen una curva de crecimiento somático, es decir, tienen una fase de crecimiento muy rápido después del nacimiento y durante la infancia –aquí *infancy*–, pero al término de este periodo de crecimiento veloz la estructura apenas es un cuarto del tamaño adulto. A este periodo de rápido crecimiento le sigue un crecimiento regular y lento hasta la temprana y media infancia –aquí *childhood* o niñez–, después otro rápido crecimiento durante la pubertad, y acaba con un periodo de crecimiento lento hasta la madurez (unos ejemplos de estas estructuras serían el paladar blando y el descenso de la laringe). En cuanto a las estructuras que están situadas en ambos planos (el vertical y el horizontal) como la longitud del tracto vocal y la longitud de la lengua, estas siguen una curva de crecimiento combinado, es decir, neuronal y somático a la vez

y es este crecimiento intermedio el que provoca que dichas estructuras sean las que mayor crecimiento muestren a cualquier edad (Vorperian et al. [2005] y Vorperian et al. [2008]).

Por tanto, algunas estructuras se acercan al tamaño adulto antes que otras: la longitud del paladar y el grosor del labio maxilar alcanzan el 80% del tamaño adulto a los 1;6 años, mientras que otras como la profundidad de la mandíbula y la longitud de la faringe alcanzan el 80% del tamaño adulto a los 6;0 años de edad (Vorperian et al. [2005] y Vorperian et al. [2008]).

El crecimiento relacional de las estructuras anteriores y posteriores (el crecimiento de las diferentes estructuras en relación al de la longitud del tracto vocal) durante los primeros dieciocho meses de vida fue prácticamente el mismo aunque después de los dieciocho meses el crecimiento relacional de las estructuras posteriores excedió al de las estructuras anteriores (Vorperian et al. [2005]).

Las variables situadas en la región oral o plano horizontal alcanzan el tamaño adulto antes que las variables de la región faríngea o plano vertical (Lieberman y McCarthy [1999], Lieberman et al. [2001], Vorperian et al. [2005] y Vorperian et al. [2008]). Según Vorperian et al. (2008), estas diferencias pueden explicarse además de por el tipo de crecimiento, porque sus orígenes embriológicos son diferentes.

A pesar de que en el primer trabajo Vorperian et al. (2005) manifestaron la ausencia de dimorfismo sexual tanto en las estructuras como en la longitud del tracto vocal a la edad de 4;0 años, tres años más tarde Vorperian et al. (2008) encontraron que algunas estructuras como la VT-H presentaban una diferencia sexual. Esta contradicción puede deberse como ya indicaban Vorperian et al. (2005) a la necesidad de tener en cuenta otras variables y no solo la de la longitud del tracto vocal sobre la que trata su estudio (en dicho trabajo se cita el trabajo de King [1952] en el que se indica que a la edad de 1;0 año los niños poseen una región faríngea mayor que las niñas de la misma edad, y el de Lieberman [2011] según el cual el crecimiento de la anchura oro-faríngea -distancia desde la pared faríngea posterior al margen posterior de la cavidad oral- es ligeramente mayor en los varones que en las mujeres entre las edades de 1;75 y 4;75 años). Vorperian et al. (2008) encontraron que las diferencias de sexo fueron significativas para todas las variables excepto para ACL (la distancia horizontal desde el comienzo del paladar duro a la intersección con una línea vertical dibujada desde la glotis al plano del paladar) indicando además que algunas variables mostraron el dimorfismo sexual antes que otras (como la ya mencionada VT-H o la OPhW a la edad de 4;0 o 5;0 años frente a otras como la VTL, el PCL y la NPhL a las edades aproximadas de 9;0, 10;0 y entre 10;0 y 11;0 años

respectivamente). Según Vorperian et al. (2008) este dimorfismo sexual es razonable si se tiene en cuenta que hay una fuerte correlación entre la longitud del tracto vocal y el tamaño del cuerpo (Bennet [1981] y Fitch y Giedd [1999]), y si se considera la velocidad y el tipo de crecimiento ya que algunas estructuras alcanzan el tamaño adulto mucho antes que otras dependiendo de su orientación y de si se trata de un sujeto de sexo masculino o femenino (por ejemplo, el OPhW es un 61% neuronal para hombres y un 75% somático para mujeres de modo que al tener una curva de crecimiento neuronal crecerá más rápido en los hombres que en las mujeres).

Las diferencias de sexo significativas que se ven sobre todo a partir de los 12;0 años (Vorperian et al. [2005] y Vorperian et al. [2008]) confirman los resultados de dimorfismo sexual de la longitud del tracto vocal que tienen lugar a partir de los 11;0 años según Fitch y Giedd (1999) y Lieberman et al. (2001).

También para Kent y Vorperian (2006) las diferentes estructuras del tracto vocal poseen distintas velocidades de crecimiento de forma que algunas configuraciones muy similares a las del adulto pueden adquirirse durante la temprana infancia y otras durante la adolescencia. Así por ejemplo el cráneo sigue una curva de crecimiento neuronal por la que el bebé en la infancia temprana alcanza dos tercios de su tamaño adulto, mientras que la cara sigue una curva de crecimiento somático según la cual tras un primer periodo de rápido crecimiento en el mismo periodo, esta tan solo alcanza un cuarto de su tamaño adulto (los autores citan en este punto a Scammon [1930]). A partir de aquí defienden que las estructuras que discurren paralelas al cráneo tendrán un crecimiento neuronal (como el paladar duro y la longitud de la mandíbula) frente a las estructuras paralelas a la cara que presentarán una curva de crecimiento somático (como el descenso de la laringe y el alargamiento de la faringe). Finalmente y de igual manera que algunas estructuras del tracto vocal alcanzan la madurez antes que otras, también algunos gestos articulatorios del bebé según Nittrouer (1995) se aproximan a un estatus adulto antes que otros.

Podríamos completar el desarrollo anatómico del bebé durante los primeros meses y años de vida con estudios como los de Hunter (1771), citado en Enlow y Harris (1964), que trata sobre la evolución de las diferentes estructuras anatómicas y según el cual la mandíbula crece en una dirección predominantemente posterior, es decir, hacia la base del cráneo a lo largo de la infancia. Además, tal y como exponen Enlow y Harris (1964) cada área de la mandíbula experimenta una serie de cambios en relación a su localización (similar a la que experimentan las distintas estructuras del tracto vocal) y según va aumentando el tamaño de la mandíbula, sus distintas partes ocupan localizaciones relativas.

Pero para la producción del sonido es necesario que los pulmones y la caja torácica del bebé presenten unas características similares a las del adulto, y que dicho bebé posea cierto control de la presión subglotal y cierto control laríngeo. Entre los cambios que sufre la anatomía encargada de la respiración y que se citan en el trabajo de Connaghan et al. (2004) encontramos el descenso de la elasticidad de la caja torácica (Bryan y Wohl [1986] y Sharp et al. [1970]), y la orientación más diagonal de las costillas con el cambio del infante a la postura vertical (Hershenson [1992]) que conlleva un aumento de la capacidad de los músculos para expandir la caja torácica. Todo ello supone una mayor contribución de la caja torácica a la respiración y una cada vez menor participación del abdomen. Esta incrementada participación de la caja torácica en la respiración permite que el bebé vaya adquiriendo un control de la respiración cada vez más fino, control esencial según Connaghan et al. (2004) para el desarrollo de diferentes comportamientos fonatorios tales como el control de la frecuencia fundamental o de la intensidad. Los cambios en las dimensiones torácicas fueron caracterizados por Howatt y DeMuth (1965) como no lineales debido a la diferente velocidad con que la anchura del tórax, la longitud del esternón y el diámetro anterior-posterior alcanzan la madurez. No obstante y a pesar de estas características peculiares, según Giesbrecht (2002) y Langlois et al. (1980), a los 0;3 años de edad el bebé posee un tórax similar al del adulto que le permite producir una fonación más prolongada en el tiempo.

En cuanto a la respiración, pueden distinguirse en el bebé dos tipos de respiraciones dependiendo de si se trata de la respiración que este mantiene en posición de descanso o de la respiración que requiere para la vocalización. El estudio de Moore et al. (2001) con bebés de 1;3 años reveló una mayor dependencia entre la caja torácica y el abdomen en la respiración en posición de descanso que la existente durante la respiración necesaria para el habla. La misma conclusión se desprende de los trabajos de Boliek et al. (1996) con bebés de entre 5 semanas y 1;0 año, y del de Boliek et al. (1997) con niños de entre 2;0 y 3;0 años. En estos estudios además se concluyó que la respiración es altamente variable según nos encontremos ante el llanto, el gimoteo o la emisión de sílabas, que el sexo del infante no conlleva diferencias en la respiración, y que las duraciones espiratorias aumentaron con la edad así como también el número de sílabas producidas por cada grupo de respiración.

Por lo que respecta al control de la laringe, Lieberman (1986) cree que debe haber un periodo crítico durante el que el niño ejercite la presión subglotal del aire y el control laríngeo ya que si no lo hiciera durante este periodo, podría haber un control fonatorio deficiente. Por su parte Clement (2004) añade que a esta edad, y gracias a la reestructuración de la caja torácica, los infantes pueden producir una mayor presión subglotal y tienen un mejor control de la duración y de la frecuencia fundamental de las emisiones gracias a la regulación de la presión de

aire. Este último autor cita a Bloom (1998) para exponer que dos meses más tarde (entre el cuarto y el sexto mes de vida) el bebé tiene un mayor dominio de los músculos intercostales lo que contribuye a un mayor control de los distintos articuladores.

### **2.1.2. Desarrollo del control motor.**

Además de conocer la configuración fisiológica con la que nace el bebé y su desarrollo o reestructuración anatómica, es importante conocer el control que tiene el niño de los órganos articulatorios necesarios para producir el habla. Smith (2010) parte de Sherrington (1906) para destacar el papel que las *motorneuron pools* (neuronas motoras que inervan los músculos) tienen en el control de la actividad de los músculos y por ende, en el control motor del habla. La importancia de la función neuronal está patente ya en el embrión por su intervención en el desarrollo de los pulmones, del sistema vascular, de los músculos y del esqueleto según Hopkins (2005). De hecho al séptimo mes, un buen número de células nerviosas ha alcanzado ya la madurez adulta aunque los patrones de la conexión entre células necesarios para adquirir las habilidades motoras no se han desarrollado aún en esta edad. Siguiendo con Smith (2010) estas órdenes a los músculos deben ser modificadas para adquirir las metas lingüísticas, y los circuitos neuronales que proveen el input a estas neuronas deben cambiar y adaptarse a los cambios anatómicos que sufre el bebé.

Las neuronas motoras tienen también para Bernstein (1967) un papel decisivo en el control articulatorio puesto que según el autor el control de los labios y el de la mandíbula por parte del bebé parece estar basado en sinergias funcionales que no son sino colecciones de neuronas motoras musculares que contraen los músculos y controlan el movimiento muscular necesario para que los niños produzcan el sonido. Son estas y no los músculos en sí, las que al activarse permitirían al infante controlar el movimiento de los labios y el de la mandíbula para crear una emisión. Añade Bernstein (1967) que según maduran los niños, estas sinergias funcionales se hacen más estables, resultando por tanto en una variabilidad menor para adquirir las metas de movimiento.

Son por tanto las redes neuronales inmaduras según Smith (2010) las responsables por ejemplo de una velocidad de habla más baja que refleja un sistema de control inmaduro característico del niño el cual requiere más tiempo para planificar y ejecutar el movimiento, y depende más del *feedback*. También en Kent y Forner (1980), Smith y Goffman (2004) y en Hopkins (2005) se refleja la menor velocidad de habla de los infantes y la menor precisión de estos en los patrones espaciales y temporales de la articulación.



En el mismo trabajo, Smith subraya que existen diferencias en el control de los articuladores y que algunos de estos alcanzan el control adulto antes que otros. Algo similar apuntaba el estudio de Nittrouer (1995) en el que a partir del análisis formántico del F1 y del F2 de vocales insertas en sílabas emitidas por diez niños de 3;0, 5;0 y 7;0 años y por adultos, se observó que el movimiento de cierre y abertura del tracto vocal era similar al del adulto a la edad de 3;0 años pero que sin embargo los gestos de la lengua estaban constreñidos por el contexto fonético hasta la edad de al menos 7;0 años. Se observó además que muchos movimientos fueron producidos de manera más lenta por los niños que por los adultos y con una variabilidad temporal mayor. Se concluye por tanto que la aparición de patrones gestuales maduros en el habla de los niños no es uniforme.

Lo mismo opinaron Green et al. (2002) cuando después de grabar el movimiento del labio superior e inferior y el de la mandíbula en niños de 1;0, 2;0, 6;0 y 7;0 años y en adultos de 35 a 37 años en la producción de las palabras *baba*, *papa* y *mama*, observaron que el control de la mandíbula de los niños fue más parecido al del adulto que el de los labios. Ello llevó a Green et al. (2002:67) a concluir que “[...] the mandible provides the fundamental patterns of early articulation that form the foundation for the learning of other, more specialized articulations”.

De la misma manera, dos años antes en el trabajo de Green et al. (2000) al analizar la coordinación entre la mandíbula y los labios en las producciones de sílabas con consonantes bilabiales en niños de 1;0, 2;0, y 6;0 años y en adultos jóvenes, se observó que en los niños de 1;0 año la mandíbula era el principal articulador frente los labios superior e inferior. Entre los 2;0 y los 6;0 años, la contribución del labio inferior aumentó significativamente (también lo hizo la del labio superior aunque no de manera significativa) al tiempo que descendió el papel de la mandíbula. Este dominio mandibular puede deberse a múltiples razones entre las que se encuentran la de que ya como feto el infante de ocho semanas puede abrir la mandíbula según Humphrey (1964), mientras que la musculatura del labio está aún desarrollándose a esta edad (Gasser [1967]). Este hecho quizás se deba según Smith (2010) a que los labios tienen grados superiores de libertad de movimiento en comparación con la mandíbula, o a que según Honda et al. (1995), Honda (1996) y Gerard et al. (2003) las contracciones de los músculos que producen los movimientos de los labios (y los de la lengua) son mucho más complejas y conllevan más dimensiones para el habla. Por otro lado, también se advirtió en el trabajo de Green et al. (2000) que a la edad de 1;0 año el movimiento del labio inferior se asimiló al patrón de movimiento establecido por la mandíbula, que a la edad de 2;0 años comenzaron a establecerse asociaciones rígidas entre el labio inferior y superior y que con el paso del tiempo se fueron diferenciando los comportamientos. Esto llevó a Green et al. (2000) a aceptar la existencia de tres etapas en el control motor del habla, sugeridas anteriormente por autores

como Fentress (1984), Kent (1992) y Lenneberg (1967), que son las siguientes: integración de estructuras y movimientos, diferenciación de los mismos y refinamiento hacia el estado adulto.

En la misma línea se encuentra el trabajo de McAllister (2009) quien subraya la dificultad a la que se enfrenta el bebé a la hora de coordinar la multiplicidad de estructuras que intervienen en la articulación (glotis, velo, lengua, labios, mandíbula, etc.) y los múltiples grados de libertad de sus movimientos. Según la autora una forma de simplificar la combinación de estructuras y movimientos consistiría en el movimiento en bloque de todos los articuladores como si de una sola unidad se tratara. Esta configuración si bien limitaría la precisión de los gestos del habla del niño, mejoraría el control motor ya que al mover todos los articuladores a la vez como si se tratara de una sola unidad se consigue un mayor control motor porque el niño realiza un solo movimiento –en bloque eso sí- pero con un control motor mayor. Para realizar este movimiento en bloque el niño se apoyaría en la mandíbula cuyo protagonismo podría deberse a razones tales como la relativa poca libertad de movimiento que trae consigo su función de bisagra y la mayor estabilidad que esta presenta debido a su mayor tamaño. Por el contrario, al ser un músculo hidrostato (sin apoyo de un esqueleto interno) la lengua es más difícil de controlar (Kent [1992]) y actúa como un participante relativamente pasivo en las primeras vocalizaciones de los bebés. Además, según Hardcastle et al. (1987), los niños poseen un control no diferenciado de las distintas partes de la lengua de forma que tienen dificultades para elevar la punta de la lengua sin elevar simultáneamente el cuerpo de esta. En cuanto a los labios, y aunque hay evidencia de algún control neuromuscular de estos por parte del bebé nada más nacer como demuestra el comportamiento de la succión (Whitaker [1973]), no está claro hasta qué punto la succión está controlada de manera voluntaria por el bebé en las primeras semanas de vida.

Según McAllister (2009) esta combinación de factores empuja al niño a emparejar los labios y la lengua con la mandíbula (para formar una unidad) siendo esta última el apoyo fundamental a la hora de realizar cualquier movimiento (esto se conoce en fonología como *principio de dominio mandibular*). Dicho dominio mandibular ha sido corroborado en otras investigaciones tales como la de Hodge (1989) que al analizar las vocalizaciones de niños de 3;0, 5;0 y 9;0 años concluyó que los gestos en la dimensión vertical, esto es hacia arriba y hacia abajo, predominan en una etapa temprana del desarrollo sobre los gestos de la dimensión anteroposterior, o en la teoría del *Frame then Content* de MacNeilage y Davis (1990) en la que se defiende que la mandíbula es el único articulador activo en las primeras vocalizaciones de los bebés, sobre todo al comienzo del balbuceo canónico, adscribiéndose el movimiento de los labios superior e inferior a la oscilación de esta.

En el mismo año y en una obra que trata sobre las bases articulatorias del balbuceo, Davis y MacNeilage (1990) reafirman que los segmentos específicos del contenido (*Content*) se desarrollan y rellenan los huecos del marco o estructura (*Frame*) según aumenta la independencia en el niño del control sobre otros articuladores del habla. Este predominio mandibular y la falta de control lingual llevan según Davis y MacNeilage (1990) a que puedan predecirse unos patrones de aparición de consonante-vocal tales como: las consonantes alveolares irán acompañadas de vocales anteriores, las consonantes labiales de vocales centrales y las consonantes velares de vocales posteriores.

Sin embargo, no todos los estudios de transcripción corroboraron esta aparición conjunta de consonante-vocal. En el trabajo de Chan (2001) se expusieron algunas de estas investigaciones y entre ellas encontramos la de Vihman (1992) en la que tras estudiar las primeras cincuenta palabras de veintitrés niños de cuatro lenguas diferentes, solo se observó el vínculo entre consonantes velares y vocales posteriores mientras que las asociaciones consonante alveolar-vocal anterior y consonante labial-vocal central fueron rechazadas. El mismo resultado obtuvieron Tyler y Lansdale (1996) tras transcribir las producciones de niños de entre 1;6 y 2;0 años de edad puesto que en sus datos solo fue evidente la asociación entre consonantes velares y vocales posteriores no cumpliéndose las otras dos. Para Oiler y Steffans (1993) las predicciones de Davis y MacNeilage (1990) se cumplieron de manera más estricta cuando la edad de los niños fue de entre 0;10 y 1;0 año disminuyendo entre los 1;4 y los 2;0 años. Algo similar ocurrió en el propio estudio de Chan (2001) donde después de transcribir el patrón CV en veintiséis niños que adquirirían cantonés y que fueron divididos en dos grupos de edad (el grupo A incluía a niños de entre 0;10 y 1;3 años y el grupo B a infantes de entre 1;4 y 1;10 años) se concluyó que la teoría de F/C es capaz de predecir los patrones de CV predominantes en el balbuceo pero no cuando el balbuceo concurre con las palabras (esto es, en una edad más avanzada). Finalmente, Levelt (1996) propuso patrones de aparición conjunta totalmente diferentes tras transcribir las producciones de niños de entre 1;0 y 2;11 años: las consonantes coronales acompañaron a vocales coronales, las labiales a vocales labiales y las dorsales a vocales dorsales (una explicación más detallada de esto último puede encontrarse en Chan [2001]). Estos patrones contrarían la base motora de la teoría del *Frame then Content*.

No obstante, los procesos de control motor del habla no están completos a la edad de diez años, según Watkin y Fromm (1984), y su equiparación con los del adulto se da en la adolescencia tardía (Walsh y Smith [2002] y Cheng et al. [2007]). En opinión de Smith y Goffman (2004:246) dichos procesos continúan madurando durante la pubertad y pasados los 16 años ya que “[...] the late maturation of articulatory motor processes is driven primarily by interaction of speech motor systems with co-developing cognitive-linguistic systems, which are

also not adult-like until this late period of development". Para Blakemore y Choudhury (2006) el cerebro continúa desarrollándose pasados los 20 años.

Continuando con el control motor del bebé y del niño pequeño, Kent y Vorperian (2006) especifican que todo control motor posee una representación neuronal o un mapa motor y que dicho mapa motor para el habla representa las capacidades de control sobre los aproximadamente cien músculos envueltos en la producción de habla. Añaden que las neuronas motoras pueden tener una función importante para el desarrollo en tanto que pueden establecer las relaciones iniciales de acción-percepción (producción-percepción del sonido).

Pero, ¿en qué áreas concretas se produce el control neural de los órganos del habla? Clement (2004) recoge la afirmación de Ploog (1979) de que las vocalizaciones del niño antes del balbuceo están relacionadas con la maduración del subcórteix involucrada en las vocalizaciones de otros mamíferos (tras diferentes estudios con monos se ha comprobado que su output vocal está controlado por el subcórteix y el medio córtex). Por tanto es según Clement (2004) la maduración de las áreas subcorticales, mediocorticales y del Área Suplementaria Motora (SMA o *Supplementary Motor Area*), que es una parte del córtex medio, la responsable de las vocalizaciones previas al balbuceo y de las del balbuceo mientras que el área de Broca comienza a ejercer su influencia en la producción temprana de la palabra.

Lo que está claro según Kent (1981, 1984) es que el control del sistema nervioso central sobre las vocalizaciones crece progresivamente con la maduración. A los seis meses el bebé ha adquirido muchas conexiones neuronales corticales y subcorticales necesarias para adquirir el control motor del habla (Kent [1981]). En trabajo de 1984, Kent recuerda la sugerencia de Ploog (1979) de que la escasa actividad neuronal en las etapas del llanto y de sonidos reflexivos aumenta en la etapa del susurro o arrullo entre el primer y el tercer mes de vida, ya que las estructuras límbico-cinguladas comienzan a ejercer algún control sobre la vocalización, y se intensifica entre los seis y los nueve meses con la llegada del balbuceo. También con el balbuceo Whitaker (1976) cree que aumenta el control neuronal que el niño tiene sobre la musculatura del tracto vocal.

Según el niño va adquiriendo un cierto control articulatorio la pregunta que se plantea es ¿disminuye con dicho control la variabilidad de movimiento muestra a muestra? En el estudio de Sharkey y Folkins (1985) se recogen de manera aclaradora los distintos puntos de vista sobre esta cuestión. Para muchos autores como Bruner (1973) una vez que el niño aprende a realizar una tarea articulatoria, la variabilidad de los movimientos muestra a muestra disminuye. Otros como Bernstein (1967) no ven la variabilidad como un error sino como un fenómeno natural que se da también en el habla adulta ya que rara vez repetimos una tarea con

gestos articulatorios exactos. Para este autor, según aumenta la habilidad articulatoria del niño se incrementan también las posibilidades exploratorias de realizar una misma tarea con diferentes gestos que son al final funcionalmente equivalentes. Para otros como Kelso y Norman (1978), Kerr y Booth (1978), y Moxley (1979) la variabilidad se explica como un juego exploratorio que facilita al bebé el aprendizaje de la lengua. No obstante, son muchos los autores que piensan que la variabilidad de los gestos articulatorios es una cuestión de madurez y que por tanto esta desciende progresivamente con la edad y entre ellos encontramos a Eguchi y Hirsh (1969), DeSimoni (1974a), (1974b), (1974c), Tingley y Allen (1975), Kent y Forner (1979) y Flege (1982). Frente a estos, el estudio de Sharkey y Folkins (1985) obtuvo resultados opuestos. En dicho trabajo se midió el desplazamiento del labio inferior y el de la mandíbula en niños de 4;0, 7;0 y 10;0 años al producir las sílabas [bæ] y [mæ] incrustadas en la frase *That's a \_\_\_d dog* y los resultados mostraron que la variabilidad del desplazamiento del labio decreció en los grupos de 4;0 y 7;0 años, no en los de 10;0 años, frente la variabilidad del desplazamiento de la mandíbula que no se redujo en ningún grupo de edad. Se demostró por tanto que la variabilidad no sigue un proceso de descenso gradual reflejando una disminución del error o un aumento de la precisión articulatoria.

### **2.1.3. Control motor y modelos artificiales de reproducción de habla.**

Las modernas técnicas de radiografía y resonancia magnética utilizadas en la observación y medición directa del bebé y del niño pequeño no solo han permitido la obtención de datos sobre la configuración anatómica que presenta el bebé al nacer y su evolución con el paso del tiempo o sobre el control motor que este posee de sus gestos articulatorios, sino que esta tecnología también ha hecho posible la creación de redes neuronales y tractos vocales artificiales que han posibilitado el estudio de la fisiología del bebé y su desarrollo contribuyendo así a una mejor caracterización de las particularidades anatómicas y acústicas de este.

Entre los primeros estudios que intentaron simular el tracto vocal del infante encontramos los trabajos de Nordström (1975, 1977) en los que se reprodujo el tracto vocal de una mujer y de un bebé a partir del tracto vocal del hombre. Aunque sus datos no coincidieron con los de Fant (1960) por múltiples razones reconocidas por los propios autores (entre las que destacan la utilización de diagramas formánticos inadecuados, la asignación a un conjunto específico de frecuencias de un símbolo concreto del IPA que puede ser algo difícil, o el asumir que las posturas articulatorias son exactamente igual en hombres y mujeres), una conclusión relevante ya destacada anteriormente por otros autores (Polgar y Weng [1979], Stathopoulos y Sapienza [1993] y Callan et al. [2000]) es que las diferencias anatómicas solo explican parte de

las diferencias formánticas entre hombres/mujeres y niños pudiéndose inferir que además de la anatomía existen diferencias funcionales entre estos grupos de edad.

Por lo que respecta a los trabajos que han reconstruido artificialmente la fisiología del bebé, encontramos el estudio de Guenther (1994) en el que se utilizó el modelo DIVA (*Directions Into Velocities of Articulators*) definido en Guenther y Perkell (2004) como una red neuronal de control que utiliza la etapa del balbuceo para aprender las transformaciones sensoriales-motoras necesarias para controlar el tracto vocal y producir así sonidos, sílabas y palabras. El funcionamiento de este modelo incluye la retroalimentación acústica para adquirir las metas oro-sensoriales del habla, la forma de los objetivos del tracto vocal (determinadas configuraciones que este puede presentar), el enlace entre las metas del tracto vocal y los movimientos del articulador, y una simulación de la coarticulación. Según Guenther (1994), de los tres niveles representados en este modelo (y que participan en la adquisición del habla), a saber, el auditivo, el oro-sensorial y el articulatorio, el nivel oro-sensorial (definido por Perkell [1980] como la información propioceptiva y sensorial sobre el estado del tracto vocal que recibe el niño) es fundamental debido a que es mediante este que se especifican los movimientos de los articuladores para alcanzar los objetivos del habla y por tanto la adquisición del lenguaje. También para Smith (2010) un modelo del desarrollo motor del habla debe incluir relaciones entre varios sistemas neuronales y además de las relaciones oro-sensoriales, acústicas y motoras, este autor propone que al menos para la producción de la primera palabra deben existir relaciones bidireccionales entre las representaciones motoras y lingüísticas de los sonidos.

Diez años después Guenther y Perkell (2004) intentaron explicar a través de este modelo el funcionamiento neuronal del bebé que subyace a la articulación. Según los autores, la producción de un sonido comienza con la activación de una célula del mapa de sonidos del habla, identificadas por autores como Rizzolatti et al. (1996a, 1996b y 1997) con las neuronas espejo, que se corresponde con el sonido a ser producido. Una vez activada dicha célula, señales de la corteza pre-motora viajan a las áreas corticales auditivas y sensoriales a través de sinapsis que codifican las metas para dicho sonido. Además del comportamiento neuronal, es conveniente subrayar de dicho trabajo por un lado la importancia que según Guenther y Perkell (2004) posee el control que ejerce el *feedback* sobre las primeras producciones vocálicas del bebé hasta que tras muchos intentos de producción del sonido el *feedforward* (especie de control sobre la orden motora) se pone a punto y permite la producción correcta del sonido, y por otro, el hecho de que la forma del tracto vocal puede variar en la repetición de un mismo sonido sin que esto conlleve un cambio en la estabilidad de la señal acústica. Según Guenther (1994, 1995) y Guenther et al. (1998), esta variación de las posiciones de los articuladores permite al hablante producir el sonido con un mínimo esfuerzo articulatorio.

También con ayuda del modelo DIVA y al igual que el trabajo de Guenther y Perkell (2004), el estudio de Callan et al. (2000:722) destaca la importancia del *feedback* como señal de entrenamiento en la adquisición del habla, y subraya el hecho de que el niño aprenda a usar diferentes configuraciones articulatorias para adquirir las mismas metas acústicas debido a los constantes cambios que experimenta el tracto vocal en los primeros meses de vida:

“[...] The changes in the structures associated with speech production that occur during development will all have varying degrees of influence as to their acoustic consequences. These changes may require that the child constantly learn to use different articulatory configurations to achieve the same speech goals”.

El artículo de Lieberman (1980), apoyándose en los datos de Ladefoged et al. (1972), Nearey (1976) y Russell (1928), ya advertía del hecho de que en la producción del habla de hablantes adultos (y por extensión puede aplicarse al habla de los bebés y niños pequeños) no existía un conjunto de formas invariantes del tracto vocal supralaríngeo usadas por todos los sujetos de una lengua particular para producir las vocales de esa lengua, sino que los hablantes se valían de diferentes maniobras para generar dichas vocales.

Otro ejemplo lo encontramos en el estudio de Ménard y Boé (2004) en el que se utilizó el *Variable Linear Articulatory Model* (VLAM), ya usado por Boë (1999), que reprodujo las siguientes variables: protusión de los labios, altura de los labios, posición de la mandíbula, posición hacia delante y hacia atrás del cuerpo de la lengua, altura del dorso de la lengua, elevación del ápice de la lengua y altura de la laringe. El estudio, que trató de la producción de las vocales orales francesas /i y u e ø o ε œ/, creó espacios vocálicos máximos (es decir, todos aquellos sonidos que el niño sería capaz de producir si usara el conjunto completo de las órdenes articulatorias del VLAM) en distintas fases del crecimiento: un recién nacido, un niño de 4;0 años, uno de 10;0 años, un adolescente de 16 y un adulto de 21 años. En cuanto a las conclusiones obtenidas, los investigadores hallaron que por un lado una misma meta acústica puede alcanzarse con varias configuraciones articulatorias (esto es, distintos gestos articulatorios para la producción de un mismo sonido), y que por otro lado es la falta de control motor y no la disposición anatómica que presenta el niño al nacer y en los años posteriores la que determina qué sonidos puede producir el bebé y cuáles no (2004:169):

“[...] Par la simulation des espaces vocaliques maximaux, nous avons montré que, si l'on suppose des capacités de controle moteur identiques de l'enfance à l'âge adulte, les différences morphologiques entre les conduits vocaux typiques des deux stades de croissance ne font pas obstacle à la production d'un espace vocalique étendu”.

Según Ménard y Boé (2004) es el control motor y no las restricciones anatómicas el que determina que unos sonidos sean más frecuentes que otros (2004:169). “[...] Les vocants et closants les plus fréquemment observés sont donc attribuables au contrôle moteur immature, et non à des contraintes imposées par la morphologie du conduit vocal”. Esta afirmación ya aparece en el trabajo de Ménard (2003) sobre las estrategias articulatorias durante el crecimiento del tracto vocal y se reitera en Ménard et al. (2004). Algo parecido encontramos en el artículo de Mowrer (1980:102) cuando hablando de las limitaciones fisiológicas del bebé nos dice que incluso siendo estas salvables “[...] neurological and motor control centers would not be sufficiently developed to permit adequate control of the muscles necessary for articulating a wide variety of sounds”.

La posibilidad de que una misma señal acústica pueda alcanzarse con varias configuraciones anatómicas ha sido establecida no solo en análisis con modelos artificiales para niños sino también en estudios de adultos como los que se citan en el trabajo de Boë et al. (1992) y entre los que se encuentran los de Schroeder (1967), Mermelstein (1967), Heinz (1967), Kadokawa y Suzuki (1969), Atal (1970), Sondhi y Gopinath (1971), Wakita (1973), Nakajima et al. (1975), Wakita y Gray (1975), Atal et al. (1978), Sondhi (1979), Wakita (1979), Charpentier (1984) y Lin (1990). Estos trabajos enfatizan la articulación compensatoria y la no unicidad del tracto vocal no solo en bebés y niños pequeños sino también en adultos, frente a la caracterización tradicional de punto único de articulación. Sin embargo, en el propio trabajo de Boë et al. (1992) los resultados arrojaron una conclusión diferente: la precisión del control de cada parámetro (localización de la constricción, área transversal de la constricción, área de la abertura del labio, etc.) que interviene en la articulación depende de la vocal que se considere (ponen como ejemplo las vocales [u] y [o] para las que no es posible ninguna compensación debido a que requieren un control muy preciso de la abertura del labio).

Otra investigación que parte del VLAM de Boë (1999), como modelo articulatorio-acústico que integra el crecimiento no uniforme del tracto vocal, es la de Serkhane et al. (2007) y en ella se intentó averiguar la exploración por parte del niño de su tracto vocal (correspondencia entre movimiento y sonidos concreto) y el grado de control motor que el infante poseía para producir los sonidos de su lengua a los 0;4 y 0;7 años de edad. Entre los motivos que llevaron a Serkhane et al. (2007) a utilizar el VLAM como medida directa del control articulatorio del bebé, encontramos por un lado la falta de medidas directas de la articulación de los niños antes de los 2;0 años de vida, a excepción según los investigadores de los estudios que han analizado los movimientos de los labios y de la mandíbula mediante grabaciones de vídeo como los de Munhall y Jones (1998) o mediante marcadores reflectantes como los de Green et al. (2000, 2002), y de aquellos otros que han registrado el movimiento de



los músculos de la mandíbula con ayuda de la electromiografía (Moore y Ruark [1996]), y la deficiencia de los estudios de transcripción y de análisis acústico para el estudio de las habilidades motoras tempranas del habla por otro. En el análisis se compararon las producciones reales de los niños de 0;4 y 0;7 años que adquirirían el inglés americano como lengua materna con aquellas creadas con el VLAM cuyas órdenes o grados de libertad articulatoria coincidían con las especificadas en el trabajo de Ménard y Boë (2004) (esto es, altura de la laringe, adelantamiento y retracción del dorso de la lengua, elevación del ápice de esta, etc.). Los resultados fueron los siguientes: a los 0;4 años de edad las vocalizaciones de los corpus reales no cubrían todos los sonidos o los sonidos posibles que se hubieran producido si el niño hubiera utilizado el rango completo de configuraciones articulatorias de acuerdo al VLAM. Las vocales a los 0;4 años se agruparon alrededor de una posición neutral definida por valores cero de todos los articuladores, y a los 0;7 años el niño exploró la parte baja del espacio acústico. A partir de estos datos la conclusión fue la siguiente (2007:336):

“[...] Considering that the remodeling of the vocal tract between 4 and 7 months is minor and cannot explain satisfactorily the vowel space expansion, the acoustic and articulatory framing results show that changes in pre-linguistic vocalization inventories over time cannot be explained by the growth of the vocal tract only, but that they also express changes in the articulatory exploitation of the vocal apparatus which may point to other aspects of the developmental process”.

Entre los trabajos que han preferido utilizar la técnica electromiográfica para especificar los movimientos de los articuladores encontramos el ya citado de Moore y Ruark (1996) en el que se midió el movimiento mandibular para comprobar la independencia de este de otros comportamientos no lingüísticos, y otros trabajos citados en Maeda y Honda (1994) como los de Honda (1992), Honda et al. (1992) y Kusakawa et al. (1993) en los que mediante la electromiografía se especificaron los movimientos de los músculos de la lengua y se estableció la influencia de estos en el F1 y F2 de las vocales (aunque en estos trabajos no se experimentó con bebés y niños pequeños sino con adultos, consideramos conveniente revisarlos debido a la concreta descripción que se hace en ellos de los músculos causantes del movimiento de la lengua en la articulación vocálica).

Según estos últimos estudios, las vocales pueden trazarse a partir de pares de músculos de la lengua que funcionan de manera antagónica. Un eje lo constituiría la actividad del músculo hyoglossus (HG) menos la actividad del músculo genioglossus posterior (GGp). El otro eje se correspondería con la actividad del styloglossus (SG) menos la actividad del genioglossus anterior (GGa). Estos músculos funcionarían de la siguiente manera según Honda (1991): el HG y el GGp contribuirían al movimiento posterior-bajo/anterior-alto del cuerpo de la lengua,

mientras que el SG y el GGa contribuirían al movimiento posterior-alto/anterior-bajo de la misma. A estos cabría añadir la actividad del OOs (Orbicularis Oris Superior), que contribuye al redondeamiento y protusión de los labios. En el análisis de Maeda y Honda (1994), además de la actividad de estos músculos que conformaría el parámetro de posición del cuerpo de la lengua (tp), encontramos otros parámetros como la posición más baja de la mandíbula (jp), la forma del cuerpo de la lengua (ts) y la posición de la punta de la lengua (tt) para constituir un modelo descriptivo consistente en tres partes: un tubo de abertura del labio, el perfil de la lengua y la laringe (asumiéndose que las paredes faríngeas, el paladar duro y blando y el velo son fijos).

El modelo pretendía calcular el F1 y el F2 para once vocales del inglés americano y los resultados de la simulación mostraron como plausibles las siguientes hipótesis: la lengua posee un sistema muscular organizado en un número de bloques funcionales para la producción del habla; los parámetros esenciales del modelo son tp (posición del cuerpo de la lengua) y ts (forma del cuerpo de la lengua) y estos describen los músculos de la lengua que se combinan de forma antagónica; la morfología del tracto vocal y de las funciones cinemáticas de los músculos de la lengua hacen posible que el niño produzca una meta acústica estable. A partir de aquí Maeda y Honda (1994) creen que son las paredes exteriores del ángulo hacia la derecha que forma el tracto vocal las regiones clave para la formación de un triángulo vocálico estable conformado por /i, a, u/. Además creen que esta estabilidad se conseguiría con un pequeño número de variables de control como son la anteriorización del cuerpo de la lengua para la activación del músculo GGp y la producción de /i/, la posteriorización del cuerpo de la lengua por la contracción del HG y la producción de /a/ y la activación del OOs para la creación de una segunda constricción en los labios característica de la /u/. Así, según Maeda y Honda (1994:27): “[...] It is remarkable that these typical configurations can be derived from extremely simple tongue position control, involving only two pairs of antagonistic muscles, because of the morphology peculiar to the human vocal tract”.

Conocidos son también los modelos que persiguen la producción vocálica basándose en las propiedades biomecánicas de los articuladores del tracto vocal. En estos modelos, como en el del Buchaillard et al. (2009) en el que se crea un modelo biomecánico para la producción de la vocal, se cree que hay una gran influencia de las propiedades biomecánicas en los patrones espacio-temporales de las señales del habla y es a partir de esta influencia que crean arquetipos de control motor biomecánicos. Para estos autores, tal y como explican Perrier et al. (2011) en una breve síntesis sobre el funcionamiento de dichos modelos, las trayectorias, las velocidades y las amplitudes de los movimientos no están dirigidas de manera directa por el sistema nervioso central sino por una combinación de factores entre los que se encuentran los patrones de órdenes motoras y su temporización, y las características biomecánicas de los articuladores. No obstante

para conocer qué propiedades de los movimientos del habla parecen requerir un control del sistema nervioso central, estos expertos llevaron a cabo investigaciones que se prolongaron durante más de quince años para establecer comparaciones entre las propiedades cinemáticas de los movimientos creados por los modelos artificiales y los movimientos articulatorios reales grabados en hablantes de diferentes lenguas (francés, alemán, inglés y chino mandarín). Las conclusiones a las que llegaron fueron que la biomecánica puede explicar mejor determinados atributos de los movimientos de los articuladores como las formas de las trayectorias de dichos movimientos, la relación entre la velocidad y la forma de la trayectoria o la relación entre la precisión del control motor y la variabilidad acústica.

No obstante, y a pesar de la considerable información que aportan todos estos modelos que intentan aclarar el funcionamiento de los movimientos articulatorios, no podemos olvidar los inconvenientes que acarrea el usar un modelo de tracto vocal adulto para averiguar el funcionamiento del habla del bebé y del niño pequeño debido a las diferencias no solo cuantitativas sino cualitativas de las que hablan Stathopoulos y Sapienza (1993:2540): “[...] Because there were both quantitative and qualitative differences, it is concluded that using an adult mechanism model for describing children’s speech production is essentially uninformative”.

Pero el control motor del habla además de estar influenciado por el funcionamiento neuronal y el dominio articulatorio, puede estar relacionado según Nip y Green (2006) con otros factores como las habilidades cognitivas y lingüísticas (e incluso con el afecto) las cuales a la vez pueden facilitar su desarrollo. Así, el estudio de Green et al. (2002) en el que se ha detectado una depresión transitoria en la estabilidad del movimiento del labio y la mandíbula a la edad de 2;0 años, muestra la influencia de dichas habilidades en el control motor puesto que dicha depresión se produce a una edad en la que los niños están lidiando con la fonología y la lengua expresiva.

Para examinar estas relaciones y siguiendo los consejos de Steeve et al. (2008), los autores Nip y Green (2006) decidieron analizar los cambios en la realización motora del habla en niños desde los 0;9 a los 1;9 años de edad, que pasaban de producir movimientos espontáneos mudos (que no requieren ningún procesamiento cognitivo y lingüístico) al balbuceo (que aunque similar a los movimientos espontáneos en cuanto a la falta de objetivos articulatorios en este, el bebé ahora debe conjugar movimientos oro-faciales, respiratorios y fonatorios para producir sonidos parecidos a los del adulto) y finalmente a las palabras (que envuelven la presencia de la fonología, la semántica y los intentos comunicativos). Los resultados fueron los siguientes: a) la velocidad articulatoria de la mandíbula y del labio inferior

aumentaron en el periodo estudiado (esta tendencia continuó pasado el primer año y hasta la edad adulta tal y como han demostrado el estudio de Smith y Gartenberg [1984] y el de Walsh y Smith [2002]); b) el control motor de cada articulador siguió cursos diferentes de forma que la velocidad articulatoria de la mandíbula se estancó a los 1;3 años, y la del labio inferior a los 1;6 años; c) los movimientos mandibulares de los bebés durante el periodo de balbuceo fueron muy similares a los del niño de 1;0 año mientras que los movimientos del labio superior e inferior se parecieron más a los del adulto entre los 2;0 y los 6;0 años de vida. Estos resultados constituyen por tanto un buen reclamo de que el control sobre la mandíbula se desarrolla antes que el de los labios y que el de la lengua, como ha sido afirmado por otros autores (Green et al. [2000] y Green et al. [2002]), y explican que el repertorio fonético inicial de los niños esté limitado a sonidos que estos pueden producir con la mandíbula (Kent [1999] y MacNeilage et al. [2000]).

El control motor requerido para las vocalizaciones (además de estar relacionado con otras habilidades lingüísticas y cognitivas como acabamos de ver) puede estar también coordinado con otros movimientos motores tal y como se demuestra en el estudio de Iverson y Fagan (2004) en el que tras grabar en vídeo a cuarenta y siete niños de entre 0;6 y 0;9 años, un aumento relacionado con la edad se produjo en la frecuencia de coordinación entre las vocalizaciones y el movimiento del brazo, por ejemplo.

Además del control articulatorio necesario para producir las emisiones, cabe preguntarse si al igual que los adultos los niños poseen estrategias de compensación motora a la hora de generar el habla. Se sabe por los estudios de Abbs y Gracco (1984), Folkins y Abbs (1975) y Kelso et al. (1984), citados en Guenther (1994), que el hablante puede apoyarse en diferentes movimientos para producir un mismo sonido de habla. Así, un emisor puede hablar valiéndose de movimientos ascendentes y descendentes de la mandíbula o puede hacerlo sujetando con esta una pipa y manejando un conjunto diferente de posiciones y movimientos de los articuladores.

Uno de los trabajos que ha examinado las habilidades de compensación motora de los niños ha sido el de Oller y MacNeilage (1983). En este estudio se les pidió a dos niños hablantes de inglés de 4;0 y 8;0 años que repitieran las vocales [i] y [æ] bajo tres condiciones: con la mandíbula libre, con los dientes apretados y con un obstáculo colocado entre los dientes. Las emisiones producidas fueron escuchadas por oyentes sin experiencia fonética, ni en emisiones de adultos ni en emisiones de niños, que tenían que decidir a cuál de las siguientes vocales pertenecían las emitidas por los niños: [i], [ɪ], [ɑ], [ɛ], [æ], [ɔ] o [a]. A pesar de que las emisiones del niño de 4;0 años eran más difíciles de interpretar que las del niño de 8;0, los resultados confirmaron que el niño poseía a la temprana edad de 4;0 años habilidades de

compensación articulatoria al menos en lo que respeta a la producción de la vocal [i] ya que todas las vocales producidas con los dientes apretados y el obstáculo en la boca fueron reconocidas como [i] por catorce de los quince oyentes (no sucedió lo mismo con la vocal [æ] que ni siquiera en la condición de mandíbula libre obtuvo el acuerdo entre los oyentes y mucho menos en el resto de condiciones).

Finalmente, y una vez descritos los condicionantes anatómicos y neuronales que debe poseer el bebé para producir el habla, y el grado de control motor del que dispone, resultaría interesante considerar si los gestos articulatorios que permiten dicha producción son gestos que el bebé desarrolla de manera independiente a otros comportamientos, o si por el contrario estos dependen de algunos ya existentes.

La teoría *Frame then content* de MacNeilage y Davis (1990) y MacNeilage (1998) defiende que el habla está directamente relacionada con los ciclos de ingestión cuya oscilación mandibular hace posible la producción del esquema consonante-vocal de la comunicación. Así pues, el marco o *Frame* lo constituiría esta oscilación mandibular, frente al *Content* o contenido que haría referencia a los elementos concretos (esto es, consonantes y vocales) que rellenan ese marco. Para MacNeilage y Davis (1990) la articulación del sonido (por otro lado, diferencia principal entre el habla y otros sistemas de llamada de mamíferos) envuelve el componente articulatorio que se da en una serie de ciclos de abertura y cierre que el niño ya usaba para otros comportamientos tales como masticar, lamer o succionar y que se concretan con dos segmentos, consonante y vocal. Según MacNeilage (1998), la descripción que dan Lund y Enomoto (1988) de la masticación define muy bien los movimientos que se producen para hablar y es a partir de aquí que este autor cree que el habla utiliza un patrón cerebral similar al que se usa para la masticación. Por otra parte, la ya mencionada importancia o dominio que tiene la mandíbula en la articulación de las primeras emisiones frente a la intervención de otros articuladores sirve a MacNeilage y Davis (1990) para refrendar su hipótesis de la dependencia del habla de otros comportamientos ya existentes.

Frente a la teoría del Marco/Contenido también defendida por autores como Kent et al. (1991), encontramos los trabajos de Moore y Ruark (1996) o Ruark y Moore (1997) según los cuales el habla surge de manera independiente a otros comportamientos. En el primer trabajo se llega a esta conclusión después de medir mediante ondas electromiográficas algunos de los músculos primarios de la masticación como el *masseter*, el *temporalis* y el vientre anterior del digástrico. Un año más tarde Ruark y Moore (1997) llevaron a cabo una descripción de la coordinación de la actividad labial en niños de 2;0 años durante la producción de habla y en movimientos en los que no se producía habla. Para ello se analizó la masticación, la protusión

del labio, la repetición de sílabas y el habla en siete sujetos de 2;0 y 2;5 años mediante grabaciones de electromiografía. Los autores encontraron que el grado de acoplamiento o emparejamiento labial (del labio superior e inferior) dependía de la tarea realizada siendo la tarea de repetición de la sílaba la que mayor grado de emparejamiento labial conllevó. Esta organización específica de la tarea (distintos grados de emparejamiento labial son empleados para distintas tareas) llevaron a Ruark y Moore (1997) a pensar que el habla se desarrolla de manera independiente a otros comportamientos.

Otros estudiosos como el ya citado Smith (2010) prefieren mantenerse al margen y simplemente plantean la cuestión en un nivel neuronal al preguntarse si el habla se basa en los CPG (*Central Pattern Generators*) ya existentes para otros comportamientos como el masticatorio, o si por el contrario posee CPG independientes.

## **2.2. Capacidades perceptivas del bebé y del niño pequeño.**

### **2.2.1. Habilidades perceptivas y evolución de estas.**

Una vez conocida la configuración anatómica con la que nace el bebé y su evolución a lo largo de los primeros meses de vida, así como el grado de control motor de los distintos articuladores y el funcionamiento neuronal que subyace a todo ello, nos disponemos a describir las habilidades perceptivas de que dispone el bebé y su desarrollo en los primeros años de vida.

En la introducción a “Experiential influences on speech perception and speech production in infancy”, Polka et al. (2007:151) describen el desarrollo de la percepción y la producción del habla como el resultado de tres factores: “[...] infants’ initial perceptual capacities and biases, the language(s) to which infants are exposed, and infants’ perceptions of the sounds they themselves produce”. En cuanto a lo que a capacidades perceptivas se refiere, conviene advertir antes de revisar los principales trabajos que la obtención de los datos se ha llevado a cabo principalmente en laboratorios debido a que para medir los distintos índices que confirman la percepción del sujeto, como son la respuesta refleja del feto, el ritmo cardíaco del bebé, etc., han sido necesarios una serie de instrumentos que el investigador no podía transportar a la casa familiar.

Observada esta puntualización, debemos señalar que dichas capacidades, las perceptivas, ya están presentes en el bebé aun cuando este no ha acabado de formarse. Así, a los cinco meses de la edad gestacional y según Lecanuet (1995), Hopkins (2005) y Kent y Vorperian (2006), el feto es capaz de escuchar los estímulos acústicos del exterior mezclados

con sonidos propios del cuerpo de la madre (sin que el latido del corazón ni los ruidos procedentes de la corriente sanguínea enmascaren según Querleu et al. [1981] y Querleu et al. [1985], la voz de la madre). Ello solo puede entenderse si a esa edad el oído humano ya está formado y existe además una cierta complejidad neuronal relacionada con la audición. Por tanto, y según los primeros autores arriba citados, a los cinco meses de la edad gestacional el órgano del oído ya está bastante desarrollado.

En cuanto al desarrollo neuronal relacionado con la percepción y según las observaciones de Moore (2002), citadas en el trabajo de Clement (2004), entre las semanas 16 y 28 de gestación el nervio auditivo y la parte auditiva del bulbo raquídeo empiezan a desarrollarse comenzando también la conducción de la información auditiva. A la semana 29 el bulbo y la cóclea han madurado neurológicamente. Una vez que el bebé nace y entre los meses cuarto y duodécimo, el procesamiento cortical de la información auditiva se hace más complejo debido a que los axones (parte de la neurona) alcanzan las capas más profundas del córtex cerebral y determinados aferentes corticales (nervios que lleva información a los órganos) maduran. Finalmente y según Moore (2002), entre los cinco y los doce años de edad se produce un refinamiento neuronal gracias en parte a que los axones del córtex auditivo crean más conexiones entre las partes del córtex y los dos hemisferios resultando en una mayor complejidad del procesamiento auditivo del habla desde esa edad en adelante.

Entre los experimentos que han demostrado la capacidad de audición del feto, encontramos el trabajo de Birnholtz y Benacerraf (1983) en el que se mostró mediante grabaciones de ultrasonido que entre la semana 28 y la 36 el feto respondía con parpadeos reflejos a diferentes sonidos altos. También fue demostrado que entre la semana 36 y 39 los fetos podían detectar el cambio de notas de piano de bajo *pitch* ya que respondían a este cambio con una desaceleración del ritmo cardíaco (Lecanuet et al. [2000]).

Asimismo, existen investigaciones que han manifestado que el bebé que no ha llegado al final del periodo de gestación no solo es capaz de percibir diferentes sonidos sino que este además tiene la habilidad de reconocer la cadena de habla a la que ha sido expuesto de manera reiterada. Es el caso del trabajo de DeCasper et al. (1994) en el que el ritmo del corazón de los fetos de 38 semanas de gestación descendió al escuchar el pasaje de un cuento cuyas madres habían leído en voz alta tres veces al día durante las semanas de gestación que fueron de la trigésimo tercera a la trigésimo séptima, frente a la audición de un pasaje no familiar. Ocho años antes DeCasper y Spence (1986) sugirieron que el feto era además capaz de retener información auditiva ya que después del nacimiento los bebés modificaron la velocidad de succión al escuchar el pasaje del cuento al que habían sido expuestos durante la gestación. Según estos

autores la influencia de las experiencias que se producen en el útero no está limitada solo a la percepción sino que puede darse también a nivel olfativo y de gusto.

Otra de las habilidades perceptivas que se ha detectado en el feto es la del reconocimiento y la preferencia de este por la voz materna. Así se mostró en el estudio de Kisilevsky et al. (2003) en el que se pidió a sesenta madres chinas que leyeran un poema a sus fetos a lo largo del embarazo. En la recta final del periodo de gestación los fetos que fueron expuestos a una grabación de audio en la que se escuchaba la voz de la madre leyéndoles dicho poema aumentaron el número de latidos por minuto frente a los fetos que escucharon una grabación en la que una voz no familiar leía el mismo poema. No obtuvieron resultados idénticos cuando los mismos autores decidieron comprobar si los bebés diferenciaban la voz de sus madres de la de una extraña a través del movimiento del cuerpo. Tampoco los bebés del estudio de Beeper et al. (1993) respondieron con un movimiento corporal cuando fueron expuestos a la voz de la madre y a la de una extraña. Estos resultados quizás respondan más a la utilización de la variable del movimiento corporal que a la capacidad de los bebés de diferenciación de la voz materna.

En un trabajo posterior, Kisilevsky et al. (2009) volvieron a demostrar la identificación por parte del feto de la voz de su madre y sugirieron que las redes neuronales que son sensibles a las propiedades de la voz de la madre y a las propiedades de la lengua nativa comienzan a formarse en el periodo gestacional. Los trabajos de DeCasper y Fifer (1980) y de Jusczyk (1997) avalaron también la preferencia de los bebés que aún están en el vientre materno por la lengua ambiente y la voz de su madre. No obstante, encontramos también publicaciones como las de Hepper et al. (1993) en las que los fetos no discriminaron la voz materna de otra voz femenina no familiar aunque ellos sí diferenciaron entre el habla directa de la madre y una grabación de su voz. Una vez que el bebé ha nacido y a la edad de entre cuatro y seis semanas, este sigue prefiriendo la voz de la madre frente a la de una extraña y además el ritmo de succión no nutritiva es más rápido cuando el habla de la madre conserva los patrones de entonación (Mehler et al. [1978])

La distinción entre dos lenguas y la preferencia por la lengua materna es otra de las capacidades de percepción que exhibe el neonato. Así, en el estudio de Mehler et al. (1988) bebés monolingües de inglés fueron capaces de diferenciar el inglés del italiano a los 0;2 años de edad. En el mismo año el experimento de Bahrick y Pickens (1988) arrojó resultados similares: la distinción de la lengua española e inglesa por parte de bebés a la edad de 0;5 años. Pero es sin duda el trabajo de Moon et al. (1993) el que ofreció más evidencias de que esta capacidad debe gestarse en el útero ya que neonatos de tan solo dos días de edad mostraron una



clara preferencia por su lengua materna aumentando las ratios de succión frente a la audición de una lengua desconocida. Según Mehler et al. (1988) la habilidad discriminadora de los niños es mayor cuanto más diferentes son las lenguas a diferenciar. Una conclusión reveladora es la que aparece en el estudio de Dehaene-Lambertz y Houston (1998) (citado en el artículo de Clement [2004]), según la cual el bebé de dos meses de edad no solo mostraba una preferencia por su lengua nativa sino que dicha preferencia desaparecería si se distorsionaba la información prosódica de las muestras de habla a las que eran expuestos. En este último campo, el de los experimentos en los que se estudia la prosodia, cabe citar el trabajo de Jusczyk et al. (1993) que muestra una inclinación de bebés hablantes de inglés americano y de noruego a los seis meses de edad por los patrones prosódicos de su lengua nativa a partir de la audición de palabras.

Otros contrastes que son capaces de percibir los bebés después de haber nacido, y que según Mattock et al. (2010:305) son claves para la percepción del habla, son por ejemplo los cambios de intensidad y los índices temporales: “[...] Sensitivity to spectro-temporal, duration, intensity, frequency, and periodicity information is paramount to speech perception”. Respecto a los primeros, el estudio de Tarquinio et al. (1990) mostró que los neonatos distinguieron sonidos con diferencias de 6 dB de intensidad, y el de Sinnott y Aslin (1985) manifestó que los bebés son capaces de distinguir la misma diferencia de intensidad en tonos de 1000 Hz entre los 0;7 y los 0;8 años. En cuanto a la distinción temporal, en el trabajo de Mattock et al. (2010) se nos habla de la importancia que tiene el procesamiento temporal para distinguir entre categorías y de como los niños son capaces de hacerlo desde muy temprano según Eimas et al. (1971).

Los experimentos de Chang y Trehub (1977), en los que se analizó la discriminación de los niños de secuencias de tonos con diferente agrupamiento temporal, y los de Morrongiello y Trehub (1987), en los que se probó la identificación de niños de 0;6 años de diferencias de duración, son prueba de esta temprana destreza. Según Mattock et al. (2010), la discriminación de la duración mejora con la edad y la discriminación similar a la del adulto aparece entre los 8;0 y los 10;0 años de edad. En cuanto a la información espectral, en el trabajo de Clarkson et al. (1988) los infantes fueron capaces de discriminar tonos complejos con idéntico *pitch* que tan solo variaban en la forma espectral. Otros estudios han mostrado como los niños tienen la habilidad de diferenciar dos tonos apoyándose en el sello espectral incluso cuando estos tonos presentan variación de otros correlatos acústicos. Así lo muestra el experimento de Trehub et al. (1990) en el que infantes de entre 0;7 y los 0;8 años y medio de edad fueron capaces de diferenciar tonos complejos con una estructura espectral específica en los que variaba la frecuencia fundamental, la intensidad y la duración. Por lo que respecta a la variación de tono, Trehub et al. (1987) demostraron mediante el procedimiento de giro de cabeza que los bebés podían extraer el contorno del *pitch* (cambios direccionales en el *pitch*) de una secuencia

multitonal o melodía. Más tarde, entre los 0;7 y los 0;10 años de edad, los niños son capaces de reconocer cambios de dirección en la frecuencia al distinguir grupos de tonos ascendentes y descendentes (Mattock et al. [2010]).

Entre los primeros estudios que han analizado la capacidad de los fetos, neonatos y bebés de distinguir los múltiples correlatos que diferencian los sonidos consonánticos, encontramos el trabajo de Eimas et al. (1971) en el que mediante la técnica de la succión se demostró que bebés de entre 0;1 y 0;4 años de edad diferenciaban /pa/ de /ba/ y que lo hacían además de manera categórica (respaldando por tanto los experimentos de Liberman et al. [1957] según los cuales la percepción de las consonantes oclusivas es categórica como veremos más adelante). En el mismo año Moffit (1971) se valió de sílabas de habla sintética para mostrar que los bebés de entre 0;5 y 0;6 años de edad eran capaces de distinguir la sílaba *bah* de *gah*. Un año más tarde Trehub y Rabinovitch (1972) demostraron que infantes de entre 0;1 y 0;4 años de edad pudieron discriminar la oclusiva sorda de la sonora (/p/ vs /b/) tanto en muestras de habla sintética como en naturales. En dicho trabajo se cita una comunicación personal de McCaffrey (1971) en la que mediante el control del ritmo cardíaco se probó que bebés de entre 0;1 y 0;7 años de edad discriminaban los contrastes consonánticos de /p/, /t/, /k/, /s/ y /n/. También el estudio de Eilers y Minifie (1975) manifestó que infantes de 0;6 años eran capaces de diferenciar las fricativas /s/ vs /ʃ/. Otros trabajos que mitigan en parte la escasez de experimentos de diferenciación consonántica son los de Trehub (1976) o los de Werker et al. (1981) y Werker y Tees (1983, 1984) que analizaron la capacidad del bebé de discriminar entre sonidos consonánticos nativos y no nativos y que veremos más adelante con motivo de la reorganización de las capacidades perceptivas del infante.

De mayor interés para nuestra investigación es la destreza que presentan los fetos y los neonatos para percibir sonidos vocálicos y que consiste según Mattock et al. (2010:305-306) en “[...] Perceiving vowels relies on the ability to detect and discriminate bands of energy in multiple, relatively invariant frequency regions (formants) and to process that information over time”. El estudio de Lecaunet et al. (1987) señaló que incluso antes del nacimiento los fetos eran capaces de discriminar pares de sílabas en las que se cambiaban sonidos vocálicos (/babi/ vs /biba/). Más tarde, el trabajo de Zimmer et al. (1993) corroboró la habilidad de discriminación vocálica de los fetos mediante un experimento en el que los bebés entre la vigésimo sexta y trigésimo cuarta semana de gestación ya eran capaces de distinguir la forma espectral al discriminar satisfactoriamente los sonidos vocálicos /ee/ y /ah/ (no obstante, necesitaron estímulos de mayor intensidad). Esta capacidad fue también identificada en bebés de entre 0;1 y 0;4 años tal y como muestra el trabajo de Trehub (1973) en el que a esa edad los infantes diferenciaron mediante la técnica de succión no nutritiva la [a] de la [i], y la [a] de la [u] tanto

en los dos experimentos en los que estas vocales aparecían aisladas como en los que aparecían en las sílabas /pa/-/pi/ y /ta/-/tu/. En el trabajo de Kuhl y Miller (1982) bebés de entre 0;5 y 0;6 años y medio tuvieron que diferenciar sonidos vocálicos producidos por hablantes de distinto sexo y edad. Los resultados mostraron que los niños fueron capaces de discriminar todas las vocales incluso aquellas acústicamente muy similares como /a/ y /ɔ/.

Distinto es el trabajo de Benders (2011) según el cual los infantes tardarían más en diferenciar dos vocales con timbre semejante que difiriesen únicamente en el rasgo de posterioridad y duración (como es el caso de la /a:/ y la /ɑ/ holandesas). Otros trabajos que han mostrado la habilidad temprana de los bebés para diferenciar sonidos vocálicos son por ejemplo los de Stevens et al. (1969), Beddor y Strange (1982), Bohn y Flege (1990) (todos citados en Polka y Bohn [1996]), y el de de Polka y Werker (1994), estudios que al tener como objetivo la manifestación de la transformación de las capacidades perceptivas de los bebés por influjo de la lengua materna, serán tratados más adelante cuando veamos la influencia de la lengua ambiente en el niño.

A la hora de diferenciar sonidos vocálicos y tal y como señalan Polka y Bohn (2011), llama la atención que la discriminación del bebé de un cambio vocálico presentado en una dirección sea significativamente mayor que la del mismo cambio vocálico presentado en la dirección inversa. A este fenómeno se le conoce con el nombre de *asimetría direccional* y fue observado cuando se probó la capacidad de los bebés de distinguir entre pares vocálicos que incluían contrastes nativos y no nativos. Según Polka y Bohn (2003), dentro de la asimetría el cambio de una vocal más central a una más periférica es más fácil de distinguir por el infante que el cambio presentado en la dirección inversa (los autores definen vocal periférica como aquella vocal que está más cerca de los límites de un extremo del espacio vocálico).

Entre los primeros trabajos que revelaron este fenómeno direccional (todos citados en Polka y Bohn [2003]) encontramos el estudio de Dejardins y Trainor (1998), en el que los niños discriminaron mejor el cambio de /i/ a /i/ que el cambio inverso, o el de Best et al. (1997) en el que bebés de entre 0;3 y 0;5 años de edad que adquirían el inglés discriminaron mejor en el contraste noruego /y/-/u/, el cambio que fue de /u/ a /y/, es decir de la vocal menos periférica a la más periférica, que el cambio que fue de /y/ a /u/. La misma asimetría fue encontrada para el grupo de niños más jóvenes en los tres experimentos del trabajo de Polka y Bohn (2011). En el primer experimento, los infantes de entre 0;6-0;8 y de entre 0;8-0;11 años y mediante la técnica de giro de cabeza, discriminaron mejor un contraste vocálico no nativo cuando la dirección en la que se presentó fue de la vocal más central, esto es /ʌ/, a la más periférica, /ɒ/. En el segundo experimento, los niños de aproximadamente 0;6 años distinguieron mejor el contraste vocálico

nativo /e/-/ε/ cuando el cambio se produjo de /ε/ a /e/ que de /e/ a /ε/. Finalmente, en el tercer experimento los niños de 0;6 años de edad discriminaron mejor el cambio que fue de la vocal menos periférica /ø/ a la más periférica /e/ que el cambio presentado en la dirección opuesta.

Estas asimetrías en opinión de Polka y Bohn (2003) podrían ser explicadas si consideramos que los bebés están predispuestos a responder de manera diferente a las vocales que ocupan diferentes posiciones en el espacio articulatorio-acústico (en el que la posición de cada vocal está definida por la altura de esta, correspondiente al F1, y por su posición en el eje antero-posterior, correspondiente con el F2). Para estos autores (2003:222) y para otros como Repp et al. (1979), Cowan y Morse (1986), Repp y Crowder (1990), Iverson y Kuhl (1995) y Sussman y Lauckner-Morano (1995) (todos ellos citados en Polka y Bohn [1996]), las vocales de los extremos son elementos que atraen a otras vocales sirviendo como puntos perceptivos de referencia: “[...] peripheral vowels may serve as important perceptual reference points for human listeners”. Para Polka y Bohn (1996) las vocales de los extremos son por ello sonidos acústicamente más estables que pueden dar indicaciones sobre el tamaño del tracto vocal (concretamente /i/ y /u/) y además pueden guiar las habilidades de producción vocálica en el desarrollo temprano del habla (para una explicación de lo anterior los autores recomiendan la revisión de Johnson [1993]).

Schwartz et al. (2005) también vieron estas asimetrías como una predisposición para las vocales periféricas en el espacio F1-F2, y relacionaron dichas vocales periféricas con la Teoría de la Focalización y con la Teoría Cuántica de Stevens (1972, 1989). Según la Teoría Cuántica existen vocales cuyos formantes están muy cercanos, ya sea F1-F2, F2-F3 o F3-F4, y ello hace que dichas vocales sean más estables acústica y articulatoriamente. Esta convergencia de formantes consecutivos es lo que Schwartz et al. (2005) llamaron *focalización* (inicialmente Boë y Abry [1986]) y es esta propiedad la que caracteriza a las vocales periféricas y la que hace que dichas vocales sirvan como ancla o referencia para la discriminación (al ser vocales más estables acústicamente pueden atraer a otras vocales hacia sí). Otra de las razones que aportan Schwartz et al. (2005) para alegar que las vocales focales o periféricas actúan como un ancla en la discriminación es la de que estas vocales parecen perdurar más en la memoria a corto plazo. Para defenderlo acuden al experimento de Schwartz y Escudier (1989) según el cual estímulos sintéticos vocálicos con convergencia formántica fueron más notorios y más estables en la memoria a corto plazo para un conjunto de oyentes franceses (la focalización por tanto ayuda a recordar los sonidos).

En un trabajo posterior, Polka y Bohn (2011:474) denominaron a las vocales periféricas *Natural Reference Vowels* o *Vocales de Referencia Natural* ya que “[...] support and guide the

development of vowel perception by attracting infant attention and providing stable perceptual forms for the language learner”. Además, el que todas las lenguas del mundo cuenten con las vocales periféricas /i, u/ apoya el hecho de que haya una tendencia natural hacia este tipo de vocales que sirven como puntos de referencia para organizar la percepción de la vocal en la adquisición: “These early vowel perception biases reveal the perceptual proclivities of the human auditory system at birth, and/or perceptual learning occurring in the first few weeks or months of life” (2011:474). Así según los autores, este conjunto mínimo de vocales provee de importantes puntos de referencia para descubrir otras categorías vocálicas adicionales y juega por tanto un papel importante en la adquisición de la L1 ya que establece un andamio para adquirir e interpretar los sonidos vocálicos.

Finalmente, Schwartz et al. (2005) relacionaron la notabilidad y estabilidad de las Vocales de Referencia Natural con la teoría del efecto de gravedad de Chistovich y Lublinskaya (1979) puesto que el resultado de la convergencia formántica es que la energía se concentre en una región estrecha del espectro y produzca prominencias en dicho espectro vocálico a las que es sensible el oído humano (recordemos que según dicha teoría el oído integra los formantes que están cerca en un solo formante o centro de gravedad que es el promedio de los dos formantes individuales).

No obstante, y tal como se advierte en el trabajo de Sussman (2001), hay muchos contrastes que los bebés no pueden percibir y que evidencian que sus habilidades perceptivas no son aún como las del adulto. En dicho artículo se recuperó el trabajo de Schneider et al. (1986) según el cual a la edad de 5;0 años la sensibilidad de los niños a la mayoría de las frecuencias es más pobre que la de los adultos. También se citó el estudio de Trehub et al. (1988) en el que se encontró que para estímulos con frecuencias centrales (entre 400 Hz y 1000 Hz) la sensibilidad máxima de los bebés no fue encontrada hasta los 10;0 años de edad. De la misma manera, el trabajo de Berg y Boswell (2000) mostró que los niños pequeños no poseen la discriminación de los adultos a la hora de diferenciar un cambio de intensidad en intensidades bajas hasta después de los 3;0 años de edad. Entre las razones, y ya que la cóclea está desarrollada desde el nacimiento según Schneider y Trehub (1992), puede estar el hecho de que los bebés y los niños pequeños necesiten una mayor cantidad de índices acústicos tales como un mayor volumen en el habla y una mayor duración del estímulo (Elliott et al. [1981] y Ohde y Haley [1997]) o necesiten cantidades mayores de información espectral (Dorman et al. [1998] y Eisenberg et al. [2000]). Este último autor junto con otros como Parnell y Amerrman (1978), Elliott et al. (1981), Sussman y Carney (1989) y Ohde and Haley (1997) sugirieron que las habilidades de percepción del habla del niño no son como las del adulto ni siquiera a los 11;0 años de edad.

A pesar de estas dificultades a la hora de percibir ciertos contrastes acústicos, y tal y como puede desprenderse si comparamos la capacidad articuladora de los bebés con las habilidades perceptivas de los fetos y neonatos que han sido expuestas en las páginas anteriores, la percepción de los sonidos en los bebés va un paso por delante con respecto a la producción de los mismos. Así lo expone Kent (1981:166) quien cree que:

“[...] Whatever the relative contributions of genetic and experiential factors might be in these early auditory discriminations, it is clear that the child can make reliable auditory discriminations of speech cues before she or he can reliably produce the same cues with the speech mechanism”.

Una vez que conocemos la habilidad de los bebés de discriminar las vocales y las consonantes, sería interesante comprender de qué forma perciben los niños estos sonidos tan diferentes. La percepción categórica del punto de articulación de las consonantes oclusivas fue demostrada por primera vez a mediados del siglo veinte gracias al trabajo Liberman et al. (1957) que se realizó en los Laboratorios Haskins y en el que sintetizaron catorce estímulos formados por un F1 y un F2 en los que se modificó la transición de este último formante. El estímulo 1 que presentaba un F2 bajo fue percibido como /b/, el estímulo 8 con un F2 medio se escuchó como /d/ y el estímulo 14 con un F2 alto fue reconocido como /g/. Sin embargo, los oyentes no percibieron consonantes intermedias en el resto de los estímulos sino que escucharon /b/ en los estímulos que iban del primero al cuarto, /d/ en los que iban del cuarto al octavo, y /g/ en los que iban del octavo al decimocuarto. Esto confirmó que la percepción de las consonantes oclusivas era categórica, es decir, que al pasar del estímulo cuarto al quinto cambiaba la categoría pero esta se mantenía en los estímulos intermedios.

Tal y como manifiesta González Álvarez (1997), la percepción categórica se probó con otros contrastes como el del VOT (*Voice Onset Time*) que no es sino el tiempo que tardan las cuerdas vocales en vibrar una vez que ha desaparecido la oclusión de los órganos articulatorios y se ha producido la explosión a la salida abrupta del aire. De nuevo el experimento pionero que demostró que la percepción de las oclusivas con respecto al VOT era categórica, fue llevado a cabo en los Laboratorios Haskins de la mano de Abramson y Lisker (1967) y fue repetido con bebés de entre 0;1 año y 0;4 años de edad por Eimas et al. (1971) cuatro años más tarde. Según el trabajo de Eimas et al. (1971), ya mencionado para hablar de la capacidad de los bebés de discriminar entre sonidos consonánticos y en el que se probó la habilidad de estos de distinguir las consonantes oclusivas sordas de las oclusivas sonoras de la lengua inglesa mediante la modificación del VOT, los bebés diferencian estos sonidos de forma categórica comportándose su percepción de este modo como la de un adulto. En opinión de González Álvarez (1997:55) la percepción categórica representa para el bebé y el adulto:

“[...] una forma de manejar la enorme cantidad de variación entre las diferentes producciones del mismo sonido que típicamente encontramos en el habla humana. Nos da un medio de manejar esta variación en tiempo real –diferencias irrelevantes entre distintas producciones del mismo sonido de habla se ignoran con el fin de acelerar el proceso de reconocimiento-”.

Frente a la percepción categórica de las consonantes, parece haber un acuerdo general en que las vocales son percibidas de forma continua al menos por los adultos (y no hay motivos para rechazar que sea esta también la percepción que los bebés presentan de los sonidos vocálicos). El hecho de que se perciban de manera continua implica según González Álvarez (1997) que cualquier variación en los formantes es percibida por el oyente. De esta forma, si escuchásemos sonidos cuyos formantes estuvieran entre los de la /e/ y la /a/, percibiríamos vocales intermedias entre estas (como por ejemplo /ɛ/ o /æ/).

Pocos son los autores que han demostrado que los estados estables de las vocales son percibidos de manera continua (entre ellos encontramos a Fry et al. [1962], Stevens et al. [1969] y Cutting et al. [1975]) pero mucho más escasos son los trabajos que han demostrado esta habilidad en los bebés y en los niños pequeños. Uno de estos pocos estudios es el de Swoboda et al. (1976) en el que se investigó la discriminación entre /i/ e /ɪ/ en niños de 0;2 años mediante el procedimiento de succión no nutritiva y cuyos resultados concluyeron que los infantes de 0;2 años de edad diferenciaban de manera continua y no categórica la vocal /i/ de /ɪ/. No obstante, la percepción continua de las vocales no es aceptada de manera universal tal y como recoge el artículo de Pisoni (1975) en el que se nos ofrece una revisión de estudios que demuestran que existe algún grado de percepción categórica en vocales (y cita los de Stevens et al. [1969], Fujisaki y Kawashima [1969] y Pisoni [1971]). También en el trabajo de Borzorne de Manrique (1977) se alude al estudio de Manrique y Gurlekian (1977) según el cual las vocales del español, a diferencia de las inglesas, se perciben de un modo categorial.

Vista en la introducción la importancia que los dos primeros formantes poseen para el reconocimiento de la vocal por parte del oyente, un aspecto clave en la percepción de la vocal por parte del bebé y del niño pequeño es averiguar de qué forma se fijan estos en los formantes vocálicos. Para la mayoría de los estudiosos, el adulto identifica las vocales apoyándose más que en las frecuencias absolutas de los formantes, y es esto según Mattock et al. (2010:305-306) aquello en lo que consiste percibir las vocales “[...] Perceiving vowels relies on the ability to detect and discriminate bands of energy in multiple, relatively invariant frequency regions (formants) and to process that information over time”, en la relación existente entre dichas frecuencias. Pero fue Richard Lloyd (1890a, 1891, 1892) a finales del siglo diecinueve quien primero formuló esta teoría a la que llamó *La teoría de la resonancia relativa* según la cual el

timbre vocálico depende de los intervalos que existen entre las resonancias, no de sus valores absolutos. Así pues, a la hora de identificar una /i/ el oyente se fija en la relación frecuencial existente entre sus dos primeros formantes, a saber, un F1 muy bajo y un F2 muy alto. Lo contrario sucedería con la /a/ para cuya identificación el oyente extraería la relación formántica a partir de un F1 alto y un F2 bajo. No obstante, para autores como Miller (1987) los valores absolutos de las frecuencias deben contar de alguna manera y para otros como Fant et al. (1974), Fujisaki y Kawashima (1968), Ainsworth (1975), Scott (1976), Traunmüller (1981) y también para el propio Miller (1953), todos ellos citados en Miller (1987), el *pitch* de la voz puede influir en la identidad de la vocal ya que puede ayudar a identificar la apertura vocálica.

Asimismo, para Syrdal y Gopal (1986) la frecuencia fundamental juega un papel importante, si no definitivo, en la identificación de la vocal tal y como se demostró en el experimento que llevaron a cabo apoyándose en la teoría del centro de gravedad de Chistovich et al. (1979) y Chistovich y Lublinskaya (1979). Dicha teoría parte por un lado de los experimentos de Delattre et al. (1952) según los cuales las vocales posteriores podían sintetizarse en un estímulo de un solo formante situado en una posición intermedia entre el F1 y el F2 debido a la proximidad de estas dos frecuencias en dichos sonidos; y por otro de la transformación de los formantes en bandas críticas o barks. Esto último descansa a su vez en diferentes experimentos psicoacústicos que según Plomp (1975), Schroeder (1975) y Schroeder et al. (1979) (todos ellos citados en Syrdal y Gopal [1986]) demostraron que el sistema auditivo está compuesto por filtros de paso banda, y se apoya también en la escala de tonalidad propuesta por Zwicker (1961) que dividía el rango auditivo humano por debajo de los 16 Khz en 24 unidades de bandas críticas que más tarde se denominarían barks (relación logarítmica entre los formantes y la audición de los mismos que los transforma en barks). A partir de aquí Chistovich et al. (1979) y Chistovich y Lublinskaya (1979) establecen el límite de 3-3,5 barks como distancia crítica que debe haber entre el F1 y el F2 de un sonido para que un estímulo sintético formado por un solo formante situado entre ese F1 y F2, sea reconocido como la vocal real de dos formantes (al menos para las vocales rusas). Al promedio que hace el oído humano entre los valores del F1 y del F2 que están próximos en frecuencia para arrojar una sola resonancia formántica situada entre los dos formantes se le denomina *Teoría del efecto de gravedad*.

Syrdal y Gopal (1986) utilizarán dicha teoría para explicar la relación existente entre la  $f_0$  y el F1 del sonido como índice de la abertura vocálica. Según estos, el oído humano diferenciaría las vocales altas (como /i, ɪ, u, ʊ/) de las medias y las bajas porque las primeras presentarían una distancia de 3-3,5 barks entre la  $f_0$  y el F1 frente a las medias y bajas en las que se excedería dicha distancia. También para Hoemeke y Diehl (1994) la altura vocálica varía con la distancia en barks entre el primer formante y la frecuencia fundamental. No obstante, los



resultados de un estudio llevado a cabo por estos últimos con las vocales del inglés revelaron que la distancia en barks entre el F1 y la  $f_0$  como principal correlato de la altura vocálica solo fue concluyente para el par /ɪ/-/ɛ/ y no para los pares /ɛ/-/æ/ y /i/-/ɪ/ para los que el F1 fue el índice más relacionado con la altura percibida de lo que lo fue la distancia F1- $f_0$ . Según los propios autores la razón de esta discrepancia puede estar en que la  $f_0$  interna a cada vocal, también llamada por estos *pitch* intrínseco a la vocal o  $1F_0$ , es más alta en las vocales altas y esto puede interferir en su reconocimiento. Para otros como Fahey et al. (1996) quizás se deba a que solo la distinción entre /ɪ/-/ɛ/ es una distinción pura fonética y fonológicamente y no lo es en las demás. En su experimento, Fahey et al. (1996) probaron la relación entre el F1 y la  $f_0$  en la identificación de la altura vocálica pero en este caso con vocales posteriores que comprendían en un primer experimento los pares /u/-/ʊ/, /ʊ/-/ɔ/, y /ɔ/-/ɑ/ y en el segundo /u/-/ʊ/, /ʊ/-/ʌ/ y /ʌ/-/ɑ/. Los resultados mostraron que tan solo para el par /ʊ/-/ʌ/ la distancia F1- $f_0$  era pertinente y entre las causas los autores señalaron la existencia en el resto de pares de otras diferencias, además de la de la altura, que podían oscurecer el efecto de la F1- $f_0$  (por ejemplo entre /ʊ/-/ɔ/ además de existir una diferencia de altura las vocales se distinguen por el rasgo tenso y por la duración).

En cuanto a cómo procesa el cerebro la relación formántica y de la  $f_0$  a la hora de identificar el sonido vocálico, y a pesar de que se sabe poco de los mecanismos neuronales que subyacen a dicha identificación, según González Álvarez (1997) es el órgano de Corti el que lleva a cabo un análisis frecuencial de la onda realizando según él una especie de espectrograma neuronal del sonido. La relación frecuencial se codificaría en el nervio auditivo y se transmitiría a la corteza auditiva.

Además de la percepción formántica y a la hora de discriminar las vocales, son útiles para Pisoni (1973) tanto la memoria fonética, que permite comparar vocales de diferente categoría fonética, como la memoria auditiva que resulta útil para determinar diferencias entre categorías y dentro de estas.

El que no existan experimentos que hayan probado la capacidad de los niños de distinguir categorías fonéticas a partir de la diferencia de barks entre el F1 y la  $f_0$ , y el hecho de que existan estudios que han demostrado que las capacidades perceptivas del bebé, incluso las del feto, son similares a las del adulto en cuanto a distinción de sonidos de diferentes lenguas, distinción de vocales y consonantes y distinción categórica de estas últimas, puede hacernos suponer que también en el niño la relación entre los formantes y la dependencia del F1- $f_0$  influyen en la identificación de la vocal.

La ya explicada teoría de la ratio formántica, según la cual el oyente identifica las vocales a partir de la relación entre sus formantes, nos sirve muy bien para introducir el concepto de normalización ya que según Peterson (1961) la mayor fuerza de esta teoría radica en su habilidad para reducir en gran medida las diferencias encontradas en las vocales relacionadas con el sexo y edad del hablante. Así pues, la normalización no es sino el proceso mediante el cual el oyente es capaz de ignorar la variabilidad acústica de los sonidos del habla debida al diferente tamaño del tracto vocal del hablante, a su diferente edad o sexo e incluso al diferente estado afectivo de este. Lieberman (1980:137) defiende dicho proceso mediante la hipótesis de:

“[...] the presence of a perceptual mechanism that is present in all normal human beings that ‘normalizes’ incoming speech sounds in terms of the presumed length of the speaker’s presumed supralaryngeal vocal-tract length and that, likewise, allows the child to ‘know’ that it has produce an equivalent signal as it listens to its own speech”.

En palabras de Disner (1980), dos son los requerimientos para los esquemas de normalización: maximizar las diferencias entre categorías vocálicas distintas y minimizar las diferencias para una misma vocal hablada por diferentes hablantes. Algunos experimentos que han demostrado que el bebé es capaz de llevar a cabo dicho proceso de normalización son por ejemplo el de Kuhl y Miller (1982) cuyos resultados manifestaron la capacidad de bebés de entre 0;5 y 0;6 años de discriminar las vocales emitidas por hablantes adultos de sexo masculino y femenino y por niños mediante la técnica del giro de cabeza. Un año más tarde Kuhl (1983) volvió a mostrar que a los 0;6 años de edad los bebés eran capaces de reconocer la similitud fonética en estímulos procedentes de distintos hablantes, y ya en la década de los noventa Walton (1991) demostró que incluso los recién nacidos poseen la habilidad de encontrar la unidad en la diversidad a partir de estímulos vocálicos producidos por diferentes hablantes. Muy cercano se encuentra el estudio de Marean et al. (1992) que manifestó que bebés de 0;2, 0;3 y 0;6 años de edad reconocieron como el mismo sonido vocálico aquellas vocales del inglés producidas por un hablante de sexo masculino y uno de sexo femenino en las que incluso se varió la frecuencia fundamental en unos 20 Hz.

El cuanto al proceso mediante el cual el oyente lleva a cabo dicha normalización, González Álvarez (1997) recoge en su trabajo “La percepción del lenguaje” dos modelos posibles: el modelo de normalización intrínseca, según el cual cada vocal porta suficiente información acústica para permitir la normalización; y el modelo de normalización extrínseca, para el que los oyentes necesitan información externa a la vocal que les permita crear un marco de referencia sobre algunas características del hablante como la edad, el sexo, el tamaño del tracto vocal, etc. Según el autor, los partidarios del primer modelo subrayan la importancia que

los formantes superiores (como el F3) y la  $f_0$  tienen en la normalización al aportar información adicional al hablante (información sobre la edad, sexo, etc.), información que no aporta la relación entre el F1 y el F2. Para los defensores del segundo, el oyente tiene en cuenta no solo los formantes de la vocal emitida sino el inventario completo de vocales de un hablante de forma que el contexto de habla se expande (este modelo parte de los experimentos de Ladefoged y Broadbent [1957] en los que tras bajar el F1 de la frase precursora el oyente identificaba la vocal /i/ como /e/ como si por contraste el F1 de la vocal fuera más alto y al contrario, al subir el F1 de la frase anterior la vocal /æ/ fue percibida como /e/, con un F1 más bajo).

Una pregunta que subyace una vez que sabemos que el niño puede estar apoyándose en la relación de los valores formánticos para identificar la vocal, es si este lo hace fijándose en el estado estable de esta o con ayuda de las transiciones formánticas características de los sonidos consonánticos que las acompañan. La escasez de trabajos en español nos lleva a citar el antiguo estudio de Borzorne de Manrique (1977) según el cual las vocales españolas son prácticamente igual de identificables cuando aparecen aisladas y cuando aparecen en el contexto /p-V-s/ no siendo completamente necesarias las transiciones formánticas para su identificación. Los resultados fueron de un 97% y un 99% de correcta identificación respectivamente frente a las vocales inglesas cuya identificación en la condición de aisladas fue del 58% según el estudio de Strange et al. (1974), frente al 83% dado por Verbrugge et al. (1974) para las mismas vocales en el contexto consonántico. Sin embargo, hay que tener en cuenta que dicho estudio fue realizado con estímulos del español de Argentina y que los oyentes eran estudiantes no graduados.

Si queremos disponer de trabajos que experimenten con niños pequeños debemos acudir a la lengua inglesa y en concreto al trabajo de Sussman (2001) en el que se probó la capacidad de once niños con desarrollo lingüístico normal, cuya media de edad era de 4;11 años, y de once niños de 6;2 años con cierta discapacidad lingüística, de identificar una vocal a partir de cuatro estímulos diferentes en los que se combinaban la parte estable del formante con las transiciones formánticas. Los estímulos fueron los siguientes: (1) Estímulo congruente compuesto por las sílabas completas [bib] y [bæb]; (2) Estímulo formado simplemente por el estado estable de la vocal; (3) Estímulo creado a partir de las transiciones de dentro y de fuera de las sílabas pero que contenían periodos de la vocal (estímulo llamado *vowelless*); (4) Estímulo contradictorio formado por las transiciones de [bib] y el estado estable de la vocal de [bæb]. Para que los niños se familiarizaran con los estímulos, se les presentó una historia sobre una hermana y un hermano llamados *Beeb* y *Baeb* y estos debían nombrar en voz alta cada estímulo después de haberlo oído. Los resultados revelaron que los niños identificaron las vocales de manera más precisa en los estímulos de estado estable y en los estímulos congruentes en un rango de entre el 94% y el 96%, manifestando por tanto que los niños utilizaban el estado estable y las

transiciones formánticas prácticamente por igual para la identificación de la vocal (el porcentaje de correcta identificación para la sílaba [bib] en los niños de desarrollo normal fue del 97%, y del 92% aproximadamente para los niños con la discapacidad lingüística; para [bæb] el porcentaje de correcta identificación para los niños con desarrollo normal fue del 93% aproximadamente y del 95% para los niños con la discapacidad. En cuanto a los resultados del estímulo del estado estable, el porcentaje de correcta identificación tanto para la [i] como para la [æ] y tanto en los niños de desarrollo normal como en los que presentaban alguna discapacidad lingüística fue del 95% aproximadamente). Los niños con desarrollo normal y los que presentaban cierta discapacidad lingüística se apoyaron en índices del estado estable para identificar la vocal en los estímulos contradictorios (en un 87% y un 79% respectivamente). Tan solo para la condición de *vowelless* los infantes parecieron apoyarse más en las transiciones a la hora de identificar la vocal y lo hicieron en un 89% y un 83,5% respectivamente.

De esta forma, comenta Susman (2001), los resultados no apoyan la hipótesis de Nittrouer y Miller (1997) según la cual los niños pequeños se apoyan más en las transiciones formánticas para la percepción mientras que los adultos y los niños mayores dan más peso a las propiedades acústicas del estado estacionario. Quizás esto se deba a que dicha teoría estaba basada en la percepción de consonantes fricativas y no había sido probada para la percepción de las vocales. Esto nos lleva a pensar según Susman (2001) que los niños pueden usar diferentes estrategias de percepción según se trate de vocales y consonantes. En el mismo artículo de Susman (2001) se cita el trabajo de Ohde et al. (1996) en el que tras varios experimentos con vocales de corta duración estos encontraron que para los niños de entre 5;0 y 11;0 años el movimiento de las transiciones formánticas no mejoraba significativamente la percepción de las vocales.

Sin embargo y al igual que Nittrouer y Miller (1997), otros estudiosos como Morrongiello et al. (1984), Nittrouer y Studdert-Kenney (1987) y Nittrouer (1992, 2005) creen que los niños identifican mejor la vocal cuando esta va acompañada de consonantes, pudiéndose apoyar así en las transiciones formánticas. En el trabajo Nittrouer (2005) este demostró que niños de 3;0, 5;0 y 7;0 años identificaron mejor las vocales /æ/ y /ʊ/ cuando estas aparecían en el contexto /dæd/ y /dʊd/ que cuando aparecían aisladas.

En el caso de los adultos, autores como Bischoff (1976), Shankweiler et al. (1977), Strange et al. (1979), Gottfried y Strange (1980) y Howell (1981) creen que la vocal se identifica mejor cuando va inserta entre consonantes. Los experimentos de este último combinaron no solo la identificación de la vocal cuando esta aparecía aislada y entre consonantes, sino precedida y seguida de un tono. En el experimento 1 se analizó en qué

condición de las siguientes: diptongos aislados ([ei] y [au]), diptongos precedidos del fonema /b/, y tono de un 1Khz más diptongo, podía identificar mejor la vocal estudiantes de entre 19 y 25 años. Los resultados mostraron que los oyentes cometieron muchos más errores a la hora de reconocer la vocal aislada que cuando esta iba precedida de una consonante o de un tono (en cuyos casos los estudiantes cometieron el mismo número de errores). En el experimento 2 aunque se cambiaron los diptongos (por los de [iə] y [eə]) y la posición del tono (que ahora aparecería inmediatamente después de la vocal) los resultados fueron los mismos. Finalmente el experimento 3 en el que los oyentes tuvieron que identificar nueve vocales ([i, ɪ, ε, æ, ɑ, ɔ, ʌ, ʊ, u]) aisladas y en el contexto /p-p/, los resultados volvieron a mostrar la mejor identificación de la vocal por parte de los oyentes cuando esta iba entre consonantes.

Por el contrario, estudiosos como Diehl et al. (1981) demostraron que doce estudiantes graduados reconocieron mejor la vocal en la condición de aislada que cuando iba acompañada de consonantes. No obstante, como las muestras de habla eran estímulos sintéticos, decidieron llevar a cabo un tercer experimento con muestras de habla natural en el que además hicieron que los estudiantes señalaran de dos formas diferentes la vocal que habían reconocido. Sus hallazgos sugirieron que era la tarea específica de identificación de la vocal la que podía favorecer que esta se escuchara mejor entre estímulos consonánticos, sin que esto se debiera a factores propiamente perceptivos.

### **2.2.2. Influencia de la lengua ambiente.**

El segundo factor que interviene en el desarrollo de la percepción y la producción del bebé según Polka et al. (2007) es “The language(s) to which infants are exposed” y este no es otro que el de la lengua o lenguas a las que está expuesto el bebé desde que se encuentra en el útero y que para muchos autores condicionarán sus producciones y el desarrollo de sus habilidades perceptivas. Se trata de la conocida controversia entre la hipótesis de la independencia defendida por Lenneberg (1967), McNeill (1970), Oller (1978), Stark (1980), Lindblom (1984), Stoel-Gammon y Cooper (1984), Holmgren et al. (1986), Koopmans-van Beinum y van der Stelt (1986), Studdert-Kennedy (1986), Vihman y Miller (1988), Roug et al. (1989), Locke y Pearson (1992), Locke (1993), Vihman (1996), Oller (2000) y Stokes y Wong (2002), entre otros, y la hipótesis de la interacción que queda reafirmada en los experimentos de Boysson-Bardies et al. (1984), Boysson-Bardies et al. (1986), Boysson-Bardies et al. (1989), Boysson-Bardies et al. (1992), Ruzza et al. (2003) y Chen y Kent (2005). Según la primera, la similitud encontrada en las emisiones de niños criados en diferentes ambientes lingüísticos lleva a predecir que son restricciones de tipo universal como las que tienen que ver con la configuración anatómica y el control articulatorio, las que guían las producciones de los bebés

en la etapa prelingüística sin que en esta influya la lengua a la que estos son expuestos. Esta es la opinión de Locke y Pearson (1992:94) que en artículo sobre el aprendizaje vocal del bebé sostienen que

“[...] that over the first year, infant’s sounds (and therefore, sound-making movements) are relatively insensitive to specific ambient experience. Moreover, many of the sounds that appear frequently in babbling resemble phonemes that occur in the majority of the world’s languages”.

También es la conclusión a la que llegan Stokes y Wong (2002:613) a partir de un estudio sobre el desarrollo de la vocal y el diptongo en niños hablantes de cantonés cuando dicen que solo a partir de los 2;0 años el efecto lingüístico comienza a ejercer más influencia que el efecto articulatorio “[...] By 24-27 months, the linguistic factor appears to have more influence than the articulatory factor”. Además según esta hipótesis, los componentes perceptivos y motores se desarrollan de manera independiente de forma que lo que percibe el bebé no influye en lo que este produce.

Para la hipótesis interaccional sin embargo, los gestos articulatorios son guiados desde la etapa del balbuceo por las capacidades perceptivas del bebé que están orientadas a la lengua ambiente (Boysson-Bardies et al. [1984], Boysson-Bardies et al. [1986], Boysson-Bardies et al. [1989] y Boysson-Bardies et al. [1992]). La influencia de la percepción en la producción puede observarse según Kent (1984) en los estudios de niños con problemas de audición en los que el balbuceo suele retrasarse y los sonidos que lo integran son diferentes a los de un niño con audición normal. En el trabajo de Kuhl y Meltzoff (1996), en el que profundizaremos un poco más adelante, se nos dice que hay evidencias de que la experiencia de oírse a uno mismo y a los demás contribuye al desarrollo de la producción del habla. Además, las autoras citan los trabajos de Oller et al. (1985), Kent (1987) y Oller y Eilers (1988) en los que se ha demostrado que los bebés sordos producen vocalizaciones diferentes de las de los bebés normales y que estas difieren además en la duración y en el *timing*. A continuación sintetizaremos brevemente algunos de los trabajos en los que se ha detectado tanto en percepción como en producción la influencia en el bebé de la lengua que le rodea.

En cuanto a la percepción, y aunque se ha sugerido que algún tipo de experiencia previa es necesaria para que los niños puedan poner en marcha las capacidades de discriminación (Butterfield y Cairns [1974]), la mayor parte de los trabajos en los que se han investigado las capacidades perceptivas del bebé en relación a lenguas no nativas (como los de Trehub [1976], Holmberg et al. [1977], Eilers et al. [1978], Eimas [1978], Morse [1978], Aslin et al. [1981] y Werker et al. [1981]) han demostrado que el bebé nace con la capacidad de

discriminar cualquier contraste fonético universal independientemente de la lengua ambiente que le rodee y sin haberlo por tanto escuchado previamente.

Uno de los estudios ya citados que ejemplifica dicha capacidad es el de Werker et al. (1981) en el que se probó la habilidad de bebés de 0;6 años de edad hablantes de inglés de distinguir dos contrastes consonánticos procedentes del hindi: el existente entre la sorda no aspirada retrofleja (/ʈa/ o /ʈa/) versus la oclusiva dental (/ta/), en el que el rasgo diferenciador es el punto de articulación, y el dado entre la oclusiva dental aspirada sorda (/tʰ/) versus la oclusiva dental sonora (/dʰ/) diferenciadas en el VOT (de unos de 131,3 milisegundos para /tʰ/ frente a los -121,5 ms para /dʰ/). El procedimiento que se empleó fue el de VRISD (*Visual Reinforced Infant Speech Discrimination*) más conocido como *Head Turning* o “giro de cabeza” que hace que el infante gire la cabeza hacia un altavoz cuando se produce un cambio de estímulo y en un tiempo determinado. Los resultados mostraron que el bebé a los 0;6 años de edad discriminó ambos pares de sonidos de acuerdo con la categoría fonémica.

Esta capacidad de distinguir contrastes no nativos nunca antes escuchados viene a corroborar la ya mencionada tesis de Kent (1981) de que las habilidades de percepción del niño van un paso por delante de las de la producción, tanto en el feto como cuando el periodo de gestación llega a su fin y nace el bebé, ya que este es capaz de percibir contrastes no nativos sin poder llegar a producirlos. Ello muestra tal y como recoge Trehub (1976:468-469), en cuyo experimento los bebés de entre 0;1 y 0;4 años fueron capaces de distinguir la sílaba [pa] de [pã] y la sílaba [za] de [řa] aumentando o disminuyendo la ratio de succión, que el desarrollo de ambas habilidades en los albores de la vida difiere: “Infants’ ability to discriminate the foreign contrasts of the present study highlights the divergent developmental processes which characterize speech perception and production, (Jakobson [1968])” .

En el mismo trabajo, Trehub (1976) acogió la propuesta que planteó McCaffrey (1971) para explicar esta asimetría. Según este autor, esta divergencia se da porque los fonemas como conjuntos de rasgos distintivos pueden ser diferenciados a partir de un solo rasgo mientras que por el contrario solo pueden ser producidos si el bebé domina la articulación de todas las propiedades que lo constituyen. De ahí se desprende según McCaffrey (1971) que ambos procesos manifiesten patrones diferentes de desarrollo.

A pesar de estas afirmaciones ninguno de los autores citados niega la relación existente entre el componente perceptivo y motor sino que más bien se limitan a afirmar que uno está más avanzado que otro. El propio Kent (1984:R890) que un principio sostiene que las habilidades de percepción y producción están inicialmente en gran medida separadas, declara que enseguida estas comienzan a coordinarse y a integrarse con la lengua ambiente:

“[...] Production and perception capabilities that ultimately lead to speech are initially largely separate, but they begin to be coordinated (integrated) within the first few months of life. The integration of the two systems also interacts with the child’s linguistic background, such that the child’s exposure to the sounds around him/her eventually influences the child’s own pattern of vocalization”.

No obstante, la capacidad de diferenciar contrastes vocálicos o consonánticos no nativos va disminuyendo a lo largo del desarrollo ya que experimentos con adultos confirman que estos tienen dificultades o incluso no pueden reconocer contrastes fonéticos que ellos no usan fonémicamente en su lengua nativa. Algunos de los estudios que han comparado las habilidades de bebés y adultos son por ejemplo el ya citado trabajo de Trehub (1976) en el que diez estudiantes de Psicología de la Universidad de Toronto fueron incapaces de distinguir el contraste de estridencia existente entre [za] y [ʃa] mientras que bebés de entre 0;1 y 0;4 años sí lo hicieron (técnica de ratio de succión). También el experimento de Werker et al. (1981) mostró que ninguno de los adultos hablantes de inglés en los que se probó el contraste del hindi /Ta/ /ta/ (la retrofleja sorda no aspirada versus la oclusiva dental donde el lugar de articulación es la dimensión crítica) fue capaz de reconocerlo y que solo unos pocos de ellos lograron observar la diferencia entre la oclusiva dental aspirada sorda y la oclusiva dental sonora *breathy* (/t<sup>h</sup>/ vs /d<sup>h</sup>/). De manera opuesta, todos los bebés hablantes de inglés de entre 0;6 y 0;7 años discriminaron el contraste sin problemas. Esto sugiere según Werker et al. (1981: 349) “[...] the possibility of a change or decline during development in some linguistic-perceptual capacities, which results in decreased adult ability to discriminate previously ‘nonrelevant’ features”.

Otros trabajos como los de Singh y Black (1966), Lisker y Abramson (1970), Goto (1971), Miyawaki et al. (1975), MacKain et al. (1980), Werker y Lalonde (1988), Best y McRoberts (1989) y Dietrich et al. (2007) también indican que hay un descenso de las capacidades perceptivas iniciales del bebé que en su desarrollo hacia la edad adulta muestra dificultades o es incapaz de reconocer contrastes de sonidos que no pertenecen a su lengua ambiente. Sin embargo, el estudio de Polka y Bohn (1996) no proveyó evidencia del efecto de la edad ni de la lengua en la discriminación de contrastes no nativos. En este trabajo veinte niños monolingües de inglés de entre 0;6 y 0;8 años y de entre 0;10 y 1;0 años de edad tuvieron que discriminar mediante el procedimiento de giro de cabeza el contraste alemán /u/-/y/ que no es fonémico en inglés, y veinte niños que adquirirían el alemán tuvieron que diferenciar el contraste inglés /ɛ/-/æ/, que no es fonémico en alemán (siguiendo el mismo procedimiento). Los resultados indicaron que no hubo un descenso con la edad asociado a la lengua ambiente que les dificultara el reconocimiento del contraste no nativo entre los 0;10 y el 1;0 año.



Si asumimos tal y como apuntan la mayoría de los estudios citados que dicha disminución de la destreza perceptiva con la que nace el bebé se produce a lo largo de su desarrollo ¿a qué edad comienza a notarse un deterioro en dichas habilidades? O, ¿en qué momento el bebé pierde por completo dicha capacidad? Los trabajos de Werker y Tees (1983, 1984) han intentado averiguar el tiempo concreto en el que se produce la pérdida al experimentar con bebés de diferentes edades para comprobar si estos eran capaces de distinguir un mismo contraste consonántico. En Werker y Tees (1984) los autores quisieron probar el contraste consonántico del thompson (lengua nativa indígena del oeste de Canadá) de glotalizada velar frente a glotalizada uvular acompañadas de la vocal [i], esto es /k'i/ frente a /q'i/, en niños de 0;6 años, de entre 0;8 y 0;10 años, y de entre 0;10 y 1;0 años de edad mediante el procedimiento de giro de cabeza hacia el altavoz cuando se producía un cambio en la categoría. Los resultados mostraron que la mayoría de los bebés de 0;6 años discriminaron el contraste del thompson mientras que solo unos cuantos de entre 0;8 y 0;10 años lo hicieron y frente a los de entre 0;10 y 1;0 años que diferenciaron el contraste de una forma pobre equiparable al modo en que lo harían los niños mayores y los adultos. Puede concluirse por tanto que entre los 0;8 y los 0;10 años hay un descenso de las habilidades perceptivas del bebé que se acentúa entre los 0;10 y los 1;0 año de edad.

El trabajo de Mattock y Burnham (2006) también indicó que se produce un descenso de las habilidades perceptivas con las que nace el bebé pero en este caso entre el sexto y el noveno mes de vida. En dichos meses los niños ingleses con los que se experimentó mostraron un descenso en la discriminación de contrastes tonales no nativos no encontrados en niños chinos de la misma edad (porque estos últimos utilizan dichos contrastes para distinguir palabras dado que el chino es una lengua tonal).

En un trabajo anterior Werker y Tees (1983) experimentaron con los mismos contrastes consonánticos del hindí utilizados en Werker et al. (1981) (recordemos que eran /t̪a/-/ta/ y /tʰ/-/dʰ/) pero en niños de mayor edad: 4;0, 8;0 y 12;0 años. De nuevo el efecto de la edad se sintió ya que los niños de 4;0 años realizaron la discriminación de manera más pobre que los bebés de entre 0;6 y 1;0 años, mientras que los niños de 8;0 y 12;0 años mostraron el mismo nivel de discriminación que los adultos. Los resultados de este experimento rechazan según Werker y Tees (1983) la hipótesis de Lennerberg (1967) de que la flexibilidad perceptiva del bebé es mantenida hasta la pubertad (de forma que los niños que aprenden una segunda lengua lo harán libres de acento si la aprenden antes de dicha edad), ya que los niños muestran un descenso en determinadas habilidades lingüístico-perceptivas antes de la adolescencia. Por otro lado, estos resultados coinciden con la afirmación de Burnham (1986) de que hay dos periodos de pérdida

de las destrezas perceptivas con las que nace el bebé: la segunda mitad del primer año de vida y aproximadamente entre los cuatro y los ocho años de edad.

Una vez explicado el descenso de las capacidades perceptivas que sufre el bebé al final del primer año de vida en cuanto a la diferenciación de contrastes consonánticos que no pertenecen a su lengua ambiente, cabe preguntarse si sucede lo mismo con las vocales. Muy pocos son los trabajos que han probado la habilidad de los niños para reconocer contrastes vocálicos no nativos y muchos menos son los que han intentado averiguar el efecto que tiene la edad en dicha destreza.

Dentro de esta escasez encontramos el estudio de Polka y Werker (1994) en el que se probó la capacidad de veinte bebés de entre 0;6 y 0;8 años de edad, y de catorce bebés de entre 0;10 y 1;0 año de edad de discriminar dos contrastes vocálicos del alemán insertos en el contexto de /d+vocal+t/, que incluía el contraste de las laxas alta anterior redondeada versus la alta posterior redondeada, esto es /dYt/ versus /dUt/, y el contraste entre las tensas alta anterior redondeada versus la alta posterior redondeada, esto es /dyt/ versus /dut/, mediante la técnica de giro de cabeza. Se observó que una proporción más alta de niños de entre 0;6 y 0;8 años de edad en comparación con los de entre 0;10 y 1;0 años de edad, pudo discriminar el contraste vocálico (menos del 20 % de los niños de entre 0;10 y 1;0 años realizaron el contraste con éxito). De este modo, podemos concluir que también en el caso de las vocales la edad va mermando las habilidades perceptivas con las que nace el bebé.

No obstante, Polka y Werker (1994) decidieron comprobar si el efecto de la edad ocurría al mismo tiempo en las vocales que en las consonantes y para ello realizaron el mismo experimento pero con bebés de 0;4 y de 0;6 años de edad (debido a la corta edad los mismos hubieron de cambiar el procedimiento de giro de cabeza por el de habituación y deshabituación de la mirada). Los resultados mostraron que aunque ambos grupos podían diferenciar los dos contrastes, se había producido un descenso en la discriminación en el grupo de niños de 0;6 años de edad. Según los autores este descenso entre los 0;4 y los 0;6 años de edad indica que el periodo comprendido entre los 0;6 y los 0;8 años marca un paso intermedio en el curso del descenso general en la discriminación de los contrastes vocálicos. Además los resultados muestran según Polka y Werker (1994:432) que dicha disminución ocurre antes para las vocales que para las consonantes:

“[...] Thus, the overall pattern of change reveals a shift in infant vowel discrimination from a language-general to a language specific pattern, consistent with previous studies of cross-language consonant perception but beginning earlier in development”.

Una atenuación de la percepción auditiva asociada a la edad fue encontrada también en el trabajo de Bosch y Sebastián-Gallés (2003a) en el que se estudió a niños de diferentes edades y hablantes de distintas lenguas: español, catalán y bilingües de español-catalán. Usando el procedimiento de habituación-deshabituación de la mirada los resultados indicaron que a los 0;4 años de edad todos los niños distinguían perfectamente el contraste /e/-/ɛ/ pero dicha habilidad descendió en los bebés de habla española entre el sexto y el octavo mes de vida debido a que el español carece del sonido /ɛ/. No ocurrió lo mismo para los niños monolingües de catalán que siguieron discriminando ambos contrastes con la misma precisión. Diferente es el caso de los niños bilingües quienes reconocieron el contraste a los 0;4 años de edad, fallaron al identificarlo a los 0;8 años, y volvieron a discriminarlo al 1;0 año manifestando así un desarrollo en forma de U.

Lo mismo ocurrió en el trabajo de Fennell et al. (2010) en el que se probó la capacidad de niños monolingües de inglés, de francés y de niños bilingües en estas dos lenguas, de distinguir el contraste del francés [y] vs [œ]. A los 0;4 años los tres grupos podían diferenciarlo pero a los 0;8 años solo lo hicieron los monolingües de francés mostrando que una reorganización de las capacidades perceptivas ya se había puesto en marcha en los niños hablantes de inglés. Al igual que en el estudio de Bosch y Sebastián Gallés (2003a), los niños bilingües volvieron a discriminar el contraste al 1;0 año mostrando una evolución en forma de U. Es en palabras de Polka et al. (2007) como si en los niños en contacto con la lengua ambiente se instalara un filtro que les permite atenuar o facilitar el acceso a diferentes categorías fonéticas.

Entre las causas que pueden explicar esta disminución, parece claro que cuanto más experiencia tienen los niños con la lengua que les rodea, más influencia ejercerá esta en las capacidades perceptivas de aquellos. Esto es lo que sugiere el trabajo de Zlatin y Koenigsknecht (1976) según el cual la experiencia con una lengua específica actúa para consolidar los límites fonémicos en los niños (pudiendo conllevar el efecto contrario que es el de la pérdida de los límites lingüísticos no nativos). También para Werker y Tees (1984:62) este descenso es una función de la experiencia con la lengua específica y no es accidental que ocurra alrededor de la edad en la que el niño está empezando a entender y producir sonidos de su lengua nativa sino que podría esperarse que “[...] this perceptual reorganization is closely related to the acquisition of phonological contrasts”. Además, según los autores esta pérdida selectiva de los límites entre las categorías facilitaría al niño el aprendizaje de una lengua específica poniendo a punto las sensibilidades iniciales con una fonología específica. Para otros autores como Maye et al. (2002), tal y como se nos dice en Bosch y Sebastián-Gallés (2009), la distribución estadística

de rastreo de los sonidos del habla en el input que recibe el bebé está detrás de esta reorganización perceptiva que ocurre al final del primer año de vida.

Esta falta de capacidad a la hora de reconocer contrastes no nativos en la edad adulta es según Werker y Tees (1984) más que a una pérdida neuronal en sí, un cambio de la atención del bebé o un cambio en las estrategias de procesamiento. Para sostener esta tesis, Werker y Tees (1984) parten por un lado del estudio de Walley et al. (1981) en el que se sugirió que a pesar de que los oyentes dejan de atender a distinciones fonéticas que no poseen significado en su lengua nativa la capacidad cortical para responder a dichas distinciones no se ha perdido por completo, y por otro lado se apoyan en el trabajo de MacKain et al. (1980) en el que se demostró que adultos japoneses para los que /r/ y /l/ son percibidas como el mismo fonema, fueron capaces de distinguirlos con un entrenamiento previo en un contexto conversacional (ocho horas al día durante más de un año). Otros estudios en los que según Werker y Pegg (1992) los adultos pueden ser entrenados para discriminar prácticamente cualquier contraste no nativo son los de MacKain et al. (1981), Pisoni et al. (1982), Tees y Werker (1984), Morosan y Jamieson (1989) y Logan et al. (1991) que manifestaron que el cambio en el desarrollo no es irreversible.

Para confirmar estos resultados que señalan que no hay una pérdida neuronal, Werker y Tees (1984) llevaron a cabo un experimento en el que hablantes de inglés de entre 18 y 40 años fueron probados en el ya mencionado contraste del hindi /Ta/ vs /ta/) y en el del thompson (/ki/ vs /qi/). Los resultados revelaron que más que una pérdida neuronal, atrofia sensorial o un cambio en el foco de atención, lo que se produce es una diferente estrategia de procesamiento fonémico que hace que el niño según va creciendo atiende más a los contrastes fonémicos de su lengua nativa pero que al no haber una pérdida neuronal, siempre que se requiera este puede modificar dicho procesamiento fonémico para atender a contrastes no nativos.

Otra influencia que ejerce la lengua ambiente en la percepción de los bebés, y frente a las asimetrías ya explicadas de Polka y Bohn (2003) y Schwartz et al. (2005), tiene que ver con las asimetrías perceptivas que estos presentan a la hora de discriminar entre sonidos prototípicos y no prototípicos de una misma categoría vocálica. Se trata de la teoría del efecto magnético perceptivo de Kuhl (1991) según la cual el bebé discriminaría mejor en la dirección que va del ejemplar no prototípico de una vocal al ejemplar prototípico de esta, que en la ruta opuesta. En cuanto a la influencia de la lengua ambiente, esta puede observarse en el hecho de que los niños discriminan mejor cuando en la dirección de vocal no prototípica a prototípica esta última resulta ser una vocal nativa que cuando es un prototipo no nativo. Esto se ve en el experimento que se lleva a cabo en Kuhl et al. (1992) en el que niños de 0;6 años que adquirían inglés americano y sueco fueron instados a diferenciar mediante el procedimiento de giro de cabeza

entre un conjunto de vocales alrededor de un ejemplar prototípico de la vocal inglesa /i/ y de la sueca /y/ (para los adultos ingleses la sueca /y/ no es prototipo de ninguna vocal inglesa y para los adultos suecos la inglesa /i/ no lo es de ninguna vocal sueca). Los resultados indicaron que los infantes del inglés discriminaron mejor cuando la dirección fue de las vocales periféricas (no prototípicas) a la vocal central (prototípica) cuando la vocal central fue la nativa (prototípica) inglesa /i/ que cuando la vocal central fue la no nativa (no prototípica) /y/. Los niños suecos mostraron el patrón opuesto siendo el efecto magnético perceptivo más fuerte para la vocal sueca que para la vocal inglesa. La lengua ambiente actúa aquí como un imán perceptivo.

Ocurriría lo mismo si el niño tuviera que discriminar entre una vocal no nativa que se acercara a un prototipo de una vocal nativa, y una vocal no nativa que no se asemejara a ninguna categoría vocálica de la lengua ambiente tal y como se demuestra en los experimentos de Polka y Werker (1994) y de Polka y Bohn (1996). Como ya hemos explicado, en el trabajo de Polka y Werker (1994) se probó la capacidad de niños de 0;6 años, de entre 0;8 y 0;10 años y de entre 0;10 y 1;0 año de edad de distinguir mediante el procedimiento de giro de cabeza dos contrastes vocálicos del alemán, /dYt/ versus /dUt/ y /dyt/ versus /dut/, en los que la vocal alta posterior en ambos contrastes fue reconocida por hablantes adultos de inglés como más similar a otra categoría vocálica inglesa resultando por tanto en la vocal prototípica. Esto hizo que los bebés de 0;6 años discriminaran mejor el contraste en la dirección de /dYt/ vocal menos prototípica (o NP) a /dUt/ vocal más prototípica (o P), y en la dirección de /dyt/ (vocal NP) a /dut/ (vocal P), que en las rutas opuestas dando muestras de que una asimetría perceptiva o un efecto de imán perceptivo estaba ocurriendo. Los niños de entre 0;8 y 0;10 años de edad y los de entre 0;10 y 1;0 año discriminaron de manera pobre (y semejante entre ellos) en ambas condiciones (tanto de NP a P como de P a NP), algo que puede explicarse porque a esa edad se intensifica el descenso de las capacidades perceptivas iniciales para distinguir contrastes consonánticos. Sin embargo, tampoco los bebés de entre 0;4 y 0;6 años de edad en el experimento segundo, en el que además se cambió el procedimiento de giro de cabeza por el de habituación-deshabituación, mostraron el efecto de imán perceptivo sin que pueda esto ser explicado por una acuse del cambio de estrategia de procesamiento perceptivo ya que los niños están en la edad en la que este es menos notorio. Esto llevó a Polka y Werker (1994) a pensar que quizás el efecto magnético perceptivo inaugurado por Kuhl fuera algo impreciso y solo se mostrara bajo determinadas condiciones de prueba (como en la del giro de cabeza con los niños de 0;6 años y no bajo la del procedimiento de habituación-deshabituación con los niños de entre 0;4 y 0;6 años).

En el también ya mencionado estudio de Polka y Bohn (1996) se muestran asimetrías perceptivas en la diferenciación de contrastes no nativos. Así, de las dos vocales que conforman

el par alemán que el bebé inglés ha de discriminar (recordemos que mediante el procedimiento de giro de cabeza), la vocal /u/ se asemeja más a una vocal inglesa con lo que se constituiría como vocal prototípica. De la misma forma, de las dos vocales dadas en el contraste inglés, la /ɛ/ es más similar a una vocal alemana de lo que lo es la /æ/ y por tanto actuaría como vocal que atrae perceptivamente. Después del experimento se demostró que los niños ingleses discriminaban mejor el contraste alemán cuando en este la dirección era de /y/ a /u/, actuando /u/ como imán perceptivo, y que los niños alemanes diferenciaban mejor el contraste inglés cuando la ruta era de /æ/ a /ɛ/ actuando esta última como imán perceptivo. No obstante, esta asimetría fue observada tanto en los niños de entre 0;6 y 0;8 años como en los de entre 0;10 y 1;0 año a pesar del descenso de las capacidades perceptivas que en teoría ocurre a esta edad.

En lo que respecta a la influencia de la lengua ambiente en las producciones de los bebés, trabajos de muy diversa índole (a nivel articulatorio, acústico y hasta prosódico) han demostrado que esta se produce. Entre los estudios que muestran el influjo ambiental en el campo articulatorio, encontramos el de Benner et al. (2007) en el que se estudió la constricción laríngea en niños hablantes de inglés canadiense, de árabe de Marruecos y de la lengua china bai. Diecinueve niños fueron grabados de una a tres veces durante el primer año de vida y se extrajeron todas las emisiones de CVC. Para todos los niños, algunos marcos de constricción laríngea (que se dan en la voz chillona o áspera, la voz *creacky*, la voz susurrada y la del cuchicheo) predominaron en los cinco primeros meses de vida debido a la posición alta de la laringe. Esta predisposición a producir sonidos constreñidos laríngeamente disminuyó a partir del quinto mes debido al descenso de la laringe y al mayor control articulatorio de los niños sobre los parámetros laríngeos que hizo que estos produjeran marcos no constreñidos (ya sea voz modal, *breathy* o falsete). Los resultados mostraron que el balbuceo de los niños arábigos fue más constreñido laríngeamente que el de los niños ingleses mostrando el balbuceo de los niños hablantes de chino el patrón de constricción laríngea más complejo.

El peso de la lengua input se deja ver también en el inventario de sonidos vocálicos y consonánticos de las emisiones de los bebés que si bien siguen una tendencia universal, ya que se han encontrado fuertes similitudes en lenguas diferentes en los tipos de sonidos y en la combinación de estos, también muestran patrones ambientales específicos. Antes de comenzar con la revisión de los principales estudios subrayamos el hecho de que en todos ellos un investigador recogió las emisiones espontáneas del bebé que se produjeron en la interacción de este con su madre o algún otro miembro de la familia y que tuvieron lugar en la casa familiar, en el laboratorio o en la guardería, pero siempre recreando situaciones habituales y cotidianas para el bebé como son los momentos de juego, de comida, etc. Por ello, no concretaremos esta metodología en cada uno de los trabajos y tan solo la especificaremos cuando dichas

circunstancias sean diferentes. Todos los sujetos, por otro lado, mostraron un desarrollo normal sin ningún problema físico, mental o sensorial.

El estudio Boysson-Bardies et al. (1980), en el que se grabó a un niño hablante de francés entre los 1;6 y los 1;8 años de vida, mostró que el porcentaje de consonantes fricativas producidas por el niño era tan alto como el porcentaje de oclusivas, aproximándose así al patrón de la lengua francesa y difiriendo de la ratio de sonidos fricativos encontrada en el balbuceo de niños hablantes de otras lenguas como el inglés (Winitz e Irwin [1958], Moskowitz [1970] y Ferguson [1963]). De este modo, el porcentaje de consonantes fricativas sordas que produjo Sebastian fue del 8,97% siendo la ratio de fricativas en francés del 7,97% frente al porcentaje de fricativas del inglés o de la lengua thai, con las que también se compararon los datos de este estudio, que fue de 3,47% y 3,17% respectivamente. Por tanto y según los autores, una adquisición “fonética selectiva” específica de la lengua tiene lugar durante la etapa del balbuceo. En el mismo trabajo, Boysson-Bardies et al. (1980) aludieron al estudio de Weir (1966) en el que el balbuceo de niños americanos fue discriminado del balbuceo de niños chinos.

Una década más tarde, Boysson-Bardies y Vihman (1991) compararon la frecuencia de aparición de las consonantes emitidas en el balbuceo y en las primeras palabras en niños de entre 0;9 y 1;5 años hablantes de francés, inglés, japonés y sueco, con la frecuencia de aparición de esas mismas consonantes en las lenguas meta de dichos niños. Los resultados indicaron diferencias significativas en el modo y lugar de articulación de dichas consonantes según la lengua ambiente a la que pertenecía cada grupo. Un año más tarde Levitt y Utman (1992) equipararon el balbuceo de un niño criado en un ambiente inglés con el de otro crecido en un ambiente francés tras registrar sus emisiones en sesiones semanales de 20 minutos a las edades de 0;5, 0;8 0;11 y 1;2 años. Aunque ambos niños produjeron oclusivas de manera más frecuente, la absoluta frecuencia de producciones oclusivas fue diferente en los dos niños achacándose esta diferencia a la lengua específica a la que estos pertenecían (ambos estudios se citan en Lee et al. [2009]).

También el trabajo de Amayreh y Dyson (2000) mostró que por un lado las producciones consonánticas de trece niños hablantes de árabe entre los 1;2 y los 2;0 años seguían un patrón universal, pero que por otro también se observaban en dichas emisiones tendencias específicas de la lengua ambiente. De este modo, la presencia de las consonantes oclusivas [b, d, t], de la nasal bilabial [m], de la fricativa glotal [h] y de la glide [w] al principio de la palabra es común a los inventarios fonéticos de niños hablantes de otras lenguas como el inglés. Sin embargo, la reiterada aparición de las consonantes [l, j] en las producciones de los

infantes que adquirirían el árabe y no en las de los niños hablantes de inglés (salvo en el estudio de Dyson [1988]), y la escasa presencia del fonema /k/ en las emisiones de los primeros frente a su predominio en las emisiones de los niños ingleses, apuntaba a que la lengua meta ya había empezado a ejercer su influencia en las producciones del bebé a la edad de 1;2 años.

Otra publicación que refleja el influjo de la lengua ambiente en el inventario de sonidos vocálicos y consonánticos es la de Kern et al. (en prensa) en la que se estudian las emisiones vocálicas y consonánticas de niños monolingües de turco, francés, rumano, holandés y árabe tunecino entre los 0;8 y los 2;1 años de edad. En cuanto al modo de articulación, en cuatro de las cinco lenguas las consonantes orales oclusivas fueron las más representadas sugiriendo una tendencia universal de producción. Sin embargo, esta no fue la tendencia para los niños hablantes de árabe tunecino en los que la fricativa glotal fue la categoría más representada. La segunda categoría más común en todas las lenguas fue la de las nasales que no obstante duplicó la media en los hablantes de francés (lengua con vocales nasales) mientras los niños hablantes de holandés y árabe tunecino apenas produjeron la mitad de la media de estos sonidos.

En cuanto al punto de articulación, en general las consonantes coronales fueron las más frecuentes con un 44,4% de apariciones en todas las lenguas. Sin embargo, los niños hablantes de francés produjeron más labiales que coronales y los de árabe tunecino emitieron más radicales que coronales influenciados por el patrón general de sus respectivas lenguas. La segunda categoría más frecuente fue la de las labiales y en general hubo una tendencia a una mayor presencia de dorsales que de guturales. Sin embargo, los niños hablantes de árabe tunecino y de holandés revirtieron dicha querencia y ellos emitieron más guturales que dorsales reflejando una vez más el papel de la lengua ambiente.

En lo que respecta a las vocales, dos o tres vocales dieron cuenta del 50% de los tipos vocálicos siendo la vocal [a] la única que presentó una frecuencia mayor al 5% en todas las lenguas. También de forma común, las vocales anteriores medias, anteriores bajas, centrales medias y centrales bajas fueron las categorías más representadas en los niños de las cinco lenguas siendo las vocales posteriores las menos identificadas. En la dimensión anterior-posterior solo los niños de dos lenguas, los hablantes de francés y de árabe tunecino, siguieron la tendencia general de vocales más anteriores que centrales. En el eje de la altura, las vocales medias fueron las más producidas por los infantes de cuatro de las lenguas constituyendo una excepción los niños hablantes de árabe tunecino que produjeron más vocales bajas que medias. Estos resultados manifestaron una vez más que además de una base universal para la producción de vocales en el balbuceo, la influencia de la lengua ambiente está presente.



Si nos fijamos en las vocales, el trabajo de Oller y Eilers (1981) refleja cierta influencia de la lengua meta en la producción de dichos sonidos. En su estudio se grabaron las emisiones de dieciséis bebés de 1;0 año de edad, ocho de los cuales adquirirían la lengua española y ocho la lengua inglesa. Los autores nos advierten de que pudo introducirse cierta contaminación lingüística en las producciones de estos sujetos debido a que la ciudad de procedencia de los niños era Miami, ciudad que presenta bilingüismo español-inglés, aunque no especificaron el grado ni el contexto. Por otro lado, los mismos autores señalaron que además del habla espontánea surgida entre el niño y la madre, tanto a los niños de habla inglesa como a los de habla española se les presentó un grupo de modelos vocálicos estándar que una vez más quedan sin especificar en este trabajo. Los resultados mostraron que los niños que adquirirían la lengua española producían una mayor cantidad de las vocales que se dan en esta lengua (/a, e, i, o, u/) y un menor número de las vocales típicas del inglés (entre ellas /æ, ɪ, ε, ʊ, ɔ, α, ʌ, ə/). Al contrario, los infantes que adquirirían la lengua inglesa emitían una mayor proporción de estas vocales así como de los diptongos que aparecen en este idioma ([ej], [ow] o [oj]).

Continuando con las vocales, Davis y Davis (2002) llevaron a cabo un estudio en el que compararon la producción vocálica de dos niños que adquirirían portugués de Brasil de entre 1;0 y 3;0 años de edad, con las vocales producidas por niños ingleses en el mismo periodo (extraídas del estudio de Davis et al. [2002]). En este caso el investigador fue el propio padre que al ser lingüista podía realizar la tarea con precisión. Los resultados mostraron que si bien los niños de ambas lenguas mostraban una preferencia por las vocales anteriores y centrales, diferían en cuanto a la altura vocálica. Así, los jóvenes hablantes de portugués de Brasil presentaron una mayor frecuencia de vocales altas frente a los niños de lengua inglesa en cuyas emisiones predominaron las vocales medias y bajas reflejando la influencia de la lengua ambiente.

En el estudio llevado a cabo por Rvachew et al. (2008) los autores informaron de que los niños que adquirirían la lengua inglesa producían más vocales altas y posteriores que los niños que adquirirían francés a la misma edad, esto es, de los 0;8 a los 1;6 años, apoyando así la hipótesis de que la producción vocálica de los niños refleja la influencia lengua ambiente en edades muy tempranas.

Encontramos también trabajos en los que se estudian las producciones de niños procedentes de más de dos ambientes lingüísticos como es el caso del trabajo de Kern y Davis (2009). En este artículo, los autores compararon la producción vocálica de veinte niños que adquirirían turco, francés, rumano, holandés y árabe tunecino entre las edades de 0;8 y 2;1 años. Los resultados mostraron una combinación de la hipótesis universal y de la de la influencia de la

lengua meta al encontrarse en las producciones de los sujetos una presencia de patrones universales, como el predominio en todas las lenguas del cuadrante izquierdo inferior del espacio vocálico debido a una frecuencia del 66% de vocales anteriores bajas y medias y de vocales centrales, y una presencia de patrones concretos de la lengua que adquirirían: los niños monolingües de francés produjeron más vocales altas que los niños de árabe tunecino; las vocales medias fueron más altas en los niños hablantes de francés y turco que en los del resto de comunidades lingüísticas; los niños rumanos registraron la mayor frecuencia de vocales altas centrales, frente a los niños de otras comunidades donde apenas estuvieron presentes.

El título del artículo de Lee et al. (2009) “Universal production patterns and ambient language influences in babbling” evidencia de nuevo la combinación de los dos factores, el universal y el de la influencia de la lengua ambiente, en las emisiones producidas en el balbuceo de los niños. En él se compararon las características fonéticas del balbuceo canónico (concretamente se extrajo el patrón CV) de niños que adquirirían la lengua coreana entre los 0;8 años y el 1;0 año de edad, con la frecuencia de sonidos vocálicos presentes en el habla dirigida a ellos por sus madres (*Infant Direct Speech* o IDS) que reflejaba las características de lengua ambiente. Los datos recogidos se compararon después con los existentes en las bases de datos de balbuceo de niños que adquirirían la lengua inglesa en el mismo periodo como los recogidos por Davis y MacNeilage (1995). Los resultados señalaron que si bien tanto los niños que adquirirían coreano como los que adquirirían inglés mostraron patrones vocálicos comunes al producir frecuencias similares de vocales posteriores bajas y centrales altas, incluso aunque las frecuencias de esas vocales eran diferentes en el IDS inglés y coreano (manifestándose de nuevo la universalidad de patrones vocálicos), las producciones de ambos también se diferenciaron reflejando las pautas de la lengua ambiente. De esta manera, los niños que adquirirían el inglés produjeron más vocales anteriores altas y bajas y centrales medias que los niños que adquirirían el coreano. En contraste estos últimos emitieron más vocales centrales bajas. Esta circunstancia muestra el influjo de la lengua ambiente ya que la lengua coreana presenta un 31% de vocales centrales bajas frente a la inglesa en la que estas vocales ni siquiera aparecen. De la misma forma en el IDS de la lengua inglesa encontramos que las vocales anteriores altas son las segundas más producidas. Por otro lado, las vocales anteriores bajas son fonemas en inglés pero no en coreano por lo que fueron raramente producidas en el IDS coreano y los niños al no escucharlas, no las produjeron de manera tan frecuente como otras vocales. De esta forma, las frecuencias de las vocales anteriores altas y bajas y vocales centrales medias y bajas fueron diferentes en los dos grupos de niños (aunque estas últimas diferencias no fueron estadísticamente significativas).

Para estos autores, las vocales manifiestan las características de la lengua ambiente antes que las consonantes puesto que, entre otras razones, las vocales son producidas con unos patrones formánticos mejor definidos que los de las consonantes de forma que aquellas se mantienen más constantes por un tiempo mayor que estas últimas. Además, según Lee et al. (2009) mientras que sonidos parecidos a los vocálicos se producen prácticamente desde el nacimiento, muchas consonantes al ser más difíciles de articular no son producidas hasta el comienzo del balbuceo cuando aparece la alternancia del cierre y apertura de la boca.

Las tendencias universales y los patrones de la lengua ambiente también se recogen en el inventario vocálico del trabajo de Chen y Kent (2010) en el que las vocalizaciones espontáneas en la estructura CV de veinticuatro niños que adquirirían la lengua del chino mandarín de entre 0;7 y 1;0 año (grupo 1) y de entre 1;1 y 1;6 años (grupo 2) fueron grabadas. A pesar de no haber un acuerdo general en el número de vocales que posee el mandarín, Cheng (1966, 1973) propuso doce vocales superficiales: /i, e, ε, y, i (con dos variantes), ə, a, u, ɤ, o, ɑ/ y una extra [ɔ] incluida en los análisis de Lin (1989). Por otro lado, todos los análisis consideran [a] y [ɑ] como variantes del mismo fonema vocálico, la única vocal baja en mandarín. Los datos mostraron que en ambos grupos las vocales más frecuentes en las vocalizaciones de los niños fueron [ε, ə, a], es decir, las vocales centrales aparecieron más frecuentemente que las vocales anteriores y estas últimas aparecieron más a menudo que las vocales posteriores. Los datos de los niños y los adultos no difirieron en el caso de las vocales anteriores pero sí en el de las centrales y posteriores ya que los niños produjeron más vocales centrales que sus cuidadores reflejando una base universal para la producción del habla debido al alto número de vocales centrales producidas por niños que adquieren lenguas muy diferentes.

En cuanto a la altura vocálica, los niños articularon más vocales bajas y medias que altas aunque el porcentaje de vocales altas producidas por estos estuvo correlacionado con el número de vocales altas producidas por sus cuidadores. De nuevo un patrón universal manifestó la tendencia de los niños a emitir más vocales medias y bajas que altas por la mayor dificultad de producción de estas últimas. Las vocales bajas tuvieron la más alta correlación entre los datos de los niños y los de sus cuidadores.

En lo que respecta a vocales concretas, encontramos que: las producciones de los niños de la vocal baja [a] estuvieron correlacionadas con los datos de sus cuidadores; que tanto los niños como los adultos produjeron más aces que vocales anteriores altas [i,y]; y que a pesar de que los adultos usaron más las vocales anteriores altas [i,y] que la vocal anterior semibaja [ε], esta última fue producida de manera más frecuente por los niños. En conclusión, la producción vocálica de los niños reflejó con más exactitud el habla de los cuidadores en términos de

posterioridad vocálica que en cuanto a la altura de la vocal puesto que los infantes no produjeron tantas vocales altas como las que estuvieron presentes en el habla de los adultos dirigida a ellos.

Por otro lado, Chen y Kent (2010) examinaron si se daban cambios en el desarrollo de la producción vocálica entre estos dos grupos de edad. Ellos observaron que entre los 0;7 y los 1;6 años de edad, la frecuencia de vocales anteriores aumentó del 16% en el grupo 1 al 36% en el grupo 2, y que la frecuencia de vocales centrales decreció de manera significativa del 81% en el grupo 1 al 46% en el grupo 2. Los autores también encontraron que el patrón de distribución de vocales en el grupo 2 se parecía más al patrón de los cuidadores adultos al hallarse en este grupo de edad un 36% de vocales anteriores, un 46% de centrales y un 11% de posteriores, y al presentar el porcentaje adulto un 33% de vocales anteriores, un 47% de vocales centrales y un 20% de vocales posteriores.

En el eje de la altura, las vocales bajas y medias aparecieron de manera más frecuente en ambos grupos de edad (entre ellas se encuentra la vocal [a] que fue la más frecuente incluso en los niños de 0;7 años reflejando el patrón de la lengua ambiente). La vocal anterior semibaja [ɛ] y la vocal central media [ə] aparecieron frecuentemente en el grupo 1 y persistieron en el grupo 2, y las vocales bajas y anteriores medias (semibajas y semialtas) fueron más comunes que las anteriores altas. No obstante, del grupo 1 al grupo 2 se produjo un aumento de la vocal media semialta que fue de un 0% a un 16%.

A partir de los datos ofrecidos puede concluirse que en las emisiones de los niños estudiados por Chen y Kent (2010) están presentes tanto la hipótesis universal como la de la lengua específica. En cuanto a la hipótesis universal, esta se refleja en el hecho de que los niños produzcan con mayor frecuencia la vocal anterior semibaja [ɛ] y la central media [ə] de forma similar a lo encontrado en estudios con otras lenguas como el inglés (los autores compararon estos resultados con los hallazgos de Irwin [1948], Buhr [1980], Kent y Murray [1982] y Kent y Bauer [1985]). Otro resultado atribuible a la hipótesis universal fue el predominio de las vocales bajas y medias sobre el de las vocales altas cuya frecuencia aumentó solo a partir del primer año de vida. Para dar cuenta de este predominio los autores citaron a Kent (1992) quien piensa que ello puede explicarse si se tiene en cuenta que el niño posee un control articulatorio inmaduro y que la lengua solo puede moverse en el eje anterior-posterior elevándose muy poco. Sin embargo, el hecho de que un 40% de las vocales producidas se correspondiera con la vocal baja [a], y el hecho de que dicha frecuencia no se haya encontrado en los estudios llevados a cabo en otras lenguas, indican claramente que un patrón de la lengua específica opera incluso en el grupo más joven.

Además del influjo de la lengua ambiente en los gestos articulatorios y en el conjunto de sonidos tanto vocálicos como consonánticos que produce el bebé, el peso de la lengua específica puede dejarse sentir también en las frecuencias de los formantes de las vocales. Así lo muestra el estudio de Boysson-Bardies et al. (1989) en el que esta influencia se ejerció tanto en el inventario de sonidos vocálicos como en sus valores formánticos. Dicha investigación examinó la influencia de la lengua meta en las vocales extraídas en el balbuceo canónico y variado (que siguieran el patrón CV) de veinte niños de 0;10 años hablantes del francés de París, del inglés de Londres, del cantonés de Hong Kong y del árabe de Argelia (cinco niños por comunidad). En cuanto al tipo de vocal, los resultados indicaron que los niños ingleses producían más vocales anteriores, que los infantes franceses emitían sobre todo vocales centrales medias, centrales bajas los hablantes de cantonés (y algunas posteriores bajas), y posteriores medias los niños que adquirirían el árabe argelino. Las mismas tendencias fueron encontradas para el habla adulta del inglés, francés, cantonés y árabe marroquí (ya que no se pudo encontrar habla adulta del árabe argelino y por la similitud de sonidos se eligió el árabe marroquí). Observamos por tanto que la lengua específica ya ha comenzado a actuar en la lengua de los bebés a los 0;10 años de edad ya que las preferencias vocálicas de estos difirieron según la comunidad lingüística a la que pertenecían.

Por lo que respecta a las frecuencias de los formantes, los valores del F1 de las vocales emitidas por los niños ingleses fueron más bajos que los hallados en los grupos del cantonés y del argelino; sin embargo, la media del F2 fue la más alta de todos los grupos. Estas frecuencias medias de los niños ingleses coincidieron en gran medida con las dadas unos años antes por Kent y Murray (1982). Los valores del F1 para los niños de habla francesa fueron también bajos y la media del F2 estuvo entre la del inglés y la del cantonés. El promedio para las vocales de los niños franceses estuvo más cerca del grupo inglés que de cualquier otro grupo. En los infantes argelinos, los valores del F1 fueron intermedios y los del F2 bajos, tan bajos como los del cantonés. Esto puede mostrar la preferencia por la vocal baja [a] en el habla adulta. Finalmente, los niños del grupo cantonés mostraron una preferencia por las vocales con un alto F1 y un bajo F2. De este modo, la media del F1 fue la mayor de cualquier otro grupo y la media del F2 fue tan baja como la de los niños argelinos.

Estos resultados solo pueden explicarse si tenemos en cuenta, también en el nivel acústico, la hipótesis de la interacción, aunque según Boysson-Bardies et al. (1989) la temprana influencia de la lengua ambiente no es incompatible con las tendencias universales sino que dichas predisposiciones universales como la tendencia a centralizar el espacio vocálico o el predominio de vocales anteriores sobre posteriores, van siendo modificadas por las características específicas de la lengua meta.

Otro artículo sobre la variabilidad de la acústica en el balbuceo a través de las lenguas, y en el que también participa Boysson-Bardies, es el de Boysson-Bardies et al. (1986). En él quedó demostrada la influencia de la lengua meta en las frecuencias de los formantes ya que se encontraron similitudes entre el espectro vocálico de niños de 0;10 años de habla francesa, árabe y cantonesa y el espectro vocálico de los adultos hablantes de cada una de estas lenguas.

El influjo de la lengua meta en las frecuencias de los formantes de los sonidos vocálicos de los bebés aparece también en el trabajo de Rvachew et al. (2006). En este se extrajeron las vocales aisladas y las contenidas en la sílaba CV de las emisiones de veintitrés niños expuestos al francés del Canadá y de veinte expuestos al inglés canadiense entre las edades de 0;10 y 1;6 años para examinar las diferencias en el primer y en el segundo formante a través de las lenguas. Los resultados indicaron que a pesar de producirse una expansión del espacio vocálico en el habla de los niños pertenecientes a las dos comunidades lingüísticas a lo largo del periodo de estudio coincidente con el expuesto en otros trabajos (Buhr [1980], Kent and Murray [1982], Rvachew et al. [1996], Robb et al. [1997] y Gilbert et al. [1997]), expansión que ponía de manifiesto un cierto carácter universal, esta expansión fue más acusada en la esquina o extremo difuso en el caso de los niños franceses y en la esquina o extremo grave en el de los ingleses. Por otro lado, un descenso del F1 se produjo con la edad en los niños que adquirían el francés pero no en los que adquirían el inglés, y un descenso de la media del F2 tuvo lugar con el paso del tiempo en los niños hablantes de inglés y no en los que adquirían el francés. Una vez más la lengua meta ejerció su influjo en varias características o índices acústicos de las tempranas producciones del niño.

Un estudio acústico más reciente es el de Chung et al. (2008) y en él se analizó la adquisición de las vocales por niños de 2;0 y 5;0 años y por adultos hablantes de inglés, coreano, griego y cantonés. Las vocales compartidas por las cuatro lenguas fueron tomadas de distintas palabras meta y en ellas se analizó el primer y el segundo formante. Los resultados mostraron que las vocales ocupaban diferentes lugares en el espacio acústico debido a las diferentes frecuencias de los formantes tanto en los niños como en los adultos (es decir, las vocales que compartían las cuatro lenguas se situaron en puntos distintos de la carta de formantes porque tenían valores formánticos diferentes de acuerdo con cada lengua).

En el artículo de Levitt y Utman (1992), mencionado en el trabajo de Kern y Davis (2009), las vocales de un bebé de 0;5 años que adquiría inglés se compararon con las de otro bebé de la misma edad que adquiría la lengua francesa. Ambos niños mostraron una relación estrecha con las frecuencias de su lengua ambiente.

La hipótesis de la interacción se deja sentir también en el ámbito de la prosodia. Entre los primeros trabajos que analizaron el efecto de la lengua ambiente en el *pitch* encontramos el de Weir (1966) según el cual un grupo de niños chinos a la edad de 0;5 y 0;6 años produjo patrones de *pitch* que fueron distinguibles de aquellos producidos por niños hablantes de ruso y de inglés. Otro ejemplo es el estudio de Sweeney (1973) en el que después de que un bebé de seis semanas fuera expuesto a las llamadas de pájaros durante un tiempo mayor del que fue expuesto a los sonidos del habla humana, se encontró que este producía chillidos de *pitch* alto similares a las llamadas de los pájaros. Casi una década más tarde el estudio de Kuhl y Meltzoff (1982) mostró que bebés menores de un año imitaron el patrón de entonación de las vocalizaciones de las producciones del adulto que ellos escucharon.

La presencia de la lengua ambiente en el *pitch* de las producciones del niño pequeño fue observada también en el trabajo de Whalen et al. (1991) en el que se compararon los patrones de entonación de vocalizaciones de dos y tres sílabas reduplicadas producidas por cinco niños expuestos al inglés americano y cinco niños que adquirirían la lengua francesa de entre 0;6 y 1;0 año. Los resultados mostraron un contorno creciente en las producciones de los niños de habla francesa que reflejaba el patrón de la lengua específica a la que pertenecían. No se encontró por el contrario dicho patrón en el grupo de niños de habla inglesa.

En el mismo año Levitt y Wang (1991) analizaron las vocalizaciones de cinco niños franceses y americanos entre los 0;4 y los 1;1 años de edad y encontraron que los de habla francesa emitían la segunda sílaba de las vocalizaciones reduplicadas con una mayor duración que la que presentó la segunda sílaba de los niños ingleses. Los resultados apuntan una vez más al influjo de la lengua materna en las emisiones tempranas de los bebés y de los niños pequeños.

También en niños de mayor edad se ha detectado la presencia de la lengua específica en los correlatos acústicos de la prosodia. Ejemplo de ello es el trabajo de DePaolis et al. (2008) en el que se examinaron las vocalizaciones de dos sílabas de niños que adquirirían americano, finlandés, francés y galés al comienzo del uso de la palabra (entre los 0;10 y los 1;6 años). Los análisis manifestaron diferencias en cuanto a la duración, intensidad y frecuencia fundamental de las vocalizaciones disilábicas reflejando los distintos ambientes lingüísticos de los que provenían.

En el trabajo de Boysson-Bardies et al. (1984) “Discernible differences in the babbling of infants according to target language” queda cuestionada la hipótesis universal defendida por Nakazima (1962), Preston et al. (1967) y Oller et al. (1976) ya que según los primeros autores las preferencias fonéticas universales descritas por Jakobson (1941) se dan en el balbuceo de niños pertenecientes a diferentes lenguas debido a los pocos estudios realizados, a que estos se

limitan a contar los segmentos dentro del balbuceo, y a las deficiencias de usar el *International Phonetic Alphabet* (IPA) para codificar las producciones infantiles.

Para arrojar algo de luz sobre esta cuestión y comprobar si verdaderamente se cumple la hipótesis universal de manera que la lengua ambiente no influye en las emisiones del bebé, Boysson-Bardies et al. (1984) llevaron a cabo dos experimentos con niños hablantes de árabe, chino y francés que tenían 0;8 años y 0;10 años en el primer experimento, y que en el segundo iban de los 0;6 a los 0;8 años por un lado, y tenían 0;10 años por otro.

En el primer experimento, los autores seleccionaron las producciones de quince segundos de duración del balbuceo marginal y canónico descrito por Oller (1980) en niños de 0;8 y 0;10 años que eran hablantes de árabe, chino y francés. Estas emisiones se agruparon en dos bloques que respondían a los dos grupos de edad y cada bloque recogía ocho pares de muestras en las que una de las producciones del par siempre pertenecía al balbuceo de los niños franceses y la otra al balbuceo de los niños que adquirirían el árabe o el chino. A continuación, las producciones se presentaron a adultos hablantes nativos de francés que tenían que juzgar cuáles de las muestras escuchadas pertenecían a su propia comunidad lingüística. Los resultados demostraron la presencia de las características de la lengua materna en las producciones de los niños a la edad de 0;8 años ya que los adultos franceses fueron capaces de reconocer el balbuceo de los niños pertenecientes a su comunidad a esa edad y descartar o etiquetar como extranjeras las emisiones de los sujetos de las otras dos comunidades lingüísticas.

En el segundo experimento dos fueron los objetivos. Por un lado, se intentó averiguar si los adultos podían identificar el balbuceo de los niños que adquirirían francés pero en este caso a las edades de entre 0;6 y 0;8 años, y a los 0;10 años; por otro, se pretendió hallar qué peso tenía el índice “no segmental” vs el “segmental” en la identificación del balbuceo en muestras extraídas únicamente de las producciones altamente articuladas del balbuceo reduplicativo a los 0;10 años de edad (a esta edad, las producciones eran pobres en patrones entonativos y ricas en rasgos segmentales frente a las muestras de los bebés de entre 0;6 y 0;8 años de edad que por el contrario envolvían patrones de *pitch* y de intensidad que favorecían los rasgos no segmentales del habla).

Como en el experimento anterior, las producciones se agruparon en pares de estímulos que entre los 0;6 y los 0;8 años, y a los 0;10 años, y a diferencia del experimento anterior, siempre incluyeron una muestra del balbuceo de la lengua francesa y otra del balbuceo de la lengua árabe debido a las dificultades que tuvieron los autores para recoger datos de la lengua china. Los resultados mostraron que los adultos identificaron mejor el balbuceo de los niños de su misma comunidad lingüística a las edades de entre 0;6 y 0;8 años, y a los 0;8 años por lo que



los juicios de estos adultos dependieron en su mayoría de los rasgos prosódicos o no segmentales de los estímulos, rasgos que eran más prominentes en las largas producciones entonadas con un porcentaje bajo de articulación (los rasgos no segmentales más usados son el tipo de fonación y la organización de los contornos del *pitch* y de la intensidad).

Estos resultados apoyan por tanto la idea de que las dimensiones perceptivas relevantes en relación con la lengua meta pertenecen al dominio de la <sup>1</sup>metafonología (Oller [1980]) no al de la fonología segmental. Además, Boysson-Bardies et al. (1984:11) creen que:

“[...] In our view, a child’s ability to produce local patterns similar to patterns found in the target language reflects (a) an early general attuning of the vocal tract to conform with the laryngeal and supralaryngeal settings that are specific to the target language, and (b) the fact that certain rhythmic and intonational properties of the target language are beginning to be acquired”.

Tal vez Vihman (1996) defendía la misma idea cuando afirmaba que los rasgos prosódicos son el primer aspecto de las vocalizaciones del bebé en desplegar la influencia de la lengua materna insinuando que los rasgos no segmentales muestran la influencia de la lengua ambiente antes de que los rasgos segmentales lo hagan. No obstante, no todos los trabajos en los que hablantes adultos debían reconocer el habla de su comunidad lingüística a partir de las vocalizaciones de bebés han demostrado que estos pudieran hacerlo. Un ejemplo es el estudio de Olney y Scholnick (1975) en el que estudiantes universitarios debían identificar los pares de vocalizaciones extraídas de bebés hablantes de chino y de inglés americano a los 0;6, 1;0 y 1;6 años. Los resultados mostraron que los estudiantes no fueron capaces de diferenciar la comunidad lingüística a la que pertenecía el bebé aunque sí que pudieron discriminar las edades de los sujetos. Lo mismo sucedió en el estudio de Thevenin et al. (1985) en el que adultos hablantes de español e inglés no pudieron discriminar el balbuceo de bebés de entre 0;7 y 1;2 años que adquirirían la lengua inglesa y la española. No obstante, en este último los estímulos que recibieron los adultos consistían en muestras de balbuceo canónico de tan solo de uno a tres segundos de duración frente al estudio de Boysson-Bardies (1984) en el que los estímulos tenían una duración de quince segundos.

Para Kuhl y Meltzoff (1996:2425) la lengua ambiente está detrás de uno de los factores que explican el cambio en las vocalizaciones de los niños de corta edad y que no es otro que el

---

<sup>1</sup> Se trata de un término acuñado por Oller (1980) para referirse al estudio de las vocalizaciones de los bebés a partir de distintos análisis acústicos.

del *vocal learning* o aprendizaje vocal ya que “[...] Human infants listen to ambient language spontaneously and attempt to produce sound patterns that match what they hear”. Según estas autoras, los niños adquieren el inventario concreto de unidades fonéticas, palabras y rasgos prosódicos empleados en una lengua particular en parte por la imitación, esto es, imitando los rasgos de la lengua ambiente que ellos perciben (capacidad compartida por el homo sapiens y por algunas aves cantoras). De esta forma para las investigadoras, la adquisición del habla está condicionada anatómicamente y guiada auditivamente por la influencia de la lengua ambiente. Pero para que se produzca la imitación vocal, los bebés han de identificar la relación entre el movimiento articulatorio y el sonido producido de modo que la pregunta que surge es ¿cuándo adquieren los bebés el enlace articulatorio-auditivo? Si revisamos los trabajos ya expuestos en el estado de la cuestión parece que antes del año de vida los bebés son capaces de imitar segmentos vocálicos y consonánticos, los valores de los formantes y hasta los rasgos prosódicos de la lengua meta.

Para intentar dar muestras de esta imitación vocal, Kuhl y Meltzoff (1996) llevaron a cabo un experimento en el que se analizaron las vocales que producían bebés de 0;3, 0;4 y 0;5 años al escuchar las palabras *hop*, *heap* y *hoop* emitidas por un hablante adulto femenino y presentadas a través de una máquina. Los sonidos que el bebé debía producir por imitación eran [a, i, u] y en ellos se analizó el primer y el segundo formante, el rasgo compacto-difuso, el rasgo grave-agudo, la frecuencia fundamental y la duración total de la emisión. Por otro lado, fonetistas entrenados debían escuchar las vocalizaciones y codificarlas en una de las siguientes vocales /i, ɪ, ε, æ, a, ʌ, ʊ, u/. Los resultados indicaron que los bebés producían más aes en respuesta al estímulo de la palabra que contenía [a], más íes en respuesta a la palabra que contenía la [i] y más úes tras escuchar el estímulo que contenía [u] demostrándose por tanto que el estímulo vocálico afecta fuertemente al tipo de vocal que los niños producen como respuesta. Se cumple así la hipótesis de la imitación vocal ya que si los bebés han demostrado que tras quince minutos de exposición a la grabación de unas emisiones pueden imitarlas y producir los mismos sonidos, mucho más influirá la lengua ambiente que les rodea, y a la que están expuestos desde el nacimiento, a la hora de guiar sus producciones.

Por otro lado, los gráficos manifestaron que los espacios vocálicos ocupados por las vocales de los bebés [a, i, u] se fueron separando progresivamente de la semana decimosegunda a la semana vigésima agrupándose las vocales en cada una de sus categorías con el paso del tiempo (la agrupación fue más fuerte en la semana decimosexta que en la decimosegunda y en la semana vigésima que en la decimosexta). Kuhl y Meltzoff (1996:2436) creen que el origen de este cambio no debía limitarse a un razonamiento anatómico según el cual un rápido crecimiento del tracto vocal se producía a estas edades, sino que también debía considerarse el

aprendizaje vocal. Según las autoras, el bebé escucharía las vocales producidas por los adultos e intentaría emitir sonidos perceptivamente semejantes a los que aquellos producen. La audición de estos sonidos llevaría al bebé a crear representaciones mentales de las vocales de su lengua ambiente que almacenadas en la memoria servirían como meta u objetivo que aquel debería alcanzar. Para ello, el bebé tendría que enlazar sus movimientos articulatorios con el sonido que escucha y ello comienza a los 0;4 años con la etapa del *cooing* resultando en:

“[...] This would result in the development of an auditory-articulatory ‘map’ relating self-produced auditory events to the motor movements that caused them. These experiences would in turn lead to the development or refinement of any perceptivo-motor linkages that were present initially”.

Una pregunta que plantea el término *imitación* es si esta aparece guiada además de por la representación mental y el enlace acústico-articulatorio, por los movimientos faciales del que habla. Para intentar dar respuesta a esta cuestión Meltzoff y Moore (1977) llevaron a cabo un experimento en el que se examinó la respuesta de bebés de entre 2 y 21 días de edad después de que estos observaran durante quince segundos los movimientos faciales de un adulto. El bebé fue expuesto durante quince segundos a uno de los siguientes gestos faciales: protrusión de los labios, abertura de la boca y protrusión de la lengua que el adulto realizaba. Finalizado este tiempo, comenzaba el periodo de respuesta en el que se observaba si el gesto o los gestos del bebé coincidían con los realizados por el adulto en los veinte segundos siguientes. Los resultados mostraron que los bebés imitaban los gestos faciales del adulto a los que estaban expuestos de manera visual ya que estos producían más gestos de protrusión de la lengua inmediatamente después de que este gesto fuera realizado por el adulto que después del gesto de abertura de la boca. Similarmente hubo más respuestas de abertura de boca después de este gesto que después de otros. Según las autoras (1977:200):

“[...] The hypothesis we favor is that this imitation is based on the neonate's capacity to represent visually and proprioceptively perceived information in a form common to both modalities. The infant could thus compare the sensory information from his own unseen motor behavior to a "supramodal" representation of the visually perceived gesture and construct the match required”.

Otro experimento realizado por Meltzoff y Moore (1989) pero en este caso con bebés de tan solo 40 horas, también manifestó la capacidad de estos de imitar los movimientos faciales e incluso el movimiento de cabeza a esta temprana edad. Asimismo Vihman (1986) cree que el hecho de que los hablantes que poseen alguna discapacidad en la audición puedan aprender a leer los labios como refuerzo a la señal auditiva por un lado, y el hecho de que los bebés den sus

primeros pasos con la lengua practicando con gestos articulatorios en principio sin fonación y por tanto sin efecto auditivo por otro, sugieren que dichos bebés atienden en gran medida a los efectos visuales producidos por las caras de los que hablan. Para Kuhl y Meltzoff (1996) sin embargo, aunque puede que el movimiento de la cara favorezca la imitación de los niños no hay que olvidarse del componente auditivo tal y como se desprende de un experimento anterior realizado por estas autoras (1982) en el que los niños que fueron expuestos a tonos que no eran del habla, aunque poseían el carácter espectral de las vocales, no produjeron sonidos del habla. Esta cuestión nos lleva a la percepción multimodal del habla defendida por algunos autores y que queda fuera del alcance de esta tesis.

Concluyen Kuhl y Meltzoff (1996) que los bebés y niños pequeños se aproximarán cada vez más a la vocal de la lengua meta según consigan relacionar las consecuencias acústicas de sus propios movimientos articulatorios con las consecuencias acústicas de los actos articulatorios de otros.

Esta explicación nos lleva al concepto de feedback, tercer factor que interviene en el proceso de producción y percepción del habla para Polka et al. (2007:151) “[...] Infants’ perceptions of the sounds they themselves produce”. De este modo, en opinión de Kuhl y Meltzoff (1996) los niños no solo tienen que ser capaces de percibir los sonidos producidos por otros sino que necesitan escuchar los resultados de sus propias producciones para aprender a hablar. Es más, sin el *feedback* auditivo de sus propios movimientos articulatorios ellos no se acercarían al circuito perceptivo-motor cuyo control es necesario para la adquisición de una lengua. Así “[...] Both hearing the sound patterns of ambient language (auditory exteroception) and the ability to hear one’s own attempts at speech (auditory proprioception) are critical to determining the course of vocal development” (1996:2436).

La importancia del *feedback* para la adquisición de una lengua quedó patente cuando en el apartado anterior explicamos los modelos artificiales de producción de voz y en concreto cuando desarrollamos el modelo DIVA de Guenther (1995) ya que dicho modelo predecía que el acceso a ambas producciones del habla, las propias y las del adulto, es crítico para el desarrollo del control motor de la misma. Además de señalar que la lengua ambiente influye en las características acústicas del habla del niño desde que este comienza a balbucear, dicho modelo defendía la importancia que la retroalimentación auditiva del habla que posee el propio niño tiene para la adquisición del control motor de la misma.

Para Kent y Miolo (1995) también es importante la relación entre la práctica o producción de *speechlike vocalizations* y el *feedback* que envuelve la audición y el seguimiento de estas producciones practicadas. Y es que cuando se repiten los movimientos que producen un

sonido, se crea una fuerte relación entre las impresiones cinestésicas (o kinestésicas) y táctiles y las sensaciones auditivas que el niño recibe de sus propias emisiones. La conciencia de estos enlaces entre los movimientos orales y la señal acústica resultante que el bebé percibe, es un prerrequisito para la correspondencia auditiva-vocal que fundamentará la producción de las palabras según los investigadores.

Otros autores como Boysson-Bardies et al. (1986) utilizan un concepto similar y hablan del auto producido *feedback* auditivo que es según ellos suficiente para entrenar una red de conexiones entre el *target* auditivo y los movimientos articulatorios incluso cuando las diversas estructuras del habla están sufriendo cambios considerables con la reestructuración típica del desarrollo (el modelo es así capaz de mantener los objetivos vocálicos durante el desarrollo y bajo condiciones de reestructuración morfológica porque se apoya en el *feedback* auditivo). Además, según estos autores la red de conexiones entre el espacio auditivo y el articulatorio puede aprenderse.

Para acabar este apartado, nos referimos a un artículo que si bien no se centra en el *feedback* sí que aporta una pequeña curiosidad acerca de este. Hablamos del trabajo de Fagan e Iverson (2007) en el que se investigó la influencia del *mouthing* o contacto de un objeto con la boca, labios o lengua (según Rochat [1989] y Ruff et al. [1992]) en las vocalizaciones. Según los autores, las vocalizaciones que ocurren durante el *mouthing* (combinación de movimientos de la lengua y los labios y los cambios relacionados con la forma de la boca) pueden contribuir al desarrollo de las relaciones entre los gestos orales y las consecuencias auditivas mediante la asociación de determinadas posturas con el sonido vocálico o consonántico producido.

### **2.2.3. Habla dirigida al bebé.**

Finalmente y en cuanto al input lingüístico que recibe el bebé, resulta interesante conocer que el habla que los adultos emplean cuando se dirigen a estos (*Infant Direct Speech*) es diferente a la que usan cuando se comunican con otro adulto (*Adult Direct Speech*). De acuerdo con Lindblom (1990), cuando los adultos se dirigen a los bebés lo hacen empleando un habla más clara, con segmentos fonéticos de mayor duración y más audibles que permiten reducir en parte la confusión de la coarticulación. Su forma de hablar es por tanto hiperarticulada.

Esta hiperarticulación del habla que es defendida por otros autores como Fernald et al. (1989), Petitto (1993) o Kuhl et al. (1997) se caracteriza por presentar una  $f_0$  más alta. Así lo demostraron los experimentos de Stern et al. (1983), Sundberg (1998) y Kitamura y Burnham (2003) quienes tras grabar el habla del adulto dirigida a niños recién nacidos y durante los

primeros años de su vida, encontraron una frecuencia fundamental generalmente más alta en el habla dirigida al bebé que en el habla dirigida al adulto. Otro rasgo que diferencia el IDS del ADS es la duración de la vocal que se presenta más larga en el habla dirigida al bebé que en el habla dirigida al adulto (Bernstein Ratner y Luberoff [1984] y Sundberg [1998]). Según Englund y Behne (2006) el aumento de la duración del sonido facilita su identificación de forma que una vocal más larga al ser acústicamente más saliente será más fácil de reconocer. Para Cooper y Cooper (2000) esto podría ser el resultado de una velocidad de habla más lenta empleada por los adultos cuando se dirigen a los bebés, velocidad esta última preferida por los niños. También Panneton et al. (2006) demostraron que los infantes de 18 semanas preferían el habla lenta a una rápida (además según estos el habla más lenta parecía aumentar la percepción del bebé del afecto vocal).

También las frecuencias de los formantes evidencian una clara diferencia entre el IDS y el ADS ya que el habla dirigida al niño presenta un espacio vocálico más expandido (Sundberg [1998]). En el estudio de Kuhl et al. (1997) se examinó el input lingüístico que recibieron unos niños hablantes de inglés americano, de ruso y de sueco. Para ello se grabó a diez hablantes nativos femeninos de cada una de las lenguas en dos condiciones: en una de ellas las mujeres hablaban con sus hijos de 0;2 a 0;5 años con palabras que contenían sobre todo /i/, /a/ y /u/; en la segunda condición, dichas mujeres hablaron con adultos. Para cada vocal se hicieron varias medidas acústicas y entre ellas se midieron la  $f_0$  y el F1, F2 y F3 al comienzo, centro y final de la vocal. Los resultados confirmaron que en las tres lenguas, las madres produjeron vocales acústicamente más extremas cuando se dirigían a los niños resultando en una expansión del triángulo vocálico en el habla dirigida a los infantes. Para Kuhl et al. (1997) este triángulo vocálico desplegado resulta beneficioso para los niños porque un triángulo vocálico expandido aumenta la distancia entre las vocales haciéndolas más distintivas unas de las otras. Además, se ha demostrado que las vocales hiperarticuladas acústica y articulatoriamente son reconocidas por los adultos como los mejores ejemplos de las categorías vocálicas (Johnson et al. [1993] e Iverson y Kuhl [1995]) y que los niños tienen mayor habilidad de categorización fonética cuando escuchan buenos ejemplos de las categorías fonéticas.

Cuatro años más tarde el trabajo de Weijer (2001) llevó a cabo un experimento con el objetivo de intentar replicar los resultados de Kuhl et al. (1997). En él se midió el F1 y el F2 de las vocales /i/, /a/ y /u/ en el habla dirigida a un bebé de entre 0;6 y 0;9 años de edad y en el habla dirigida a los adultos. Las vocales fueron extraídas de palabras con contenido (sustantivos, verbos, adverbios, etc.) como en el trabajo de Kuhl et al. (1997) pero a diferencia de este se utilizaron también palabras vacías de contenido. Los resultados mostraron que los triángulos

vocálicos en el habla dirigida al niño se expandieron en comparación con los del habla dirigida a los adultos, pero solo en el caso de las palabras contenido.

Los trabajos de Sundberg (1998) y de Liu et al. (2003) también demostraron la existencia de un mayor espacio vocálico en el habla de la madre cuando esta se dirigía a su hijo que cuando hablaba con otro adulto. En este último se estableció una correlación positiva entre el espacio vocálico más expandido en el IDS de las madres y la facilidad de los niños de discriminar sonidos del habla después de realizar un experimento de giro de cabeza. Sin embargo, el artículo de Englund y Behne (2006) nos remite a un experimento de estos mismos autores (Englund y Behne [2005]) en cuyos resultados no estaba claro que el espacio vocálico del IDS fuera mayor que el del ADS.

Antes de finalizar este apartado y debido a la multitud de autores y trabajos citados a lo largo del mismo, creemos conveniente sintetizar brevemente los principales asuntos tratados para facilitar su comprensión al lector. En primer lugar esta sección ha dado cuenta del desarrollo anatómico de los órganos responsables de la percepción así como de las habilidades perceptivas con las que cuenta el feto antes de que el periodo de gestación llegue a su fin (como la capacidad para diferenciar lenguas o la preferencia por la voz de la madre) y al término de este periodo, es decir, nada más nacer. Diferentes trabajos han manifestado las similitudes perceptivas entre el bebé y el adulto en cuanto a la diferenciación del infante de las vocales y las consonantes, en cuanto al diferente modo de percepción dependiendo de la naturaleza del sonido percibido (vocal y consonante), en relación a la percepción de la vocal en sí (relación formántica) y en cuanto al uso de estrategias perceptivas por parte de ambos como es la de la normalización. Posteriormente, se han presentado las dos principales hipótesis acerca de la influencia de la lengua ambiente en las destrezas perceptivas y en las emisiones del bebé y niño pequeño a partir de los sonidos emitidos por estos (hipótesis de la independencia y de la interacción). Con ayuda de experimentos se ha evidenciado la capacidad de los bebés y de los niños pequeños de discriminar entre sonidos no pertenecientes a su lengua y nunca antes escuchados. Ha habido estudios que han demostrado la influencia de la consideración de la vocal (prototípica y no prototípica) a la hora de una mejor discriminación de la misma. Otras pruebas han manifestado que con el paso del tiempo dichas capacidades perceptivas quedan mermadas y la habilidad del bebé para diferenciar sonidos de una lengua no nativa se ve reducida debido a la influencia que ejerce en este la lengua materna. Finalmente, incluimos algunos trabajos en los que se examinaron las características del habla dirigida al bebé (IDS) y se observó que esta era hiperarticulada, esto es, que sus segmentos eran de mayor duración y poseían una  $f_0$  más alta.

## **2.3. Revisión bibliográfica sobre el análisis vocálico.**

### **2.3.1. Consideraciones iniciales.**

Acercándonos un poco más al objeto de nuestro estudio, nos disponemos a revisar en este apartado por un lado aquellos trabajos en los que la observación del habla del bebé y niño pequeño ha permitido establecer el inventario de sonidos vocálicos que caracteriza las emisiones de estos durante los primeros años de vida, y por otro aquellos estudios en los que los sonidos vocálicos se han analizado además desde el punto de vista acústico mediante el examen de sus valores formánticos. Antes de comenzar dicha revisión debemos tener en cuenta dos cuestiones de muy diversa índole acerca de estas investigaciones:

La primera cuestión atañe a la escasez de trabajos puesto que son pocos los estudios que han emprendido la tarea de registrar de manera longitudinal los sonidos vocálicos producidos por bebés y niños pequeños, y son escasos también los trabajos que han analizado dichas emisiones acústicamente. Una de las mayores dificultades radica, tal y como comentan Palethorpe et al. (1996), en la propia obtención de los datos y en concreto en la búsqueda de sujetos ya que debido a la corta edad de estos el investigador debe tratar con los padres y lograr que estos acepten las condiciones del estudio que a veces pueden resultar tediosas puesto que no solo sus hijos serán grabados de forma regular sino que ellos, como padres (o al menos uno de ellos), deberán estar presentes en cada una de las sesiones con lo que se verán forzados a participar en el estudio. Esto conlleva tal y como observan Bond et al. (1982) que los estudios longitudinales sean mucho más escasos que los transversales al ocuparse estos últimos de un periodo de tiempo más corto.

Otros inconvenientes intrínsecos a la edad del sujeto, y que contribuyen considerablemente a la escasez de publicaciones, son por ejemplo la dificultad que presenta el transcribir muestras de habla de sujetos tan jóvenes tanto por lo complicada que resulta la identificación de los sonidos de los bebés, cuyas características acústicas no suelen coincidir con las de los sonidos adultos, como por las carencias que presentan los actuales sistemas de transcripción que como el Alfabeto Fonético Internacional están diseñados para la transcripción de las emisiones adultas, incluso con la utilización de todos sus diacríticos (Kent y Bauer [1985], Oller [1986], Gilbert et al. [1997] y Robb et al. [1997]). En relación a las deficiencias que según Oller (1986) tiene el método de transcripción a partir de categorías adultas, Vihman (1996:100) comenta: “[...] Nevertheless, its disadvantages, particularly for the younger infant, are clear: Such a transcriptional system interprets infant productions through the limiting filter of adult segmental categories”. Ello puede resultar según Oller (1986) en una combinación o



ecuación ilusoria de dos tipos de elementos, los sonidos de los niños y los elementos fonéticos adultos.

Estas limitaciones pueden ser subsanadas según el propio Oller (1986) acompañando el estudio de las vocalizaciones de un análisis acústico que informe de la duración, la frecuencia fundamental, la amplitud y las frecuencias de resonancia (formantes) de dichos sonidos (medidas todas objetivas y cuantificables). No obstante, estas indicaciones acústicas carecerían de sentido si no se interpretaran bajo el marco apropiado. Por ello Oller (1986) diseña la denominada “metafonología” o “infrafonología” basada en el análisis de la frecuencia, la resonancia, la intensidad y el tempo para establecer cuáles de las vocalizaciones que emite el bebé durante los primeros meses de vida son unidades fonéticas bien formadas. Con este modelo, la caracterización de las vocalizaciones que produce el bebé en el proceso de adquisición de la lengua materna ya no dependerá únicamente de la mera transcripción fonética sino de estos índices acústicos que permitirán determinar qué vocalizaciones son inmaduras (como las producidas entre los 0 y los 0;2 años, también llamadas “cuasi-vocales”) y cuáles presentan características fonéticas adultas (como las que emite el bebé entre los 0;3 y los 0;8 años de vida que se denominan vocalizaciones de “plena resonancia” y entre las que se encontrarían los sonidos vocálicos de nuestro estudio).

Las limitaciones de la transcripción fonética de los segmentos producidos por el bebé fueron también consideradas por Stark et al. (1975) y por Stark et al. (1978) quienes siguiendo los pasos de Oller (1986) desarrollaron un sistema basado en comportamientos no vocales que acompañaban a las vocalizaciones entre los que se encontraban la dirección de la mirada, la expresión facial, los movimientos corporales o postura del bebé y la respuesta de la madre. Las vocalizaciones producidas en estos contextos se analizaron acústicamente y de manera auditiva y se obtuvo como conclusión que muchas de las características del habla adulta aparecen ya en estas tempranas emisiones (como la sonoridad, la oclusión glotal o la modulación del *pitch*). Este sistema fue muy útil para caracterizar las primeras vocalizaciones de los bebés (hasta la edad aproximada de 0;4 años) ya que permitió a los autores diferenciar entre sonidos vegetativos, de llanto, de disconformidad o de confort.

Koopmans-van Beinum y Van der Stelt (1986) también evitaron usar las categorías adultas para clasificar las vocalizaciones de los bebés y crearon un método basado en la unidad del ciclo respiratorio como unidad básica. Dentro de esta unidad, los autores se fijaron en la fonación, que podía ser continua o interrumpida, y en la articulación, observando si se producían o no movimientos articulatorios supraglóticos.

Tras revisar estos modelos, Alexandra Karousou (2003) decidió estudiar en su tesis doctoral las vocalizaciones del bebé que aparecían antes de la primera palabra, combinando factores articulatorios (Oller [1986]), de entonación, duración y ritmo para observar cómo la evolución y la interacción de todos estos parámetros (nivel segmental y suprasegmental) preparaban el camino para la palabra adulta.

A partir de los métodos explicados (articulatorio, de ciclo respiratorio y de comportamientos no vocales) se llega a una agrupación de las vocalizaciones de los bebés en distintas etapas según las características que presenten dichas vocalizaciones, y aunque estas etapas difieren en la nomenclatura y a veces en el periodo de tiempo que abarcan, todas ellas coinciden en el contenido. Un ejemplo lo encontramos en la primera etapa en la que se agruparon las vocalizaciones de los bebés que fue denominada por Oller (1986) como “etapa de fonación” y que tenía lugar según el autor en el primer mes de vida, mientras que para Stark et al. (1975, 1978), Stark (1980) y Koopmans-van Beinum y van de Stelt (1986) la primera etapa abarcaba los dos primeros meses que seguían al nacimiento y se denominaban “etapa reflexiva” y “etapa de fonación ininterrumpida” respectivamente. Sin embargo, en todas ellas dominan los sonidos vegetativos, de desagrado y de llanto que aparecen de manera involuntaria y que son producidos en general por la realización de otras funciones del cuerpo como la respiración, la ingestión o digestión y el tragado (véase Vihman [1996] para una completa caracterización de cada una de las etapas según los tres modelos).

En la lengua española, han sido Miquel Serra et al. (2000) quienes han agrupado las vocalizaciones de los bebés en diferentes etapas a partir del criterio articulatorio. Según estos autores, existen cinco etapas dentro del periodo prelingüístico:

La primera etapa es la *etapa de fonación o de vocalizaciones reflejas* que tiene lugar entre los 0 y los 0;2 años y que reúne por un lado los sonidos que producen los bebés en relación a los cambios debidos a la nutrición, la respiración, la succión y el tragado, tratándose por tanto de sonidos vegetativos que tienen que ver según Vihman (1996) con el control inmaduro de las funciones esenciales del cuerpo, y por otro agrupa aquellos sonidos que el bebé emite en situaciones de confort o placer y de dolor.

La segunda etapa es la de *gagueo o sonrisas* que va de los 0;2 a los 0;4 años y que recibe este nombre por el aumento de las producciones velares (gagueo) producidas normalmente por las sonrisas y el habla que el interlocutor dirige al bebé.

La tercera etapa es la etapa de *expansión fonética o juego vocal* (entre los 0;4 y los 0;6 años) y en ella parece que el bebé ha aumentado el control del mecanismo laríngeo y

articulatorio con lo que es probable que este produzca sonidos voluntariamente. En este periodo comienzan a manipularse rasgos prosódicos como el cambio de tono o la intensidad y aparecen por primera vez sonidos vocálicos de resonancia completa (vocales muy similares a las del adulto). El bebé es también capaz de producir sonidos consonánticos acompañados de ruidos, gritos, gruñidos y de otras manifestaciones expresivas.

La cuarta etapa es la del *balbuceo canónico* (a los 0;6 años) y en esta encontramos por primera vez el tipo de unidad silábica, constituido por un cierre (consonante) y una abertura (vocal), parecida a la del adulto. Dentro del balbuceo canónico encontramos el *balbuceo reduplicativo*, que se produce cuando los bebés repiten gestos articulatorios de manera constante a distinta velocidad y con diferente entonación, como en *bababa*, y observamos también el *balbuceo melódico*, que se da cuando en las emisiones del balbuceo el bebé modifica la vocal y la consonante e introduce diferentes contornos melódicos como en *aaagadá*.

La última etapa es la que se denomina *balbuceo variado o conversacional* (a los 0;10 años) y se caracteriza por un aumento notorio de las consonantes y de los tipos de estas, por una mayor aparición de segmentos propios de la lengua que rodea al bebé, y por un aumento de los registros tonales y melódicos. Se le llama *conversacional* porque parece que el niño puede seguir turnos alternativos de habla.

A pesar de estos intentos por caracterizar las producciones tempranas de los bebés y niños pequeños teniendo en cuenta el tipo de fonación, la articulación de los segmentos, la unidad de ciclo respiratorio, distintos elementos prosódicos o incluso comportamientos no verbales, la mayoría de los estudios optan por una transcripción fonética de los sonidos vocálicos y consonánticos de las emisiones infantiles.

Estudios más actuales como el de McAllister (2009) señalan las desventajas que conlleva la mera transcripción fonética de los segmentos y creen que hay muchas razones para pensar que una caracterización del habla del niño basada en este sistema podría contener distorsiones y omisiones que la distanciarían de manera notable de la verdadera realidad fonética. Ello lleva a los autores a valerse del análisis acústico para completar los resultados de la transcripción.

No obstante, si bien el análisis acústico sirve para paliar en cierta medida las insuficiencias de la transcripción segmental, la dificultad que presenta su interpretación sobre todo a la hora de identificar los formantes, es sin duda uno de los factores que más influye en la escasez de publicaciones. La complejidad del examen acústico viene dada según Kent (1976)

por las peculiaridades que presenta el habla de los bebés y de los niños tan jóvenes. Entre estas particularidades se encuentran según este autor las siguientes:

La tendencia de los bebés y de los niños pequeños a emitir sonidos con voz no modal del tipo *breathing* (empañada) o *creaky* (crepitante), voz que contribuye a oscurecer otros índices acústicos como la identificación de los formantes (Kent [1976]).

La nasalización de los sonidos producidos por el bebé, nasalización que tiene lugar sobre todo en los cuatro primeros meses de vida debido a los factores ya explicados en el punto 2.1.1 (como el de la alta posición de la laringe, la aproximación del velo del paladar y la epiglotis, y la incapacidad de los músculos del velo del paladar de elevarlo debido a su ubicación más alta), que se refleja en una serie de resonancias y antirresonancias alrededor de los formantes las cuales dificultan en gran medida su interpretación tal y como señalan Kent (1976) y Kent y Murray (1982) (para una mayor comprensión de los correlatos acústicos de los sonidos nasales véase el capítulo cuarto sobre la caracterización articulatoria y acústica de las vocales del español).

La alta frecuencia fundamental del habla de los bebés que provoca que los armónicos estén más separados y puedan confundirse con los formantes (Fant [1960], Lindblom [1961], Kent [1976], Bond et al. [1982], Bickley [1989], De Clement y Wijnen [1994], Robb y Cacace [1995], Kuhl y Meltzoff [1996], Palethorpe et al. [1996], Gilbert et al. [1997] y Pols et al. [2006]). Según Ménard et al. (2004), en una  $f_0$  más alta de 3,5 barks (unos 339 Hz) los dos primeros armónicos estarían espaciados por una distancia mayor de la distancia crítica de integración espectral (Chistovich et al. [1979]) con lo que la primera prominencia espectral no correspondería con el primer formante sino con el primer armónico. Esta dificultad añadida ha llevado a Gilbert et al. (1997) a justificar la ausencia de publicaciones que analizan los dos primeros formantes frente a otras que estudian la  $f_0$ , por ejemplo, y ha contribuido según los autores a la falta de estudios que se centren en las vocalizaciones de los niños antes de los tres años siendo más frecuentes los trabajos a partir de esta edad (como los de Eguchi y Hirsh [1969], Bennett [1981], Hillebrand et al. [1995], Busby y Plant [1995], Lee et al. [1999] y Whiteside [2001]). No obstante, las modernas técnicas han permitido disminuir este problema mediante la utilización de espectrogramas de anchos de banda más amplios y análisis LPC (Kent [1976], Palethorpe et al. [1996] y Gilbert et al. [1997]).

Todos estos factores contribuyen a que la estructura formántica quede pobremente definida y resulte complicado localizar los formantes (Pols et al. [2006]).

Otro obstáculo que nos gustaría comentar y que fue detectado a la hora de realizar el análisis de nuestros datos, fue el de las fuertes y constantes subidas y bajadas de la frecuencia fundamental, típicas en los bebés que están experimentando y jugando con el habla, que distorsionaban los armónicos provocando que el estímulo sonara a sintético o que se saturara la señal.

El hecho de que no se pueda establecer una comparación entre los formantes de los adultos y los de los bebés debido a que las diferencias morfológicas provocan que las frecuencias de estos últimos sean mucho más altas (tal y como veremos en el apartado de los resultados, discusión y conclusión), es otro inconveniente.

Todas estas dificultades entorpecen el análisis acústico de las emisiones tempranas de los bebés y de los niños pequeños y además son las responsables de que no dispongamos de muchas referencias con que comparar nuestros datos para encontrar posibles respuestas a nuestros resultados.

La exigüidad de trabajos longitudinales que estudian las emisiones del bebé y del niño pequeño queda entonces justificada por el considerable periodo temporal que abarca este tipo de estudios, por las deficiencias del sistema de transcripción fonética y por las dificultades que presentan otros análisis como el acústico.

La segunda cuestión que queremos aclarar antes de proceder a la revisión bibliográfica de todos aquellos trabajos longitudinales que han indagado en el habla del bebé, es la de la clasificación de dichos trabajos según el tipo de estudio realizado. De esta forma, encontramos investigaciones que se encargan de registrar y transcribir los sonidos del habla del niño para un estudio lingüístico posterior, y aquellos otros en los que la frecuencia vocálica y la transcripción fonética de los segmentos presentes en el habla del niño a lo largo del tiempo vienen acompañadas, como es el caso de esta tesis, de un análisis acústico. Creemos necesaria la inclusión bibliográfica de estos dos tipos de estudios en este trabajo al ser ambos el objeto de estudio de la presente investigación.

En cuanto a estas investigaciones, llama la atención tal y como podrá comprobar el lector en las páginas siguientes, la ausencia de estudios que hayan analizado la lengua española (ausencia tanto de estudios de tipo longitudinal que hayan realizado el inventario de los sonidos vocálicos de bebés y niños pequeños que adquieren la lengua española durante un considerable periodo de tiempo, como de estudios que hayan analizado acústicamente los dos primeros formantes).

En relación a los primeros, si bien existen trabajos que han analizado el habla del bebé y niño pequeño en diferentes edades como los de Laura Bosch (2004), Conxita LLeó (1997), Susana López Ornat (2005, 2013) y Alexandra Karousou (2003, 2013), estos no ofrecen un registro detallado del tipo vocálico ni de la frecuencia vocálica de los sonidos vocálicos que emite un niño que adquiere la lengua española durante el periodo de tiempo comprendido entre los 0;4 y los 3;0 años de edad. Por otro lado, y aunque existen algunos corpus que recogen el habla de bebés y niños pequeños hablantes de catalán, español y gallego (como el CHIEDE o Corpus de Habla Infantil Espontánea del Español que recoge el habla de niños que adquieren el español como lengua materna y que fueron grabados a partir de los 3;0 años, el corpus Koiné que incluye el habla de niños que adquieren el español y el gallego entre los 1;10 y los 4;0 años, y el proyecto Phon.cat y Phon.es perteneciente al corpus CHILDES que registra el habla de niños pequeños monolingües y bilingües que adquieren catalán y español entre los 0;11 y los 4;8 años), lo que encontramos en estos es una transcripción fonética de las emisiones de los niños sin que se hayan hecho estudios en los que se analice cuáles son las primeras vocales en aparecer, cuáles aparecen después, a qué edades aparecen unas y otras, cuáles son las vocales más frecuentes, y si estas preferencias se mantienen a lo largo de un periodo de tiempo concreto.

En cuanto al segundo tipo de estudios, aquellos que analizan acústicamente los dos primeros formantes de las vocales de niños que están adquiriendo la lengua española entre los 0;4 y los 3;0 años de vida, y a excepción del trabajo de Casal et al. (2002) donde se miden los dos primeros formantes de las vocales de niños con desarrollo fisiológico normal y se compraran con los de niños con problemas fisiológicos cuya media de edad es de 2;3 años (y al que nos referiremos en el capítulo octavo de la discusión cuando utilicemos los valores de los formantes obtenidos por estos autores para compararlos con los nuestros), la ausencia es total.

Por otro lado, tanto los trabajos que han registrado el inventario vocálico de los bebés y niños pequeños que adquieren la lengua española, como los estudios que han analizado acústicamente los dos primeros formantes, siguen en parte una misma metodología ya que ambos se han realizado bajo circunstancias similares de grabación. De esta forma, las grabaciones de estos estudios recogían el habla espontánea producida por el bebé en interacciones naturales con su madre o algún otro miembro de la familia, la mayoría de ellas tenían lugar en la casa familiar o en el laboratorio, y en todas se recreaban situaciones habituales de juego, alimentación, y otras tareas cotidianas (por ello no especificaremos ninguna de estas circunstancias en cada uno de los trabajos que revisaremos a menos que estas cambien y sea por tanto pertinente hacerlo).

Finalmente, y tal y como comprobará el lector si acude a los trabajos originales que revisaremos a lo largo del siguiente apartado, el formato de las vocales que aparecen en dichos estudios depende de la consideración que tengan de estas los autores que las han estudiado: como auténticos fonemas, caso en el que aparecerán entre barras oblicuas, o como segmentos vocálicos, caso en el que se mostrarán entre corchetes. Al no haber acuerdo general en cuanto a su consideración, y al ser estas analizadas en todos estos trabajos desde una perspectiva segmental, se ha optado en esta tesis doctoral por encerrarlas entre corchetes.

### **2.3.2. Revisión de estudios de inventario vocálico.**

Entre los estudios longitudinales que intentan establecer el inventario fonético del bebé, debemos considerar en primer lugar aquellos a los que nos referimos cuando hablamos de la influencia de la lengua ambiente en el habla del niño por presentar estos una descripción del repertorio vocálico de los bebés y de los niños pequeños (Boysson-Bardies et al. [1980], Oller y Eilers [1992], Davis y MacNeilage [1995], Davis y Davis [2000], Rvachwe et al. [2008], Kern y Davis [2009], Le et al. [2009], Chen y Kent [2010] y Kern et al., en prensa). En todos estos estudios se manifiesta la tendencia de los bebés a producir sonidos vocálicos en la parte central media y baja, y en la parte anterior de la cavidad bucal.

A mediados del siglo veinte Irwin (1948) (citado en Chen y Kent [2010]) realizó un estudio longitudinal en el que las vocales anteriores [ɛ, ɪ] y la central media [ə] daban cuenta del 70% de la producción vocálica del primer año de vida de noventa y cinco niños que adquirían la lengua inglesa y que fueron grabados cada dos meses desde el nacimiento hasta los 2;6 años (entre estas vocales la [ɛ] era la que más a menudo aparecía). Según este estudio, la tendencia en el desarrollo vocálico de estos niños era ir hacia un aumento de la producción de vocales posteriores desde la edad de 0;11 a 2;6 años de vida.

Veintidós años más tarde, Cruttenden (1970) estudió a sus hijas gemelas y encontró que las vocales [æ], [a] y [ə] predominaron a lo largo del balbuceo.

El mismo patrón presentaron ocho niños que adquirían la lengua sueca y que fueron grabados por Landberg y Lundberg (1989) entre el 0;1 y los 1;5 años de edad en intervalos de tiempo de dos semanas. Las vocales aisladas que dominaban las producciones tempranas de estos niños eran las vocales medias, centrales y anteriores. Los resultados mostraron que las vocales [ə, æ] fueron las más frecuentes durante el primer mes, ampliándose a [i, e, æ, a] al finalizar el primer año de vida. Además, las vocales no redondeadas predominaron sobre las

redondeadas durante todo el periodo de estudio y las vocales posteriores no aparecieron hasta el segundo año de vida.

Si ampliamos el inventario vocálico producido por el bebé y más allá de la vocal aislada consideramos aquellas vocales que aparecen en el interior de la palabra, tienden a observarse por lo general, los mismos resultados. En el trabajo de Landberg y Lundberg (1989) se cita el de Bickley et al. (1986) en el que las vocales de las palabras tempranas de catorce niños de entre 1;0 y 2;0 años de edad parecen agruparse una vez más en el cuadrante inferior izquierdo vocálico retardándose igualmente la aparición de las vocales posteriores.

En el estudio de Davis y MacNeilage (1990) si bien se mantiene la propensión hacia las vocales centrales y anteriores dentro del espacio vocálico, algunas vocales posteriores aumentan su presencia de manera considerable en comparación con otros trabajos. Después de extraer los sonidos vocálicos insertos en emisiones monosilábicas y disilábicas emitidas por bebés que adquirirían el inglés americano entre los 1;2 y los 1;8 años de edad (grabados semanalmente durante un periodo de dos horas), se observó que entre las vocales de los monosílabos no solo era frecuente la vocal neutra [ə] y la anterior alta [i] (siguiendo el patrón universal), sino que también lo eran [u, o]. En la posición de primera vocal en las palabras bisílabas, la vocal central baja [a] predominaba sobre el resto y [æ] y [ʌ] fueron relativamente frecuentes (sin que lo fuera la [ɛ], muy presente según reconocen los propios autores en otros estudios del balbuceo). En la posición de segunda vocal en las palabras bisílabas, la vocal anterior alta [i] y la vocal neutra [ə] fueron las más representadas.

Una posible explicación de la presencia más notoria de las vocales posteriores alta y semialta en las emisiones de los niños grabados por Davis y MacNeilage (1990) puede estar en el contexto de aparición vocálica. De esta forma, la extracción de la vocal a partir de la construcción de CV tanto monosilábica como disilábica frente a su consideración de forma aislada puede traer consigo que una vocal sea más común que otra porque el niño repita, sean cuales sean los motivos, la palabra en la que dicha vocal aparece y por tanto se repita en mayor medida dicha vocal. Este hecho podría explicar que la [o] sea una vocal usual en las producciones del niño a esta edad ya que según los autores el infante mostró una fuerte preferencia por la palabra *no* que apareció desproporcionadamente en su habla.

Cinco años más tarde Davis y MacNeilage (1995) volvieron a dar cuenta del tipo de vocal que de manera más usual el niño empleaba en el contexto de CV y VC tanto en monosílabos como en polisílabos, pero esta vez el periodo de observación y registro lingüístico fue mayor. Para ello extrajeron todas las vocales precedidas de una consonante, siendo la consonante una obstruyente, una glide o una sonorante, en seis niños que adquirirían la lengua



inglesa desde el comienzo del balbuceo canónico (que varió entre los 0;6 y los 0;8 años según el niño) hasta los 3;6 años (las grabaciones duraban una hora y eran semanales). En cuatro de los seis sujetos, dos o tres vocales dieron cuenta del 50% de la producción mientras que en los otros dos, esta frecuencia se distribuyó en distintos tipos de vocales. Una vez más, las vocales anteriores bajas y medias y las vocales centrales fueron las más articuladas por los sujetos siendo los porcentajes concretos de aparición vocálica (según la media de porcentajes entre los cinco transcritores) los siguientes: un 30,1% de aparición de vocales altas, un 27% de vocales medias y un 42% de vocales bajas (siendo estas nuevamente las que presentaron un porcentaje más alto). En relación a la anterioridad y posterioridad, los porcentajes fueron de un 69% para las vocales anteriores, de un 28% para las vocales centrales y de un 2,9% para las vocales posteriores (intuyéndose de nuevo un patrón universal de producción vocálica).

Las vocales centrales medias y bajas y las vocales anteriores fueron también las que predominaron en las producciones de trece niños de entre 1;2 y 2;0 años que adquirían la lengua árabe y que fueron grabados en el trabajo de <sup>2</sup>Amayreh y Diyson (2000), ya citado a propósito de la influencia de la lengua ambiente. Ellos examinaron el tipo de vocal que aparecía acompañada de una consonante en distintas posiciones de la palabra y exactamente el porcentaje medio de aparición de las vocales a lo largo del periodo estudiado fue el siguiente: la [a] presentó una frecuencia del 36%, y fue seguida de la [a:] con un 19,5%. La [i] se observó un 13,2% y la [e] un 8,9% quedando la [u] reducida a un 5,5%. Menor fue la frecuencia de estas últimas vocales cuando presentaban una mayor duración ([e:] un 4,3%, [u:] un 2,9% e [i:], 2,5%). Las vocales breves fueron por tanto más frecuentes que sus contrapartidas largas.

Los autores Stokes y Wong (2002) si bien se centraron en la correcta realización de la vocal por parte de cuarenta niños de entre 0;10 y 2;3 años que adquirían la lengua cantonesa, asunto que por salirse de los límites de esta tesis no trataremos, también documentaron la presencia de algunas vocales en determinadas edades. Según los autores el orden de aparición de las vocales en el bebé depende de dos factores: el factor articulatorio y el lingüístico o de la lengua ambiente. Por un lado, el criterio articulatorio establece que las vocales que se articulan con mayor facilidad aparecerán antes que las que se producen con una dificultad mayor; por otro y según sugirieron Pye et al. (1987), la frecuencia de aparición de un determinado sonido en la lengua ambiente puede determinar el orden de adquisición de los mismos. En base a estos

---

<sup>2</sup> Aunque como ya hemos anunciado en los artículos originales que se citarán a partir de ahora los autores utilizan para referirse a las vocales tanto barras oblicuas como corchetes (incluso dentro de un mismo trabajo), en esta tesis se emplearán estos últimos al referirse a sonidos como tales.

dos criterios y puesto que las vocales más frecuentes en cantonés son según Fok (1979) y por este orden /a/, /ɔ/, /i/, /ɐ/, /ɛ/, /œ/, /y/, /u/, Stokes y Wong (2002) predijeron que /a/, /ɔ/ e /i/ deberían ser vocales de desarrollo temprano.

Para comprobarlo grabaron las producciones de cuarenta niños hablantes de cantonés que fueron separados en los siguientes grupos de edad: de 0;10 a 1;1 años (grupo 1), de 1;3 a 1;6 años (grupo 2), de 1;8 a 1;11 años (grupo 3) y de 2;0 a 2;3 años (grupo 4) y transcribieron las vocales insertas en las palabras que dichos niños emitían (conocemos este dato porque los propios autores aclaran que el grupo de niños más jóvenes –grupo 1– no fue incluido en el estudio porque muchos de sus componentes no produjeron palabras reales).

A partir de sus grabaciones (de una hora de duración) se observó que entre los 1;3 y los 1;6 años las vocales [a, ɛ, ɔ] estaban presentes de la misma forma que entre los 1;8 y los 1;11 años las vocales [i, ɛ] habían aparecido. Entre los 2;0 y los 2;3 años las vocales [œ, u, y] estaban apareciendo. A la luz de estos resultados dos conclusiones pueden extraerse: por una parte que como en el trabajo de Bickley (1986) y frente al de Davis y MacNeilage (1990) la tendencia de los niños que adquieren cantonés en diferentes edades es la de producir un mayor número de vocales anteriores y centrales incluso dentro de la palabra; por otra, que las vocales [a] e [ɔ] aparecen en la primera etapa de grabación (1;3-1;6 años) y son seguidas por [i] que se presenta en la etapa inmediatamente posterior quedando la vocal [u] relegada al último periodo. Este resultado respalda la opinión de Stokes y Wong (2002) de que las vocales [a, ɔ, i] aparecen antes que otras debido por un lado a la facilidad de articulación que influye en las vocalizaciones del bebé entre los 1;3 y los 1;6 años, y por otro a la influencia que ejerce la lengua ambiente, en la que dichas vocales son más frecuentes, en las emisiones del bebé.

Es el factor de la lengua meta también el que puede favorecer el hecho de que la vocal [u] aparezca más tarde además de por presentar esta una mayor dificultad de realización (recordemos que se observa entre los 2;0 y los 2;3 años de edad). Un ejemplo similar lo encontramos en el trabajo de Tse (1991) en el que tras registrar el comportamiento vocálico de niños que también adquirirían la lengua cantonesa, se observó que la vocal posterior alta no apareció de manera significativa en las primeras etapas del desarrollo.

Un año más tarde, el estudio de Zmarich y Miotti (2003) manifestó las tendencias de centralismo y anterioridad vocálica de los trabajos ya comentados pero en la lengua italiana. Para ello los autores grabaron las producciones espontáneas en tres sujetos italianos a los 0;10, 1;0, 1;2, 1;4 y 1;6 años de edad en sesiones de cuarenta y cinco minutos que tenían lugar cada dos semanas. Los criterios que se siguieron para considerar las emisiones como balbuceo fueron los de Oller (1986) y se utilizaron los de Vihman y McCune (1994) para las palabras. Las

estadísticas mostraron una preferencia por la vocal central baja [a] y por la anterior semialta [e] con ausencia de vocales redondeadas posteriores al menos hasta la etapa de la segunda palabra.

El estudio de Chen (2005), al que no hemos podido tener acceso y el cual solo conocemos porque se menciona en otros trabajos, viene a corroborar las preferencias vocálicas generales de los niños puesto que en él la vocal anterior semibaja [ɛ] y la central media [ə] fueron las más comunes en el balbuceo característico de un niño que adquiriría el mandarín antes de los 0;8 años. Por otro lado, el trabajo de Dawn Warner-Czyz (2005), anteriormente citado para subrayar la importancia de la percepción en la producción temprana, manifestó que también los niños que fueron sometidos a un implante cloquear a la edad de 0;6 años produjeron entre los 0;11 y los 1;1 años vocales centrales medias y bajas con una mayor frecuencia coincidiendo con las producciones de los niños con desarrollo normal con los que fueron comparados.

Dos artículos que sirven de puente entre aquellas investigaciones que se limitan a registrar el inventario de los sonidos vocálicos de la lengua del bebé y aquellas otras que analizan acústicamente el primer y segundo formante, son el de Kent y Bauer (1985) y el de Selby et al. (2000). En el primero, se realizaron análisis acústicos de las emisiones vocálicas pero no se midieron los formantes; en el segundo, y a pesar de que sí se analizaron acústicamente los formantes, no se consideraron sus frecuencias para situar las vocales en el espacio formántico ni se contempló si se había producido un descenso significativo de los mismos a lo largo del tiempo (tan solo se emplearon dichos valores para que el transcriptor identificara mejor la vocal que percibía). Veámoslos con más detenimiento.

En cuanto al primer estudio, Kent y Bauer (1985) grabaron las emisiones producidas por cinco niños que adquirirían la lengua inglesa a la edad de 1;1 años en contextos de interacción natural madre e hijo (una hora de grabación de cada niño servía como material de análisis). A continuación, extrajeron las vocales aisladas, los diptongos y las construcciones de CV, VC y CVC y realizaron un análisis acústico para observar los contornos de entonación, y uno fonético o de transcripción para establecer qué vocales emitía el bebé. Los resultados mostraron que las *vocants* (término empleado por algunos autores para denominar los sonidos vocálicos que aún no son como los del adulto y a cuya utilización ya nos remitíamos en el capítulo dedicado a la metodología) que más frecuentemente aparecían en la forma de vocal aislada o de vocal más vocal, esto es diptongo, eran por este orden: [ʌ], o [ə], [ɛ], [æ], [ʊ], [e], [ɑ] y [o]. Las vocales centrales y anteriores se prefirieron a las posteriores, y las vocales medias y bajas predominaron sobre las altas. Llama la atención según Kent y Bauer (1985) el hecho de que que las vocales [i] e [u] apenas fueran producidas por los niños a los 1;1 años de edad. Estos resultados siguen la

tendencia hallada en otros estudios ya revisados en cuanto a la frecuencia de determinadas vocales según la dimensión anterior-posterior y el eje alto-bajo de la lengua, y en cuanto a la escasa aparición de las dos vocales altas de los extremos. Los propios autores compararon el producto de sus análisis con el de otros estudios como el de Winitz e Irwin (1958) quienes observaron que en la producción de las primeras palabras por veintitrés niños de entre 1;1 y 1;2 años, la [i] y la [u] ocurrían el 4% o 5% de las veces mientras que la [ɛ] se aproximaba al 16%, la [ʌ] al 18% y la [ɑ] al 20%.

Por otro lado, en el trabajo de Selby et al. (2000) se grabaron las emisiones de cuatro niños hablantes de inglés americano desde los 1;3 a los 3;0 años a las siguientes edades: a los 1;3, 1;6, 1;9, 2;0 y a los 3;0 años (una sola sesión de grabación de una hora de duración fue realizada en cada edad). Las vocales se extrajeron de vocalizaciones o palabras inteligibles y no inteligibles y podían ir acompañadas de una o varias consonantes (vocales no aisladas). Los resultados fueron los siguientes: las vocales que aparecían un 75% de las veces en las producciones de los cuatro niños a lo largo del tiempo fueron: a los 1;3 años, [ɑ, ɪ, ʊ, ʌ]; a los 1;6 años, [ɑ, i, u, ʊ, ʌ, ɔ, æ]; a los 1;9 años, [ɑ, i, ɪ, ε, u, o, ʌ, ɔ]; a los 2;0 años, [ɑ, i, ɪ, ε, e, u, o, ɔ, æ]; y a los 3;0 años, [ɑ, i, ɪ, ε, e, u, ʊ, o, ʌ, ɔ, æ, ɜ].

Tres observaciones pueden hacerse a partir de los datos anteriores: por un lado, los tipos vocálicos van aumentando conforme avanza la edad de los niños de forma que de los 1;3 años a los 3;0 años el número de vocales fue de cuatro a doce; por otro lado, se produjo un aumento de las vocales ya que se pasó de una única vocal a los 1;3 años, vocal [ʌ], a seis vocales a los 3;0 años, vocales [ɛ, e, o, ʌ, ɔ, ɜ] (hay que recordar que dicho estudio solo recoge las vocales que aparecieron al menos dos veces en cada sesión de grabación, criterio que se tomó para que una vocal fuera incluida como parte de su inventario y que impide observar si el niño produjo otras vocales medias a los 1;3 años las cuales no han quedado reflejadas posiblemente porque no las repitió durante el tiempo de grabación); y finalmente y de forma llamativa, se encontró que el grupo de vocales posteriores fue más alto que el de vocales anteriores a los 1;3 años y a los 1;6 años de edad. Entre las razones que pueden explicar este dominio de las vocales posteriores sobre las anteriores en determinados meses de edad quizás se encuentre como en el trabajo de Davis y MacNeilage (1990) la de que al haber extraído la vocal de vocalizaciones sin sentido pero también de palabras inteligibles, el niño puede manifestar en estos meses la preferencia por una determinada palabra que contenga vocales posteriores (y que puede representar un nuevo hábito adquirido, un nuevo juguete, etc.) y al repetirla con mayor frecuencia se registre un mayor número de este tipo de vocales en esos meses.

Finalmente, puede observarse que las vocales de los extremos [i, u, a] fueron articuladas por todos los niños a los 1;6 años de edad, circunstancia que pudo observarse también en los trabajos de Wellman et al. (1931), Templin (1957), y Davis y MacNeilage (1990). En esta ocasión, Selby et al. (2000) consideraron temprana la aparición de las vocales de los extremos y citando a Kent (1992) comentaron que este hecho podría facilitar al bebé la adquisición de nuevas vocales debido a los límites acústicos y articulatorios que marcan dichas vocales.

### **2.3.3. Revisión de estudios de análisis acústico.**

Por lo que respecta a las investigaciones en las que se ha llevado a cabo un análisis acústico de las vocalizaciones de los bebés y de los niños pequeños, es necesario que indiquemos que todos los trabajos que hemos revisado y que nos disponemos a comentar, a excepción del artículo de Kent y Murray (1982) donde simplemente no se concreta y del de Kuhl y Meltzoff (1996) donde las vocales aparecen separadas de las consonantes, han analizado los formantes de los sonidos vocálicos acompañados de otros consonánticos ya sea en sílabas aisladas (del tipo CV o CVC), ya insertas en la palabra. Este hecho, tal y como comentamos en la metodología, si bien aumenta el número de vocales que pueden analizarse (pues a partir de un periodo determinado son más comunes las vocales acompañadas de consonantes que las aisladas), también establece una diferencia sustancial entre nuestro estudio que analiza los formantes y el espacio vocálico a partir de sonidos vocálicos aislados, y los expuestos a continuación. No obstante, la ausencia casi completa de estudios similares a los nuestros nos ha obligado a contemplar los resultados de estos últimos para poder establecer una comparativa con nuestros datos. Presentamos a continuación una revisión de estos trabajos según un criterio cronológico tal y como venimos haciendo en este capítulo del estado de la cuestión.

El primer estudio en el que se analizaron los formantes de las vocales de los niños pequeños fue el de Peterson y Barney (1952). En él, setenta y seis sujetos hablantes de inglés americano, entre ellos treinta y tres hombres, veintiocho mujeres y quince niños (algunos de ellos nacieron fuera de Estados Unidos y otros aprendieron una lengua extranjera antes que el inglés) tenían que producir una lista con diez palabras monosilábicas que comenzaban con la aspirada [h] y acababan con [d], y que diferían solo en la vocal: [i, ɪ, e, ε, æ, a, ɔ, ʌ, u] (después esta lista fue escuchada por distintos oyentes a los que se les pidió que escribieran los sonidos que habían percibido). Los análisis mostraron que las frecuencias de los formantes de los niños eran sobre la mitad y una octava mayores que las de los hombres. Esta diferencia en los valores formánticos puede deberse, según los estudios comentados al principio de esta revisión de conocimientos, a las características particulares del tracto vocal de los niños que provocan un aumento de los valores formánticos en relación a los del adulto.

El siguiente artículo llevó a cabo un estudio longitudinal de las vocales [i], [æ], [u], [ɛ], [a] y [ə] en sujetos de entre 3;0 y 13;0 años que adquirirían el inglés americano (el sonido de [a], no encontrado normalmente en muchos dialectos americanos, fue tomado de la primera porción del diptongo [ai] en la palabra *I* “yo”). Se trata del trabajo de Eguchi y Hirsh (1969) en el que ochenta y cuatro sujetos agrupados en las edades de 3;0, 4;0, 5;0, 6;0, 7;0, 8;0, 9;0, 10;0, 11;0, 12;0 y 13;0 años debían producir las oraciones *He has a blue pen* “Él tiene un bolígrafo azul” y *I am tall* “Yo soy alto”, y repetir las en cinco ocasiones (el grupo de edad de sujetos más jóvenes primero hubo de escucharlas para después repetir las). De estas repeticiones se extrajeron las vocales en las que se midieron los dos primeros formantes para observar su evolución y realizar posteriormente algunos análisis estadísticos como el de la desviación estándar de los valores entre sujetos y dentro de un mismo sujeto (inter e intra-sujeto). Los resultados mostraron un claro descenso de los dos primeros formantes a lo largo del periodo de estudio, siendo este más acusado entre los 3;0 y los 5;0 años de edad. Por otro lado, el segundo formante tendió a bajar más que el primero en todas las vocales excepto en la [a].

Estos datos se desprenden de los valores concretos del F1 y F2 dados para cada una de las vocales y de entre los cuales nos interesan aquellos pertenecientes a las vocales compartidas por el español que son precisamente los de las tres vocales de los extremos.

**Tabla 2.1. Valores de los dos primeros formantes de las vocales de Eguchi y Hirsh (1969).**

		i	a	u
3;0	F1	484	986	578
	F2	3318	1982	1664
4;0	F1	444	879	472
	F2	3050	1858	1528

En cuanto a la desviación estándar entre sujetos, la edad no resultó ser un factor determinante de forma que la media individual de los formantes no se hacía más parecida a la de otros sujetos conforme estos se hacían mayores. Por tanto, otras razones además de las de crecimiento anatómico deberían aducirse para explicar este hecho. En lo que respecta a la variabilidad dentro de un mismo sujeto, esta decreció para todas las vocales a lo largo del tiempo estudiado, es decir, se produjo un descenso con la edad, aunque la variabilidad de las frecuencias del primer formante fue mayor para las vocales medias y para la vocal baja [a], que para las vocales anteriores y posteriores altas ([i, u]). Alrededor de los 11;0 y los 12;0 años la variabilidad de las distintas emisiones de un mismo niño fue similar a la de los adultos. Según

los autores esto evidenció que los niños más jóvenes no son tan exactos en sus posiciones articulatorias como lo son los niños mayores y los adultos (1969:44):

“[...] The implication is clear that 3-year-old children do not move their tongues to exactly the same position for a particular vowel as it occurs in repetitions of the same sentence, at least not so exactly the same as older children and adults. The variability decreases until about age 11”.

Finalmente, y aunque quede fuera de nuestro objeto de estudio, acabamos la revisión de este artículo aludiendo al hecho de que los formantes de las niñas de 13;0 años fueron muy similares a los formantes de los hablantes femeninos adultos, mientras que los de los niños de 13;0 años tuvieron frecuencias más altas que las de los formantes de los adultos varones.

Gilbert (1970) llevó a cabo un estudio transversal para comparar los formantes de niños con desarrollo lingüístico normal con los de niños que presentaban cierto retraso. Para ello se analizaron las emisiones de cinco niños de 4;0 años de edad que adquirirían el inglés americano y que presentaban habilidades lingüísticas normales según distintas pruebas tales como el test estándar de articulación del sonido del habla de Templin-Darley (1960), la prueba del nivel de logro lingüístico medida por un test de vocabulario de Dunn (1965), y un test de habilidades psicolingüísticas de McCarthy y Kirk (1961). Por otro lado, los niños mostraron un umbral de escucha normal según las pruebas de la *International Standards Organization* (1964). Después se eligieron otros cinco niños de la misma edad con un nivel de audición normal pero con cierto retraso en las pruebas de articulación del sonido y de logro lingüístico. Los dos grupos de niños tenían que pronunciar cuatro palabras monosilábicas que comenzaban por el sonido [h] y acababan con el sonido [d]. Las palabras fueron *heed* (vocal [i]), *had* (vocal [æ]), *hod* (vocal [a]) y *who'd* (vocal [u]) y estas fueron asignadas a cuatro dibujos sin sentido con los que se familiarizó a los pequeños.

La media de los formantes de las vocales emitidas por los niños con desarrollo normal y compartidas por el inglés y el español fueron: [i] tuvo un F1 de 442 Hz y un F2 de 2510 Hz; [a] registró un F1 de 693 Hz y un F2 de 1246 Hz; y [u] presentó un F1 de 539 Hz y un F2 de 1255 Hz. Los valores del grupo de niños que manifestó cierto retraso en las pruebas realizadas fueron: [i] tuvo un F1 de 555 Hz y un F2 de 2613 Hz; [a] registró un F1 de 727 Hz y un F2 de 1261; y [u] alcanzó un F1 de 533 y un F2 de 1336 Hz. Los ANOVA (*Analysis Of Variance* o ‘Análisis de Varianza’) realizados sobre los conjuntos de datos revelaron que no había diferencia significativa en las medias de los formantes de ambos grupos de niños. Dicho resultado implicaba según Gilbert (1970) que el F1 y el F2 no estaban bien correlacionados con el nivel de desarrollo lingüístico sino con el nivel de crecimiento fisiológico entroncando directamente

con la teoría acústica de Stevens y House (1961) según la cual la frecuencia de los formantes de las vocales depende de las dimensiones del tracto vocal.

Una década más tarde Lieberman (1980) combinó grabaciones semanales y bisemanales para recoger las emisiones espontáneas de cuatro niños que adquirían la lengua inglesa entre las edades de 0;4 y 5;0 años en contextos naturales: al principio recogían las emisiones del bebé mientras sus padres los bañaban o los alimentaban, después las emisiones del bebé se produjeron mientras la madre le leía un cuento con dibujos, y en los niños mayores el habla se obtuvo de las conversaciones entre estos y cualquier adulto que les rodease. A continuación ofrecemos un resumen de los resultados obtenidos después de que un test de percepción y un análisis espectrográfico de banda ancha en el que se midió el F1 y el F2 de las vocales aisladas, en el balbuceo e insertas en palabras con significado, fueran realizados para cada uno de los niños. No obstante, nos resulta imposible mostrar los resultados de cada niño en el mismo periodo temporal puesto que el estudio solo nos muestra los resultados obtenidos en un periodo de tiempo concreto y diferente para cada sujeto (de ahí que los mostremos por separado).

Los datos mostrados para el sujeto L. S. cubren la franja temporal que fue desde los 0;4 a los 1;4 años abarcando todo el periodo de balbuceo (luego las vocales analizadas para este niño se extrajeron del balbuceo canónico y variado). Las vocales de los extremos [i], [a] y [u] fueron infrecuentes en las vocalizaciones tempranas y ello coincide según los autores con los datos de George (1978) quien documentó una completa ausencia de dichas vocales en veintinueve niños con desarrollo normal antes de los 0;3 años, y una escasa aparición de estas en las semanas siguientes. En este periodo no hubo un descenso significativo del F1 ni del F2 y según Lieberman (1980) esto se debió a que el tracto vocal supralaríngeo apenas aumentó su longitud en dicho periodo. Por otro lado, el niño no tiene forma de determinar la configuración supraglótica del tracto vocal que subyace a la producción de un sonido específico por lo que tiene que imitar la señal acústica que percibe para producir ese sonido. Los valores de los formantes de este sujeto quedan fuera del rango de valores que se da en el habla del adulto y esto es consistente con la corta longitud del tracto vocal supralaríngeo del niño. En cuanto al desarrollo del espacio vocálico, a los 0;4 años encontramos una clara superposición de las vocales [ɪ, ɛ, æ, ʊ]. A los 0;6 años parece como si las vocales producidas por L. S. se estuvieran moviendo hacia unas posiciones acústicas que se asemejan a un rudimentario triángulo vocálico. A los 0;10 y 1;3 años el espacio vocálico sigue expandiéndose. A los 1;4 años un total de cuarenta y ocho vocales fueron no identificadas como vocales del inglés americano.

Para el sujeto R. C., que en este caso era una niña, los análisis acústicos que se realizaron entre los 0;9 y los 1;5 años mostraron que el espacio vocálico de dicho sujeto se



expandió entre estas edades (como en el caso de L. S, las vocales fueron extraídas del contexto de balbuceo).

El análisis espectrográfico de las emisiones del sujeto J. B. fue realizado sobre vocales extraídas de palabras con significado en el periodo que fue desde los 1;4 a los 3;0 años. De nuevo una expansión del espacio vocálico se produjo entre esas edades acercándose más dicho espacio al espacio acústico adulto.

En el caso del niño F. R. las vocales se extrajeron de oraciones con sentido entre los 2;7 y los 3;6 años. Los análisis mostraron una expansión de su espacio vocálico también entre estas edades. Hubo un descenso del F1 para la [i] y la [ɪ] a lo largo del periodo de análisis y todas las vocales se movieron hacia una posición más correcta en relación al espacio vocálico de los hablantes adultos de inglés americano (en particular las vocales [ɪ, ε, æ] estuvieron mejor alineadas en el eje agudo, y las vocales [ʌ, o] en el grave. Este resultado prueba según Lieberman (1980:151) que: “[...] The trend towards ‘perfecting’ the acoustic vowel space thus continues from babbling well into the phonologic and syntactic stages of language acquisition”.

Al comparar los resultados de los niños en distintas edades se observó que los datos de los formantes de la vocal [i] para el hablante J. B. entre los 1;4 y los 2;3 años poseían un mayor rango de variación que los valores de los formantes obtenidos en el sujeto F. R. entre los 2;7 y los 3;4 años de edad. La duración de las vocales de J. B. fue también mayor que la de F. R. que se aproximó más a la de los niños mayores y adultos. Por otro lado, los valores de los formantes de las vocales de F. R. fueron más bajos que los de J. B. debido a que aquel era mayor en edad y en tamaño que J. B. Del mismo modo, el F3 para L. S. fue de unos 6.5 kHz en la vocal [i] entre los 1;0 y los 1;3 años, de 5.5 kHz para el hablante J. B. entre los 1;8 y los 2;0 años, y de 4.5 kHz para el sujeto F. R. entre los 2;7 y 3;6 años. Es decir, “[...] as the children grow, their formant frequencies fall” (1980:155). La edad también es según Lieberman la responsable de la mejora gradual y consistente en la producciones de los niños de las vocales del inglés que se produce desde el balbuceo hasta los 3;0 años.

En el mismo año, Buhr (1980) llevó a cabo un estudio vocálico de carácter longitudinal de un solo sujeto al que se grabó dos veces por semana en interacción espontánea con su madre, y entre las edades de 0;4 y 1;4 años. En dicho estudio se realizaron análisis espectrográficos de banda ancha y banda estrecha en los que se midieron los tres primeros formantes de las vocales del habla del bebé tomadas unas veces aisladas y otras extraídas del contexto CV. Se hizo también un doble test de percepción. En cuanto a la frecuencia de aparición de estas vocales, la [ε] fue la vocal más frecuente apareciendo unas doscientas noventa y cinco veces (el 31% de la muestra), seguida de la [æ] que apareció ciento sesenta y cuatro veces (17%), de [ɪ] que se

mostró unas ciento catorce veces (12%), a esta le siguió la [ʌ] con ciento seis veces (11%), después la [ʊ] con ciento cinco veces (11%), la [i] apareció cuarenta y nueve veces (5%), muy cerca estuvo la [a] con cuarenta y siete veces (también un 5%), le sucedió la [e] con cuarenta y seis veces (de nuevo 5%), y finalmente la [u] y la [ɔ] que aparecieron cinco y cuatro veces respectivamente.

Por lo que respecta a la relación entre el F1-F2-F3, los resultados fueron los siguientes: para la [ɛ] y la [æ] el F1 se mantuvo constante a lo largo del periodo de estudio frente al F2 y F3 que subieron y bajaron constantemente sin una tendencia monótona. El mismo patrón presentó la [ʊ] con la diferencia de que en esta vocal el F3 varió unos dos mil hertzios a lo largo del tiempo analizado. En el caso de la [i], tanto el F2 como el F3 presentaron una gran variación con un rango de entre 2250 Hz y 3875 Hz para el F2, y de 4000 Hz y 6250 Hz para el F3. Las vocales [ʌ] y [e] apenas mostraron cambios significativos en ninguno de los tres formantes. Llama la atención la alta variación de los valores de los tres formantes encontrada para la vocal [i] yendo el F1 de 250 Hz a 1000 Hz, el F2 de 3000 Hz a 4500 Hz, y el F3 de 4750 a 7500 Hz. Lo mismo ocurrió en la [a] donde se observó una variación del F1 de unos 875 Hz (variando más que el F2 y el F3). Estos resultados chocan con la consideración de la [i] y la [a] de vocales cuánticas según la teoría de Stevens (1972). Por el contrario, la [u] pareció cumplir con la condición de vocal cuántica ya que los valores de sus formantes no cambiaron de manera significativa. En el caso de la [ɔ] y la [o] y debido al reducido número de muestras, no se pudieron establecer conclusiones sobre el cambio en los formantes.

A partir de los gráficos de evolución formántica de los diferentes sonidos vocálicos (o cambios en los valores de los formantes de las distintas vocales), podemos observar que la relación entre el F1 y el F2 se hace consistente para ciertas vocales a partir de determinados meses. Es el caso de la [a], [ʌ] y [æ] cuya relación formántica se mostró más estable a mitad de los 0;9 años, de la vocal [ɛ] a partir de los 0;8 años, y de la vocal [i] pasado el primer año de vida.

En cuanto a la posición que ocuparon las vocales en el espacio formántico, percibimos con ayuda de las cartas de formantes una superposición de todas las vocales en el espacio acústico, excepto de la [i], a la edad de 0;4 años. A los 0;6 años pareció surgir un rudimentario espacio vocálico delimitado por la [u], por la [æ] y por la [i] (esta última vocal se superpuso a la [i] a la edad de 0;10 años). A partir de los 0;10 años el esparcimiento de las vocales comenzó a estrecharse con excepción de la [ɛ], y apareció la [i] como última vocal de la dimensión aguda. Al final del periodo también se observó la incorporación de las vocales del eje posterior, sonidos estos que tardaron más en aparecer (a excepción de la [ʊ]).

Si nos detenemos en la presencia de determinadas vocales, advertiremos que a partir de los 0;4 años la *schwa* hace su aparición y con ello puede inferirse según Buhr (1980) que la lengua ha adquirido algún grado de flexibilidad para poder replicar el tubo resonador necesario para articular esta vocal. Esto contrasta con la poca movilidad y capacidad para acomodarse de la lengua en las primeras semanas que siguen al nacimiento debido a que su gran tamaño con respecto a la cavidad bucal, restringe el movimiento de esta (según Crelin [1972] la lengua mide al nacer unos 4 cm, y ocupa un espacio de 7,5 cm, aproximadamente, según Lieberman et al. [1971]). Todo ello sugiere según el autor que la infancia temprana puede caracterizarse por un desarrollo gradual de la habilidad neuromuscular de la lengua junto con una expansión y un cambio en la forma de la cavidad oral acompañados de cambios en el movimiento de la mandíbula para la producción de las emisiones.

Respecto a la [i], y a pesar de que se presentó de forma intermitente, esta apareció a lo largo del corpus sugiriendo que el eje agudo del triángulo vocálico tiene como base la porción del tracto vocal cuya existencia y control están desde el nacimiento. La temprana aparición de vocales anteriores cuyos rangos formánticos son además similares, como es el caso de [i], [e] y [ɛ], es consistente con las observaciones de Lindblom y Sundberg (1969) quienes justificaron el pronto desarrollo de dichas vocales alegando que estas debían ser producidas únicamente por la manipulación de la posición de la mandíbula. Según Burh (1980), causas anatómicas podrían estar detrás de este desarrollo temprano del eje vocálico agudo frente al grave tales como que la mandíbula y la parte frontal de la lengua parecen desarrollarse a una velocidad mayor que el descenso de la laringe y la parte de atrás de la lengua, procesos estos que pueden llevar varios años. Dicha observación también coincide con los resultados de Irwin (1946, 1948) quien se dio cuenta de que en la temprana infancia las vocales anteriores eran las más frecuentes. Sin embargo, algunas diferencias fueron halladas entre el trabajo de Irwin (1948) y el de Burh (1980): en el estudio de Irwin (1948) existieron menos muestras de [v] antes de los cinco o seis meses, o de [æ] antes de los 0;11 o 1;0 año, mientras que en el sujeto de Buhr (1980) se detectaron estos dos sonidos desde los 0;4 años y presumiblemente desde algún tiempo antes. Este resultado puede ser debido a simples preferencias o diferencias individuales.

En lo que confiere a la [u], las pocas muestras de esta vocal pueden explicarse debido a que la coordinación del movimiento de los labios, la lengua y la mandíbula necesaria para su producción aún no se ha desarrollado de una manera consistente. Sin embargo, en el trabajo de Buhr (1980) se indica que el movimiento longitudinal de la lengua necesario para el tragado emerge en el periodo postnatal (Bosma [1975]) y que la musculatura responsable del movimiento de la mandíbula también parece estar desarrollándose durante este periodo como evidencia la producción de sonidos agudos. Por ello, quizás una explicación más convincente de

la escasez de este timbre vocálico sea la de que los músculos faciales que inervan el *orbicularis oris* y que son los responsables de que los labios se plieguen o doblen, no se hayan desarrollado de la manera que se requiere para la producción de la [u]. Ello no imposibilita que la [ʊ] sea producida debido a que para la articulación de esta vocal la protusión de los labios no es requerida. Puede ser por tanto y según el autor, que el control de los movimientos del labio no se desarrolle hasta más tarde a pesar de que el movimiento de este que se requiere para la succión de la leche materna por ejemplo, esté presente desde el nacimiento.

Kent y Murray (1982) analizaron las emisiones de veintiún sujetos que adquirirían el inglés americano en tres edades diferentes: a los 0;3, 0;6 y 0;9 años de edad (uno de los sujetos de 0;9 años excedía esta edad ya que tenía nueve meses y veintinueve días el día de la prueba). Los sujetos fueron grabados en presencia del investigador mientras estaban sentados en unas sillas, que se colocaron junto a sus madres, en una única sesión de grabación en la que se utilizaron juguetes para estimular la producción de habla. Análisis espectrográficos de banda ancha (450 Hz) y de banda estrecha (45 Hz) sirvieron como herramientas para examinar los formantes, la duración y los patrones de entonación de las emisiones vocálicas o *vocalic utterances* que según los autores se corresponderían con las *vocants* en la terminología propuesta por Martin (1981), analizándose por tanto vocales aisladas tal y como se procede en nuestro estudio.

Los resultados mostraron la media de los valores formánticos en cada una de las tres edades y esta fue la siguiente: a los 0;3 años el F1 presentó unos valores que fluctuaban entre los 500 Hz y los 1500 Hz para todos los sonidos vocálicos emitidos, y el F2 registró unos valores que se situaron entre los 1800 Hz y los 3800 Hz; a los 0;6 años el F1 varió entre los 500 Hz y los 1700 Hz situándose el F2 en el mismo periodo entre los 1600 Hz y los 3800 Hz; a los 0;9 años el F1 se movió entre los 500 Hz y los 1800 Hz, y el F2 estuvo entre los 1400 Hz y los 4100 Hz. Estos datos indican que los valores del F1 y del F2 aumentaron con la edad aunque según los autores la mayoría de los sonidos vocálicos producidos a los 0;9 años de edad tuvieron el mismo patrón formántico que los emitidos a los 0;3 años. Además, aunque puede apreciarse una ampliación del espacio vocálico, delimitado por el F1 y el F2, conforme va aumentando la edad de los bebés, es arriesgado según los autores hacer inferencias sobre estos datos ya que el centro de dichos formantes (F1-F2) no cambió de forma marcada a lo largo del periodo de tiempo. Por otro lado, según Kent y Murray (1982) los valores obtenidos por Buhr (1980) están dentro de los que ellos dan en los mismos periodos de tiempo.

Después de reflejar las frecuencias de los formantes de las vocales seleccionadas en cada mes en gráficos que simulaban cartas de formantes, se observó que la mayoría de los

sonidos vocálicos producidos por los bebés entre los 0;3 y 0;9 años se articulaban en la parte central o anterior media de la boca. También se contempló la presencia de vocalizaciones más nasalizadas a los 0;3 años que a los 0;9 años, hecho que se debe según Kent y Murray (1982) a razones anatómicas como la modificación de la relación velo-faríngea entre los 0;3 y los 0;5 años de edad (Sasaki et al. [1977]).

En el mismo año Bond et al. (1982) dirigieron un estudio longitudinal en el que se recogieron las emisiones espontáneas de una niña que adquiría inglés americano. En cuanto a la metodología y a diferencia de los trabajos anteriores, en este estudio las grabaciones no se realizaron en la casa familiar sino que tuvieron lugar en la universidad de Ohio y en una clínica de problemas de audición. La niña fue grabada a los 1;5, 1;7, 1;10, 2;2 y 2;5 años de edad, periodos que se agruparon en tres etapas: la etapa I fue de los 1;5 a los 1;10 años, la etapa II comprendió los 2;2 años, y la etapa III los 2;5 años. Se analizaron espectrográficamente los dos primeros formantes de todas las vocales anteriores y posteriores insertas en palabras identificables por los adultos (por lo que no se analizaron los sonidos vocálicos aislados ni aquellos que aun acompañados por consonantes no constituyeron palabras reconocibles por los investigadores). Los resultados revelaron una superposición de los formantes de todas las vocales, incluyendo los formantes de las vocales de los extremos [i], [æ], [ɑ] y [u], en las etapas primera y segunda, superposición que descendió considerablemente en la etapa tercera. Además, en esta última etapa las vocales de los extremos se hicieron relativamente distintas. Los autores recuperaron las observaciones de Fletcher (1973) según las cuales la superposición entre las vocales muestra los imprecisos gestos articulatorios que poseen los bebés en las dos primeras etapas. Esta falta de precisión descendió en la etapa tercera sugiriendo un mayor control articulatorio que se reflejó en un espacio vocálico más definido.

En el trabajo de Kent y Bauer (1985) se grabaron las emisiones de cinco niños de 1;1 años de edad en contextos familiares en los que la madre debía hablar con el niño sobre una colección de juguetes y sobre una serie de comidas que el investigador le había facilitado a esta previamente (luego la grabación fue de habla espontánea pero al mismo tiempo fue de habla guiada o dirigida). Los resultados mostraron que las vocales que más frecuentemente aparecían eran por este orden: [ʌ], o [ə], [ɛ], [æ], [ʊ], [e], [ɑ] y [o]. Así, las vocales centrales y anteriores prevalecieron sobre las vocales posteriores, y las vocales bajas predominaron sobre las vocales altas ya que la [i] y la [u] apenas fueron producidas por los niños. En cuanto a los análisis acústicos, y aunque se realizaron espectrogramas de banda ancha y de banda estrecha, estos solo se utilizaron como refuerzo a la transcripción de los segmentos por lo que no se aportaron los valores concretos ni del F1 ni del F2.

Recuperamos en este punto el estudio de Boysson-Bardies et al. (1989), ya comentado en el apartado de la influencia de la lengua ambiente en las vocalizaciones de los niños, para tratar ciertos aspectos acústicos. En dicho artículo veíamos que las tendencias universales como la propensión a la centralización del espacio acústico y la mayor frecuencia de las vocales anteriores frente a las posteriores comenzaban a ser modificadas a los 0;10 años por las características de un ambiente lingüístico determinado en niños de cuatro comunidades lingüísticas diferentes. En este apartado sin embargo, nos detendremos en los valores de los formantes que arrojaron los análisis acústicos y que según Boysson-Bardies coincidieron con los ofrecidos en los estudios de Lieberman (1980), Buhr (1980) y Kent y Murray (1982).

Según el análisis de varianza conocido como *test de Student* que llevaron a cabo estos autores, la variabilidad inter-sujeto fue más alta que la intra-sujeto tal y como reflejaron los siguientes datos: la media del F1 de todos los sonidos vocálicos producidos por el niño francés fue de 758 Hz frente a los 1173 Hz de un niño de Hong Kong; por otro lado, la media del F2 fue de 2061 Hz en el niño de Hong Kong frente a los 2805 Hz del niño inglés. Estos resultados pueden reflejar la preferencia de los niños pertenecientes a comunidades lingüísticas diferentes por una vocal determinada. De esta forma, el niño hablante de cantonés parece manifestar una preferencia por las vocales bajas, la cual condiciona los valores de sus formantes provocando que estos sean más altos en el caso del F1 y más bajos en el del F2. Debe tenerse en cuenta por tanto esta inclinación hacia determinadas vocales o esta influencia de la lengua meta a la hora de comparar los valores formánticos de las emisiones de bebés provenientes de diferentes ambientes lingüísticos porque en el caso de que no coincidieran, esta podría ser una explicación convincente.

Por otro lado, y aunque no tan significativa como la variabilidad hallada entre los sujetos que adquieren una lengua meta diferente, en el trabajo de Boysson-Bardies et al. (1989) se observaron diferencias en las frecuencias de los formantes de los niños pertenecientes a una misma comunidad lingüística. Así por ejemplo, el F2 de uno de los niños ingleses fue de 2505 Hz frente al F2 de uno de sus compañeros que fue de 2805 Hz. Algunos hertzios más separaron el F2 de un niño argelino, que fue de 2170 Hz, del F2 de otro niño de su misma comunidad lingüística que registró los 2599 Hz. Pero es sin duda en el grupo de cantonés en el que se encontraron las mayores diferencias puesto que el F2 de uno de los cuatro niños fue de 2061 Hz y el de otro fue de 2566 Hz (en el F1 las diferencias no fueron tan altas). Este hecho no puede ser explicado por una lengua meta diferente ya que todos los niños pertenecían a la misma comunidad lingüística y ello nos hace pensar que quizás las preferencias individuales de cada infante por unas u otras vocales dentro del mismo ambiente lingüístico puedan ser las causantes de este desbarajuste formántico.

Dicha variabilidad intra-sujeto también fue hallada en el estudio de Clement y Wijnen (1994) que incluyó cuatro niños con desarrollo normal entre los 2;0 y los 2;6 años, cuatro niños con desarrollo normal entre los 4;0 y los 4;6 años, cuatro niños con retraso fonológico de la misma edad (4;0-4;6 años) y un grupo de cuatro adultos. Los sujetos, hablantes de holandés, debían producir tres palabras que no existían en dicha lengua y que se correspondían con tres figuras de animales: /tup/, /tap/ y /top/ (luego la vocal no se analizó aislada sino en el contexto de CVC). Los resultados mostraron una amplia variabilidad entre los individuos de un mismo grupo de edad tanto en los parámetros espectrales, esto es en el F1 y el F2, como en los parámetros temporales de la producción del habla. Los autores aludieron a Eguchi y a Hirsh (1969) que sugirieron que la variabilidad de los formantes refleja la precisión con la que los hablantes colocan sus articuladores para producir unas metas acústicas determinadas.

El estudio también incluyó los valores medios de los formantes para cada una de las vocales según el grupo de edad. Estos fueron:

**Tabla 2.2. Valores de los dos primeros formantes de las vocales de Clement y Wijnen (1994).**

Sujetos	Edad	Formantes	a	o	u
Niños con desarrollo normal	2;0-2;6	F1	1297	998	480
		F2	2280	2000	1560
Niños con desarrollo normal	4;0-4;6	F1	1120	834	441
		F2	2208	1600	1216
Niños con retraso fonológico	4;0-4;6	F1	1148	1004	468
		F2	2152	1959	1438

Por otro lado, y aunque queda fuera del alcance de nuestro estudio, los autores se opusieron a la idea de Gilbert (1970) de que el nivel de desarrollo lingüístico no influye en el desarrollo de la estructura espectral puesto que para Clement y Wijnen (1994) está claro que la combinación del desarrollo fonológico, el anatómico y el auditivo condicionan la configuración espectral de las vocales. Estas posibles diferencias entre ambos trabajos pueden deberse a la diferente metodología que se siguió en cada uno de ellos ya que según los autores los sujetos de Gilbert (1970) no eran menores de 4;0 años, ni este se fijó en las estructuras espectrales de las vocales relativamente cerradas.

También la investigación de Palethorpe et al. (1996) mostró que los formantes de las vocales de los adultos son más altos que los de las vocales de los niños tras analizar diez vocales

del inglés australiano en el contexto de CVC producidas por niños de 0;4 años que fueron grabados dos veces por semana en un periodo de cuatro meses.

Retomamos en este punto el experimento de Kuhl y Meltzoff (1996) en el que se analizaron las vocales articuladas por bebés de entre 0;3 y 0;5 años tras escuchar las palabras *hop*, *heap* y *hoop* emitidas por un hablante femenino a través de una máquina, artículo que ya tratamos cuando hablamos del aprendizaje vocal de los niños (apartado 2.2.2 que trataba sobre la influencia de la lengua ambiente). La razón de que regresemos a dicho trabajo no es otra que la de examinar los formantes de las vocales emitidas por los bebés cuyo promedio en cada uno de los meses analizados aparece en la tabla siguiente:

**Tabla 2.3. Valores de los dos primeros formantes de las vocales de Kuhl y Meltzoff (1996).**

Edad	Formantes	i	a	u
0;3	F1	781	933	731
	F2	3121	2606	2198
0;4	F1	739	1044	757
	F2	2887	2499	2156
0;5	F1	778	945	675
	F2	2947	2393	2335

Si echamos un vistazo a los datos presentados en la tabla anterior, no parece que los formantes de cada vocal sufran descensos significativos a lo largo del periodo de estudio. Por otro lado, podemos observar una vez más que los formantes de los bebés son más altos que los de cualquier hablante adulto debido a las características de su tracto vocal. En cuanto a este último punto y en relación a los formantes de la vocal [u], Kuhl y Meltzoff (1996) destacan el hecho de que a pesar de que los formantes que presentó esta vocal en los bebés fueron muy diferentes a los de la misma vocal emitida por cualquier adulto o niño mayor hablantes de inglés, esta vocal fue escuchada como [u] por todos ellos debido a que la identificación de una vocal no está determinada por el lugar que esta ocupa en la carta de formantes sino por la información relacional que tiene en cuenta la  $f_0$  y las relaciones entre formantes tal y como han sugerido algunos autores como Syrdal y Gopal (1986), Miller (1989) y Nearey (1989).

En el estudio aparecía también la evolución del espacio acústico que como ya comentamos en el apartado 2.2.2 revelaba que las áreas ocupadas por las vocales [a, i, u] se iban separando progresivamente entre los 0;3 y los 0;5 años de vida debido a cuestiones anatómicas y de aprendizaje vocal.



Un año más tarde, Gilbert et al. (1997) analizaron las vocales de cuatro niños con desarrollo postnatal normal que adquirían el inglés americano y que fueron grabados durante sesiones de una hora de duración realizadas a los 1;3, 1;6, 1;9, 2;0 y 3;0 años de edad. Se extrajeron los sonidos vocálicos que aparecían aislados o acompañados de consonantes, se hicieron espectrogramas de banda ancha (450 Hz) y se transcribieron los segmentos extraídos usando la transcripción ancha del Alfabeto Fonético Internacional. Para simplificar el número de clasificaciones fonéticas, cada vocal identificada se incluyó en una categoría amplia como la de la altura de la lengua en el eje vertical, es decir, vocal alta, media o baja, y el avance de la lengua en la dimensión horizontal, esto es, vocal anterior, central o posterior. Las vocales altas fueron percibidas como cualquiera de las siguientes [i, ɪ, u, ʊ]; las vocales medias como [ɛ, ʌ, ə, ɔ, o, ə̃]; y las vocales bajas como [ɑ, æ]. Por otro lado, las vocales anteriores fueron percibidas como una de las siguientes [i, ɪ, ɛ, æ]; las vocales centrales como [ʌ, ə, ə̃] y las vocales posteriores como [u, ʊ, o, ɔ, a]. Se seleccionó también el primer segmento de los diptongos y las transcripciones se completaron con la visualización acústica de la estructura formántica para ayudar a la identificación.

En cuanto a la media de los valores formánticos obtenida en cada una de las edades, los resultados fueron los siguientes:

**Tabla 2.4. Valores de los dos primeros formantes de las vocales de Gilbert et al. (1997).**

Edad	F1	F2
1;3	1078	2558
1;6	1087	2501
1;9	1061	2579
2;0	871	2304
3;0	645	1938

Para determinar si los formantes descendieron de manera significativa a lo largo del periodo de estudio, se realizaron análisis de varianza (ANOVA) que manifestaron que no hubo descensos significativos ni en la media del F1 ni en la media del F2 entre los 1;3 y los 1;9 años, pero sí entre los 1;3 y los 3;0 años de edad. Estos análisis se completaron con otros más concretos en los que se observó la relación entre el descenso del F1 y del F2 y el aumento de la edad en las vocales que atendían a la clasificación según la dimensión anteroposterior de la cavidad bucal. Los resultados mostraron que el F1 y el F2 no descendieron con la edad en las vocales anteriores, centrales y posteriores entre los 1;3 y los 1;9 años de vida pero que sí lo hicieron entre estos periodos y los 2;0 y 3;0 años. Además, el promedio del F1 y del F2 fue

significativamente más bajo a los 2;0 y a los 3;0 años de edad, comparado con los valores entre los 1;3 y los 1;9 años, y las correlaciones más altas estuvieron en las vocales posteriores (mayor dependencia en estas vocales entre el F1 y el F2, y el crecimiento anatómico). En cuanto al eje vertical, los análisis realizados para evaluar la relación entre el descenso del F1 y el F2 y el aumento de la edad en las vocales altas, medias y bajas indicaron como en el caso de la dimensión de anterioridad y posterioridad que el F1 y el F2 de estas vocales descendió significativamente a los 2;0 y a los 3;0 años en comparación con los valores obtenidos entre los 1;3 y los 1;9 años. En esta ocasión, la correlación más fuerte entre el descenso del F1 y el F2 y el desarrollo anatómico se dio en las vocales altas.

Entre las razones de la ausencia de descenso del F1 y del F2 antes de los 2;0 años de vida puede señalarse según los autores la de la nasalización de las vocales de los niños más pequeños que trae consigo la aparición de una resonancia adicional de baja frecuencia que aparece por debajo del F1 debido a la asociación de la cavidad nasal con la oral y con la faríngea (Fant [1960] y Kent y Read [1992]). Ello pudo oscurecer los valores reales del F1 y del F2 de las vocales analizadas ya que el investigador pudo confundir dicha resonancia nasal con el F1, confusión que enmascararía cualquier descenso de dicho formante relacionado con la edad. Por otro lado, y tal y como opinan Kent y Read (1992), el hecho de que los formantes no descendan antes de los 2;0 de edad puede deberse a un crecimiento irregular o no proporcionado del tracto vocal. El tracto vocal del bebé, que aquí se describe como con forma de llamarada, muestra en el extremo más dilatado la influencia de las cavidades nasales y orales, y en el extremo más estrecho la influencia de la cavidad faríngea (Sasaki et al. [1960]). Según el niño se desarrolla, se produce una reducción de la influencia de la cavidad nasal, un descendimiento de la laringe y un aumento de la longitud de la cavidad faríngea (Sasaki et al. [1960]). Por ello los autores proponen que (1997:265):

“[...] It is the trading relationship between the decreasing flare (i. e., decrease in nasal cavity influence) and increasing vocal tract length and pharyngeal space that may contribute to the acoustic effect of relatively stable F1 and F2 values during the 15- to 21-month age period. These dynamic changes in vocal tract anatomy may override the effects of general increases in vocal tract length. However, by 36 months of age, the increasing length of the vocal tract is dominant and contributes to a lowering of F1 and F2”.

Por tanto, es la distinta naturaleza de los cambios que ocurren en el tracto vocal a lo largo de este periodo y no un simple proceso de alargamiento de este, lo que impide que un descenso de los valores formánticos ocurra. Crelin (1987) añade que el tracto vocal del niño

continúa reconfigurándose hasta los cinco años de edad, tiempo en el cual su estructura es similar a la del tracto vocal de un adulto.

Si examinamos los datos con mayor detenimiento, encontramos cierta variabilidad inter individual. De esta forma, a los 1;3 años la media del F1 del niño dos es de 910 Hz mientras que la del niño cuatro es de 1321 Hz, esto es, cuatrocientos once hertzios mayor que la del sujeto dos. En cuanto a la media del F2 en el mismo periodo, la del sujeto uno es de 2733 Hz frente a la del sujeto cuatro que es de 2201 Hz, es decir, quinientos treinta y dos hertzios menor que la del sujeto uno.

Una vez más, los valores de los formantes de los niños son mayores que los de los adultos y estos además se corresponden según los autores con los datos formánticos ofrecidos por otros estudiosos. De este modo, la media de los valores del estudio de Buhr (1980) y la media de los formantes del trabajo Kent y Murray (1982) desde los 0;3 a los 1;4 años de edad es de 1000 Hz para el F1 y de 3000 Hz para el F2, siendo la media del F1 de los niños de este estudio a los 1;3 años de 1078 Hz, y la del F2 de 2558 Hz (Gilbert et al. [1997] unen los periodos que comprenden ambos estudios). Por otro lado, los valores observados por Eguchi y Hirsh (1969) en los niños de 3;0 años de edad fueron de los 484 Hz a los 802 Hz para el F1 y de los 1485 Hz a los 3318 Hz para el F2, datos muy similares a los que presentaron los niños de 3;0 años de este trabajo y que fueron de 645 Hz para el F1 y de 1938 Hz para el F2. Estos resultados parecen llenar según los autores el vacío en la literatura entre el periodo de 1;0 a 3;0 años de edad.

En el mismo año Robb et al. (1997) llevaron a cabo el estudio de veinte niños con desarrollo normal que adquirirían la lengua inglesa entre los 0;4 y los 2;1 años de edad y cuyas emisiones fueron recogidas a los 0;4, 0;8, 1;3, 1;6, y 2;1 años en sesiones de grabación que duraban una hora. Los investigadores analizaron las *vowel-like* que aparecían dentro de una espiración acústica e hicieron análisis de *Linear Predictive Coding* (LPC) y espectrogramas de banda estrecha sobre la parte estacionaria del sonido con el objetivo de medir los formantes y el ancho de banda de estas frecuencias. Este criterio de análisis implica que las vocales que deben ser analizadas pueden ser en ocasiones vocales aisladas, como a la edad de 0;4 años, pero otras veces pueden ir acompañadas de consonantes como en la fase que precede a la palabra y en la de la palabra. Por otro lado, no se realizó una transcripción de los segmentos extraídos por dos razones: la primera razón reside en el hecho de que según los autores no pueden etiquetarse las vocales que aparecen en la fase de *preword* puesto que no se trata de palabras inteligibles sino de un sistema lingüístico inmaduro al que no pueden asignarse vocales concretas; la segunda, se basa en la consideración del IPA por parte de los autores como un sistema de transcripción

deficiente para el tratamiento de los sonidos de los bebés puesto que puede resultar insuficiente para capturar ciertas cualidades de las producciones vocálicas de los niños.

Los valores formánticos de las vocales de los cuatro niños incluidos en cada grupo de edad fueron los siguientes:

**Tabla 2.5. Valores de los dos primeros formantes de las vocales de Rob et al. (1997).**

Edad	F1	F2
0;4	768	2407
0;8	775	2494
1;3	815	2367
1;6	881	2417
2;1	821	2347

Como podemos observar, el F1 experimenta un aumento a los 1;6 años de edad que según los autores puede reflejar un punto de transición en el desarrollo vocal resultando de un cambio en la maduración anatómica o lingüística. Por otro lado, el promedio de los datos del F1 fue de los 768 Hz a los 881 Hz y el del F2 de los 2347 Hz a los 2494 Hz. Estos valores están a medio camino entre los datos ofrecidos por Kent y Murray (1982) en el periodo de 0;3 a 0;9 años y los obtenidos por Eguchi y Hirsh (1969) en los niños de 3;0 años.

En cuanto a si los valores del F1 y del F2 cambian con el paso del tiempo, parece que estos no experimentan cambios significativos entre los 0;4 y los 2;1 años de edad. La ausencia de un descenso significativo de los formantes con el paso del tiempo puede deberse, tal y como explicaron los autores en Gilbert et al. (1997), a la presencia de la nasalización en las vocalizaciones de los bebés y al crecimiento irregular del tracto vocal en el que además de producirse un alargamiento del mismo los cambios dinámicos favorecen que el F1 y el F2 no cambien de manera significativa antes de los 2;0 años de vida (a los 3;0 años de edad el aumento del tracto vocal es significativo y contribuye a la bajada del F1 y F2).

Por el contrario, los resultados sí que indicaron un descenso del ancho de banda del F1 y del F2 a lo largo del tiempo. Según los autores, este hecho puede atribuirse por una parte a la influencia de la nasalización y por otra a los cambios en la composición de los tejidos que forman el tracto vocal. En cuanto a esto último, Sinclair (1985) sugiere que en la composición de los tejidos de los niños se encuentra la retención de agua o viscosidad que decrece con la edad. Dicha viscosidad aumenta la absorción del sonido según Fletcher (1992) de forma que al

descender la cantidad de agua de los tejidos se produce un descenso de la proporción de absorción de sonido que conlleva un estrechamiento del ancho de banda.

El siguiente trabajo se centra en la adquisición de las vocales [i, u] por un grupo de niños con desarrollo normal y por otro grupo de niños con riesgo de tartamudeo. Se trata del artículo publicado por Fosnot (1997) en el que niños que adquirían el inglés americano fueron grabados a los 0;6, 0;9, 1;0, 1;3, 1;6, 1;9, 2;0, 2;3, 2;6, 2;9 y 3;0 años de edad en interacciones espontáneas con sus padres. Los análisis de LPC y de FFT mostraron que en los niños con riesgo de tartamudeo no había una clara distinción entre las vocales anteriores y posteriores, y que sus valores no eran los típicos de los de los niños que se desarrollaron con normalidad. Estos niños mostraron un F1 más alto tanto para [i] como para [u], sugiriendo según el autor que la altura de la lengua o posiblemente de la mandíbula es más baja para las vocales altas, y exhibieron valores formánticos del F2 también más altos para ambas vocales, reflejando una posición de la lengua más adelantada para las vocales anteriores y posteriores en los niños con riesgo de tartamudeo.

En la lengua italiana, Zmarich (1997) emprendió el estudio del balbuceo y de la primera palabra de un sujeto que aprendía italiano entre los 0;10 y los 1;4 años de edad. El niño fue grabado cada dos semanas en situaciones de juego entre este, la madre y el experimentador. Se extrajeron todas las vocales precedidas o seguidas de consonante y se realizaron análisis de FFT y de LPC para determinar los dos primeros formantes. La transcripción de los segmentos fue hecha por un transcriptor con experiencia.

En cuanto a la mayor o menor aparición de determinadas vocales, se observó un predominio de vocales anteriores en los primeros meses apareciendo la [a] al año de edad y aumentando esta su presencia en los meses posteriores convirtiéndose en la vocal más articulada a los 1;2 y 1;4 años y reflejando quizás la alta presencia de esta vocal en la lengua italiana. Como en estudios anteriores, las vocales bajas y medias caracterizaron las primeras producciones mientras que la serie posterior y las vocales redondeadas no aparecieron en los primeros meses de análisis. Zmarich (1997) acudió a los trabajos de Kent y sus colaboradores para explicar que el tipo de vocales producidas depende del progreso en la articulación de la lengua cuyo posicionamiento en los primeros meses es de reposo favoreciendo el movimiento en el eje vertical y en la dimensión anterior. El aumento de vocales posteriores en los últimos meses de análisis evidencia la habilidad del sujeto para explotar el componente dorsal de la lengua.

A los 1;2 años se diferenciaron de forma notable las vocales cardinales [a, i, u] y el triángulo vocálico se alargó hacia los lados. A la edad de 1;4 el triángulo evolucionó hacia el

modelo adulto distinguiéndose la posible área de todas las vocales excepto la de la [o] que solo fue registrada una sola vez. Un año más tarde Zmarich junto con Lanni (1998) publicaron un artículo donde estudiaron la combinación de vocales y consonantes en el mismo sujeto y en las mismas edades.

Mayores fueron los sujetos del estudio de Smith y Kenney (1998) cuyo objetivo era determinar los formantes de la vocal [ɪ] en siete sujetos femeninos que adquirirían el inglés americano y que fueron grabados en intervalos de año y medio a las siguientes edades: en el tiempo 1, las niñas tenían entre 8;0 y 9;0 años de edad (media de 8;7 años), en el tiempo 2 tenían entre 9;0 y 10;0 años (media 10;1 años), y en el tiempo 3 ellas tenían entre 11;0 y 12;0 años (media de 11;5 años). Cada uno de los sujetos debía repetir veinte veces la palabra *sissy* de manera sucesiva. Con ayuda de los análisis de FFT y LPC se obtuvieron los siguientes valores formánticos como media de las producciones de todos los sujetos en cada una de las edades: en el tiempo 1, el F1 fue de 535 Hz, el F2 de 2301 Hz y el F3 de 3692 Hz; en el tiempo 2, los valores fueron de 540 Hz, 2130 Hz y 3408 Hz respectivamente; y en el tiempo 3, el F1 fue de 487 Hz, el F2 de 2101 Hz y el F3 de 3198 Hz. Los resultados del test de Friedman indicaron que los cambios a lo largo del tiempo fueron significativos para el F2 y para el F3 mientras que el descenso del F1 no alcanzó significatividad.

Por otro lado, en los tiempos 1 y 2 ninguno de los valores del primer formante estuvo dentro del rango del de los adultos mientras que cinco de los valores del F1 en el tiempo 3 sí alcanzaron los valores adultos. A pesar de que los datos de esta investigación son muy limitados al centrarse únicamente en una vocal y al ser los intervalos de grabación tan espaciados, hemos optado no obstante por incluirlos ya que según los propios autores diferentes estudios transversales han mostrado que varios aspectos de la producción del habla están aún desarrollándose en estas edades (Tingley y Allen [1975], Kent y Forner [1979] y Smith y Kenney [1998]).

A partir de la base de datos de Miller et al. (1996), que recoge el habla de cuatrocientos treinta y seis sujetos que adquirirían inglés americano entre las edades de 5;0 y 18;0 años y el habla de cincuenta y seis adultos hablantes de la misma lengua, Lee et al. (1999) decidieron analizar acústicamente las vocales y los diptongos insertos en diez palabras. Las palabras utilizadas para las diez vocales monoptongadas fueron *bead* (/ɪY/), *bit* (/IH/), *bet* (/EH/), *bat* (/AE/), *pot* (/AA/), *ball* (/AO/), *but* (/AH/), *put* (/UH/), *boot* (/UW/) y *bird* (/ER/). El análisis de los tres primeros formantes así como el de la  $f_0$  fue dado para todas las vocales analizadas tanto en los sujetos de sexo masculino como en los de sexo femenino a lo largo de las diferentes edades. Los resultados mostraron que la variabilidad espectral de las producciones

descendió con la edad de manera progresiva y convergió con los valores adultos alrededor de los 14;0 años para ambos sujetos (hombres y mujeres). Las comparaciones indicaron que dicha variabilidad se redujo de manera significativa entre los 5;0 y los 9;0 años y entre los 10;0 y los 14;0 años de edad. Estos resultados sugieren según los autores que los niños menores de 10;0 años no han establecido metas articulatorias óptimas o estables.

También el trabajo de Huber et al. (1999) examinó los formantes de las vocales en niños mayores aunque el periodo de estudio en este caso alcanzó los treinta años de edad. Diez sujetos mujeres y diez sujetos varones hablantes de inglés americano fueron grabados a los 4;0, 6;0, 8;0, 10;0, 12;0, 14;0, 16;0, 18;0 años y entre los 20 y los 30 años de edad, produciendo una fonación sostenida durante dos o tres segundos de la vocal /a/ en los siguientes tres niveles de intensidad vocal medidos en decibelios: fonación confortable, fonación alta (10 dB por encima de la confortable) y fonación baja (5 dB por debajo de la confortable).

La media del primer y del segundo formante de la vocal producida en la condición de fonación confortable de los sujetos de sexo femenino fue la siguiente:

**Tabla 2.6. Valores de los dos primeros formantes de los sujetos femeninos del trabajo de Huber et al. (1999).**

Edad	F1	F2
4;0	1139	1817
6;0	1030	1817
8;0	945	1738
10;0	935	1607
12;0	921	1522
14;0	865	1500
16;0	882	1560
18;0	858	1381

En cuanto a los sujetos varones, las medias del F1 y del F2 de la vocal producida en la condición de fonación confortable fueron las siguientes:

**Tabla 2.7. Valores de los dos primeros formantes de los sujetos masculinos del trabajo de Hubert et al. (1999).**

Edad	F1	F2
4;0	981	1676
6;0	992	1681

8;0	944	1611
10;0	938	1569
12;0	856	1525
14;0	739	1292
16;0	670	1273
18;0	623	1209

A partir de estos datos se infirió que para los varones el crecimiento del tracto vocal se aceleró entre los 12;0 y los 14;0 años y que a este le siguió un aumento más pequeño a la edad de 16;0 o 18;0 años. Para las mujeres, la maduración estuvo completada en su mayoría a la edad de 12;0 años aunque el F2 y el F3 descendieron de manera leve hasta la edad de 18;0 años. Puede concluirse por tanto que el crecimiento del tracto vocal en las mujeres parece estabilizarse a una edad más temprana que en los varones. Por otro lado, los formantes de las niñas se hicieron similares a los de las mujeres adultas a la edad de 12;0, lo que sugiere que estas poseen a esta edad a un tracto vocal más maduro, mientras que para los varones los formantes descendieron de manera significativa a la edad de 14;0 años y continuaron descendiendo a lo largo del periodo de estudio apuntando a una estabilización del crecimiento del tracto vocal más tardía que en las mujeres.

En cuanto a esta distinción sexual, Huber et al. (1999) citaron en su artículo el trabajo de Lee et al. (1999) según el cual las diferencias de sexo no son visibles hasta alrededor de la edad de 15;0 años para el F2 y el F3, y también aludieron al artículo el de Bennet (1981) quien tras estudiar el F1, el F2, el F3 y el F4 en hombres y mujeres entre los 7;0 y 8;0 años encontró que los hablantes de sexo masculino tendían a producir formantes más bajos que los de sexo femenino y que estas diferencias se debían principalmente a distinciones en la longitud faríngea. Otros estudios citados por Huber et al. (1999) son el trabajo de Boersma et al. (1979), según el cual la velocidad de crecimiento de hombres y mujeres es diferente, y el estudio de Walker (1994) en el que se sugirió que las diferentes proporciones de crecimiento del hueso en varones y mujeres resultan en diferentes formas faciales y ello puede quizás afectar a los formantes. Para otros autores como Merow y Boradbent (1990) las diferencias de sexo no se producen hasta la pubertad.

Uno de los últimos estudios de la década de los noventa que llama la atención por las características de los sujetos estudiados, es el trabajo de Feld (1999) en el que se analizaron los formantes de la vocal [a] producida por niños prematuros de menos de 1500 gramos nacidos en Buenos Aires (Argentina). Los resultados indicaron que la diferencia entre el valor más alto del



F2 y el más bajo del F1 para cada emisión fue siempre mayor en los niños prematuros que en los niños nacidos al término del embarazo. Además, los formantes de la vocal [a] de los niños prematuros presentaron una mayor dispersión en frecuencia que la de los niños nacidos al final del periodo de gestación. Estas diferencias pueden entenderse a partir de las características anatómicas que presentan los tractos vocales de los niños que han nacido después de haberse completado el periodo de gestación y los de los prematuros.

Resulta interesante el trabajo de Lyakso et al. (2005), que en principio se aparta de nuestro objeto de estudio ya que pretende averiguar si los adultos son capaces de reconocer las palabras con significado y las oraciones de niños que están adquiriendo la lengua rusa a los 3;0 años de edad (grabados en intervalos de tres meses), por ciertas aportaciones acústicas. Entre estas se encuentra el hecho de que los triángulos formánticos de las vocales [a, u, i] no fueron significativamente diferentes para las vocales en posiciones acentuadas y no acentuadas evidenciado según los autores la falta de formulación de la propiedad acentuada y no acentuada en el tercer año de vida. Por otro lado, Lyakso et al. (2005) encontraron que los formantes de la [o] alcanzaron valores adultos al final del tercer año de vida, algo hasta ahora no contemplado en ninguno de los estudios mencionados. Quizás esto indique que los niños han alcanzado a esta edad el desarrollo neuronal y motor característico de los adultos al menos para la producción de esta vocal.

Siguiendo con la lengua rusa, el trabajo de Van der Stelt et al. (2006) pretendió determinar el espacio vocálico de cuatro niños que adquirirían la lengua rusa a la edad de 0;6, 1;0, 1;6 y 2;0 años. Los sujetos fueron grabados cada tres meses y no está claro si las vocales en las que se analizó el F1 y el F2 aparecían de manera aislada, acompañadas de consonantes o en ambos contextos ya que lo único que se concretó en el artículo fue que se analizó el núcleo de la sílaba. En la lengua rusa hay seis fonemas vocálicos, /a/, /o/, /u/, /e/, /i/ e /y/, aunque algunos lingüistas consideran que son solo cinco pasando el fonema /y/ a ser un alófono de /i/. Los resultados mostraron los timbres vocálicos que predominaron en los niños en las diferentes edades y aunque las tendencias universales pudieron observarse (por ejemplo la alta frecuencia de la vocal central y anterior en los primeros meses de análisis), también encontramos preferencias individuales.

A los 0;6 años, las vocales centrales y anteriores fueron las más frecuentes en los cuatro niños, sobre todo las centrales, aunque el niño tres mostró un mayor predominio de vocales posteriores que el resto de los niños. A la edad de 1;0 año, las vocales centrales siguieron predominando y en concreto las vocales bajas y las medias lo hicieron en los niños tres y cinco. El niño tres presentó de nuevo una clara preferencia por las vocales posteriores, en este caso

medias, y el niño dos por el contrario articuló más vocales anteriores altas y medias. A los 1;6 años de edad, las vocales centrales continuaron prevaleciendo en todos los niños siendo bajas y medias en los niños dos y tres, y medias y altas en el niño cuatro. A esta edad, el número de vocales anteriores se redujo en el sujeto dos y aumentó en el sujeto cinco. El niño tres presentó más vocales anteriores que posteriores y el niño cuatro exhibió el patrón contrario. A los 2;0 años de edad, las vocales centrales volvieron a predominar sobre el resto de vocales especialmente en el niño cinco. El niño dos presentó prácticamente el mismo número de vocales anteriores y posteriores, en el niño tres observamos más vocales posteriores que anteriores y en el niño cuatro más anteriores que posteriores.

Según los autores la [a] y la [e] fueron las primeras vocales en aparecer seguidas de [i], [o] e [y], siendo esta última la vocal más difícil de pronunciar para los que adquieren ruso y siendo quizás también la más difícil de articular para los niños.

Finalmente, al comparar los resultados de este estudio con los hallados en la lengua holandesa, Van der Stelt et al. (2006) observaron una similitud entre los dos grupos de niños en cuanto a la correcta producción de la vocal [i] que según los autores podía deberse, al menos en holandés, a la frecuente escucha de los niños de vocales palatalizadas debido a su fuerte presencia en el diminutivo holandés (e. j. ‘je’, ‘tje’, ‘kje’, y ‘pje’).

Otra publicación que investigó la adquisición del ruso por parte de niños pequeños es la de Lyakso et al. (2006) y en ella se comparó el habla de niños de 2;0 años hablantes de ruso con la de niños de la misma edad pero hablantes de húngaro y holandés. Las sesiones de grabación para el holandés y el ruso tuvieron lugar en la casa de los sujetos cuando los niños conversaban con sus madres, mientras que los datos de la lengua húngara fueron recogidos en el laboratorio. Para provocar las emisiones se utilizaron marionetas a las que se les asignó nombres con la siguiente estructura silábica CV:CV: (ej. /pi:pi:/). Los resultados mostraron que las vocales del holandés variaron más y se situaron más periféricamente en el espacio acústico que las vocales de los niños húngaros y rusos. Encontramos pues diferencias formánticas en las mismas vocales al ser estas emitidas por hablantes de distintas lenguas.

Un resultado similar fue hallado en el estudio de Rvachew et al. (2006) tras examinar las diferencias en las frecuencias del primer y segundo formante de vocales compartidas por dos lenguas. En dicho trabajo, ya comentado en el apartado de la influencia de la lengua ambiente en las vocalizaciones de los niños, se grabaron las muestras de habla de veintitrés niños expuestos al francés de Canadá y veintitrés expuestos al inglés del mismo país entre las edades de 0;10 y 1;6 años (todos los niños nacieron entre las 38 y las 42 semanas de gestación). Las sesiones de grabación se extendieron hasta que el niño produjo sesenta emisiones o hasta los

treinta minutos de duración. Análisis de FFT y de LPC se realizaron en las vocales aisladas y en aquellas que formaban parte de sílabas canónicas. Según Rvachew et al. (2006) hay que tener en cuenta que las producciones vocálicas del adulto de estas lenguas indicaron que las vocales del francés y del inglés de Canadá presentaron diferencias acústico-fonéticas incluso cuando existía un solapamiento fonológico entre ambas lenguas tal como mostraron estudios recientes como los de LaCharite y Paradis (1997), Martin (2002) y Escudero y Polka (2003). Quizás estas diferencias pueden manifestarse también en las características acústicas de las vocales producidas por los niños que están aprendiendo dichas lenguas.

Los resultados indicaron un descenso a lo largo del periodo de estudio en la media del F1 en los niños que adquirían francés pero no en los que adquirían inglés, y un descenso en la media del F2 en los niños que adquirían inglés pero no en los que adquirían francés.

Ambos grupos de niños mostraron según fueron desarrollándose una expansión del tamaño del espacio vocálico a lo largo de la dimensión difusa-grave, aunque la expansión en el extremo difuso fue mayor para el grupo de niños que aprendía francés y la expansión en el extremo grave fue mayor para el grupo de niños que aprendía inglés. Estos hallazgos podrían sugerir según los autores una dilatación en el tiempo de estudio de las áreas del espacio vocálico que tradicionalmente se han asociado con la retracción y el avance de la lengua en la articulación adulta. Por otro lado, la compresión del espacio vocálico en el extremo agudo, particularmente para el grupo francés, apunta a gestos de abertura de la mandíbula menos extremos con la edad.

Esta expansión del espacio vocálico coincidente según los autores con la de otros estudios como los de Buhr (1980), Kent y Murray (1982), Rvachew et al. (1996), Gilbert et al. (1997) y Robb et al. (1997) es predecible como consecuencia de los cambios que ocurren en la habilidad del niño de controlar la posición de la punta, el cuerpo y el dorso de la lengua independientemente de la altura de la mandíbula (esto último es indicado a partir de los resultados obtenidos en el segundo experimento de este trabajo en el que los autores simularon el tracto vocal de un niño entre las edades de 0;6 a 1;2 años).

En cuanto a la dispersión de los valores formánticos, ningún grupo mostró cambios relacionados con la edad en la desviación estándar del F1. Sin embargo, los niños de habla inglesa mostraron un incremento significativo en la dispersión del segundo formante según aumentó la edad, mientras que los que adquirían la lengua francesa no mostraron ningún cambio pertinente.

Finalmente, revisamos el trabajo de Rvachew et al. (1996) al que aluden los propios autores y en el que se describieron las vocales producidas por nueve niños que adquirirían la lengua inglesa entre los 0;6 y los 1;6 años de edad. A partir de este estudio longitudinal se observó que la media y la desviación estándar del F1 fueron estables a lo largo de este periodo frente a los valores del F2 que aumentaron con el paso del tiempo. Por otro lado, la media de los valores del primer formante fue similar a la dada por Kent y Murray (1982) mientras que la media del segundo formante fue considerablemente más baja (unos 2400 Hz menor).

La adquisición de las vocales de los niños pertenecientes a la comunidad lingüística del mandarín ha sido muy poco estudiada a lo largo del tiempo. Uno de los trabajos que intentó mitigar este vacío fue el de Chen (2006) quien investigó las producciones vocálicas de niños que adquirirían dicha lengua entre los 0 y los 2;0 años de edad. Se realizaron análisis acústicos en los que se midió el F1 y el F2 de las vocales aunque no se explicitó si se analizó la vocal aislada o en contexto consonántico. Los hallazgos de este estudio longitudinal fueron los siguientes: no hubo cambio significativo en el rango de valores del F1 y del F2 durante estos dos primeros años de vida; las vocales más comunes en los bebés de 0 a 0;6 años fueron las vocales centrales medias mientras que entre los 0;7 y los 2;0 años la vocal [a] fue la que apareció más frecuentemente; otras vocales se desarrollaron a partir de los 1;6 años pero las nuevas vocales como [i, u] fueron difíciles de identificar debido a la baja amplitud y corta duración que las caracteriza; las categorías vocálicas que más aparecieron mostraron un patrón de los formantes más estable en el segundo año de vida; y finalmente, las diferencias individuales en el primer año de vida se hicieron menos obvias en el segundo año.

Un importante trabajo longitudinal por el amplio periodo de tiempo que comprendió el estudio aunque algo particular en cuanto a la metodología empleada, es el que llevaron a cabo Vorperian y Kent en el 2007. En él, los autores recopilaron un total de quince trabajos en los que se investigó la adquisición de las vocales por parte del niño (publicados todos ellos en las últimas cinco décadas) y analizaron el F1 y el F2 para determinar su evolución y la del área vocálica a lo largo del tiempo. Los estudios seleccionados fueron: Peterson y Barney (1952), Eguchi y Hirsh (1969), Pentz y Gilbert (1983), Hodge (1989), Childers y Wu (1991), Kent y Murray (1992), Zahorian y Jagharghi (1993), Busby y Plant (1995), Hillenbrand et al. (1995), Yang (1996), Kuhl y Meltzoff (1996), Hagiwara (1997), Lee et al. (1999), Assman y Katz (2000) y Perry et al. (2001).

Las peculiaridades del mismo son evidentes dado que cada uno de estos estudios siguió una metodología diferente con la resultante heterogeneidad de condiciones empleadas y sujetos analizados. Como parte de esta mescolanza encontramos por ejemplo que las vocales

seleccionadas en cada una de las publicaciones provienen de diferentes dialectos de la lengua inglesa (del inglés canadiense como en el trabajo de Hodge [1989], del inglés hablado en Australia como en el de Busby y Plant [1995], etc.), que en cada trabajo varió el número de sujetos estudiados así como el sexo de los mismos, que cada publicación seleccionó un tipo de vocales para ser estudiadas aunque en la mayoría se analizaron las tres vocales de los extremos [a, u, i] y en muchos se incluyó la [æ], o que las vocales analizadas se extrajeron de diferentes contextos (de vocales aisladas como en el trabajo de Kuhl y Meltzoff [1996], de los contextos CV, hV, de palabras completas o de combinaciones de dos palabras como en el de Hodge [1989], o se tomaron directamente de distintas oraciones como en el trabajo de Eguchi y Hirsh [1969]).

No obstante, todos los estudios debían responder a una serie de características que aportaban algo de unidad dentro de esta amalgama y que eran las siguientes: los estudios debían ofrecer datos cuantitativos de hablantes que estuvieran en un periodo de desarrollo, esto es, de niños, o que por otro lado fueran maduros, esto es, adultos; los sujetos debían ser hablantes de inglés; las vocales estudiadas debían ser al menos las tres de los extremos prefiriéndose que fueran las cuatro (es decir, que incluyeran la [æ]); finalmente, se prefirió los estudios de grupo a los estudios que analizaban las emisiones de un solo sujeto. A partir de aquí Vorperian y Kent (2007) tomaron el F1 y el F2 resultante de los análisis de todos estos estudios e intentaron establecer el área del cuadrilátero vocálico según las diferentes edades (desde los 0;7 a los 11;0 años y desde los 4;0 años a la edad adulta), y según el sexo (áreas vocálicas para hombres y para mujeres).

Los resultados apoyaron la predicción de la teoría acústica de que los formantes decrecen según crece el tracto vocal con la edad. Dicha teoría se cumplió tanto en los valores de los formantes de los niños, cuyas cartas de formantes mostraron claramente un descenso del área vocálica (y por tanto de dichos formantes) entre los 0;7 y los 11;0 años de vida, como en los datos ofrecidos para los hablantes adultos de ambos sexos tal y como exhibieron las áreas de los cuadriláteros trazadas para estos sujetos entre los 4;0 años y la edad adulta. No obstante y a lo largo del periodo estudiado, los autores encontraron cambios no lineales en los formantes puesto que se produjeron saltos en los valores formánticos en las edades o momentos correspondientes con los *spurt* o periodos de aceleración del crecimiento anatómico ya fuera del tracto vocal en su conjunto o de alguna de sus partes. Estos periodos de incremento del crecimiento provocaron cambios bruscos en los formantes.

Sin embargo, los datos del F1 y del F2 con los que contaban Vorperian y Kent (2007) pasaban de los 0;7 años al 1;0 año y de ahí a los 3;0 años sin que se hubieran realizado análisis

de los valores formánticos del segundo año de vida. Esto quizás llevó a nuestros autores a hablar de los trabajos de Robb et al. (1997) y de Gilbert et al. (1997) como de una excepción a la teoría acústica estándar. Recordemos que en ellos los formantes de las vocales de los niños no descendían de manera significativa durante los dos primeros años de vida a pesar de que según las imágenes de resonancia magnética de los estudios de Vorperian (1999, 2000 y 2005) el tracto vocal del bebé crecía durante este periodo. Para explicar este hecho tal vez sea necesario examinar según Vorperian y Kent (2007) algo más que el aumento de longitud del tracto vocal a esta edad, y observar los cambios en la anchura y la longitud de la cavidad faríngea y su relación con los formantes o el ancho de banda de estos (según los autores la reducción del ancho de banda del estudio de Robb et al. [1997] podría ser el resultado de la disminución de la nasalización, del cambio en determinadas propiedades biomecánicas de los tejidos del tracto vocal o del cambio de volumen en la región faríngea y velo-faríngea).

En cuanto a la variabilidad de los promedios del F1 del F2 y del F3 que ofrecen cada uno de los quince estudios para las distintas vocales a lo largo del tiempo, esta es alta debido a las múltiples fuentes de datos usadas para construir los gráficos. No obstante, en general hay una mínima variabilidad en el F1 para las vocales altas pero la variabilidad en el F2 aumentó particularmente para la vocal posterior alta [u] a lo largo del periodo estudiado. Este aumento puede indicar según los autores que dicha variabilidad radica en diferentes factores tales como la influencia del dialecto, la variabilidad articulatoria y la variabilidad debida al crecimiento no uniforme del tracto vocal, particularmente en la región posterior, que ha sido típicamente la responsable de los valores del segundo formante (los autores citan en este punto a Fant [1975] para recordar que el crecimiento de la región faríngea versus el de la región oral, no es uniforme). Esta variabilidad formántica se redujo con el paso tiempo aunque el F1 alcanzó la estabilidad antes que el F2.

Vorperian y Kent (2007) también observaron una transición en las vocales de los niños que pasaban de ser casi exclusivamente nasales a presentar un alto grado de resonancia oral. A pesar de la escasez de estudios entre el nacimiento y los 3;0 años de edad parece que el cierre velo-faríngeo para las emisiones del habla está todavía desarrollándose a los 0;6 años de edad según Thom et al. (2005) y puede seguir cerrándose hasta los 3;0 años según Thompson y Hixon (1979) y Leeper et al. (1998).

En cuanto a las diferencias de sexo, estas empezaron a emerger en el espacio acústico a los 4;0 años de vida especialmente en las vocales bajas cuyos valores de F1 fueron entre 150 y 200 Hz más bajos para los varones que para las mujeres. Esta diferencia se hizo más pronunciada con la edad de forma que el solapamiento formántico se fue reduciendo con el

tiempo y desapareció a la edad de 16;0 años. Este dimorfismo sexual puede deberse a diferencias en el tracto vocal en cuanto a forma y longitud, diferencias que parecen aflorar a los 4;0 años de edad y que se hacen más notables a la edad de 7;0 u 8;0 años, tiempo en el que los niños presentan formantes más bajos que las niñas en todas las vocales (Bennet [1981] y Busby y Plant [1995]). De acuerdo con Crelin (1973), la diferenciación sexual en el tamaño de la laringe comienza a notarse a la edad de 3;0 años. Por el contrario, Eckel et al. (1999) subrayaron la falta de diferencias en el tamaño de la laringe en la temprana infancia. En lo que respecta a la longitud de las cuerdas vocales, según Kazarian et al. (1978) las diferencias de sexo aparecen entre los 6;0 y los 7;0 años. Sin embargo, estas diferencias anatómicas no parecen reflejarse según los autores en la  $f_0$  que no es manifiestamente diferente en hombres y mujeres hasta la pubertad, tiempo en el que el tamaño de la laringe, en concreto el de la dimensión antero-posterior del cartílago tiroideos, aumenta el triple en los varones junto con aumentos en la longitud de las cuerdas vocales y diferencias en su composición.

Por otro lado, las diferencias en los formantes entre niños y niñas quizás también pueden reflejar distinciones articulatorias además de las anatómicas ya comentadas. Así pues, las diferencias en las frecuencias del F1 para las vocales bajas pueden significar que los chicos producen estas vocales con una mandíbula relativamente más abierta. No obstante, las diferencias para las vocales altas y posteriores parecen deberse más a diferencias en la longitud, anchura y volumen oro-faríngeo.

Como habrá podido comprobar el lector a lo largo de esta revisión de estudios, los distintos trabajos ofrecen diferentes conclusiones sobre el comienzo del dimorfismo sexual en las emisiones del habla. Según Vorperian y Kent (2007) estas diferencias pueden achacarse a que los trabajos se han centrado en diferentes aspectos y partes del sistema. Finalmente, estos autores nos advierten, citando a Kent y Read (1992), de que la generalización de estos resultados y su extensión a los de otras lenguas han de hacerse con cautela debido a que las vocales que se consideran que son fonéticamente equivalentes en dos lenguas pueden tener distintos patrones formánticos (tal y como se expuso en el trabajo de Rvachew et al. [2006] de la mano de autores como LaCharite y Paradis [1997], Martin [2002] y Escudero y Polka [2003]).

En la adquisición del japonés como lengua materna, es necesario que comentemos el ambicioso trabajo de Ishizuka et al. (2007) en el que siguiendo en parte un método longitudinal y en parte uno transversal, los autores estudiaron las vocales de un sujeto varón entre los 0;4 y los 0;6 años, entre los 0;8 y los 1;8 años, a los 1;10 años, a los 2;0 años, a los 2;1 años, a los 2;6 años, a los 2;10 años, a los 3;4 años, a los 3;8 años y a los 4;4 años de edad, y las de un sujeto

femenino entre los 0;4 y los 1;10 años, y a los 2;0, 2;1, 2;6, 2;11, 3;4, 3;9, 4;2, 4;7 y 5;0 años de edad. Las vocales se extrajeron de la base de datos de Amano et al. (2002) en la que aparecen las emisiones de niños japoneses. Hay al menos una sesión de grabación por mes y en ella se recogió el habla espontánea del niño. Por otro lado, dos tipos de análisis se realizaron: en el primero, se llevó a cabo un análisis acústico de los dos primeros formantes mediante la técnica de LPC al que se le unió una transcripción de las vocales que fueron extraídas de la cadena fónica y que podían aparecer por tanto en contextos consonánticos. Dicha transcripción fue realizada con la ayuda de dos investigadores que se valieron no solo de los índices auditivos sino también de los espectrales para etiquetar los sonidos vocálicos según los cinco segmentos vocálicos del japonés, a saber, /a, e, i, o, u/. En el segundo, tan solo se usaron los datos del habla que fueron etiquetados por los investigadores encargados de efectuar la transcripción como sonidos vocálicos excluyéndose así las consonantes. No obstante, los resultados de ambos tipos de análisis coincidieron tanto en el desarrollo del espacio vocálico como en las diferencias inter individuales.

Los resultados acústicos mostraron variaciones según la vocal y el sujeto. En lo que respecta al primer formante de las cinco vocales se observó lo siguiente: los valores del F1 para la [a] en ambos niños decrecieron hasta la edad de entre 1;8 y 2;6 años, permaneciendo sin cambios entre los 2;1 y los 3;4 y decreciendo de nuevo después de los 3;4 años. Para la vocal [e] el F1 descendió en ambos sujetos hasta los 2;6 años, edad a partir de la cual los valores del F1 del sujeto de sexo femenino continuaron decreciendo mientras que los del sujeto varón permanecieron casi sin cambiar. El primer formante para la vocal [i] descendió en ambos sujetos hasta los 1;8 años, edad a partir de la cual los valores del sujeto varón permanecieron sin cambios mientras que los de la mujer continuaron decreciendo. Para la vocal [o], los valores del F1 se mantuvieron sin cambios durante todo el periodo para el sujeto varón mientras que en la niña los valores del F1 fueron constantes hasta los 1;8 años y después descendieron. Finalmente, el primer formante de la vocal [u] del niño decreció hasta la edad de 1;8 años y después permaneció sin cambiar frente a los valores del F1 de esta vocal en la niña que permanecieron sin cambios hasta la edad de los 1;8 años y después decrecieron.

En cuanto al F2, los valores de esta frecuencia en ambos niños para la vocal [a] decrecieron de forma rápida hasta alrededor de los 1;8 años, momento después del cual permanecieron casi sin cambiar. Para la vocal [e], el F2 del sujeto de sexo masculino permaneció sin cambiar hasta alrededor de los 1;8 años, descendió a partir de esta edad y hasta los 2;6 años y después de este momento apenas sufrió modificaciones. En el sujeto de sexo femenino, el F2 de la [e] apenas cambió entre los 0;4 y los 5;0 años de edad. El segundo formante de la vocal [i] en el niño permaneció sin cambiar durante todo el periodo estudio



mientras que en la niña aumentó con la edad. Para la vocal [o], los valores del F2 de ambos infantes decrecieron hasta alrededor de los 1;6 años y después permanecieron sin cambiar. El mismo patrón presentó la vocal [u] ya que el F2 de ambos niños disminuyó hasta los 1;6 años y después se mantuvo sin cambios.

Ishizuka et al. (2007) intentaron explicar esta evolución formántica mediante la relación anatómico-acústica. Según los autores, el valor del F1 depende de la longitud del tracto vocal y de la altura de la posición de la articulación de la lengua de forma que las vocales altas tienen un bajo F1 y las vocales bajas poseen uno más alto. Así, la diferenciación en los valores del F1 de las vocales hasta alrededor de los 1;8 años puede reflejar el alargamiento del tracto vocal y la elevación vertical de la lengua algo que puede llevar a la distinta articulación de las vocales altas, medias y bajas (y a la diferente evolución de sus formantes). Los autores creen que el aumento de la longitud del tracto vocal y la elevación de la lengua ocurrieron por el rápido crecimiento de la cavidad faríngea y el rápido descenso de la laringe y el hiedes producido entre los 1;3 y los 1;8 años de edad según Vorperian et al. (2005). En opinión de estos, el descenso en los valores del F1 en los últimos periodos de desarrollo, sobre todo en las vocales emitidas por el sujeto femenino, puede considerarse un reflejo del aumento gradual de la longitud del tracto vocal a partir de los 1;8 años. Por otro lado, los valores del F2 se corresponden con el avance horizontal de la lengua en la cavidad oral. El rápido crecimiento de la lengua en cuanto a longitud producido hasta los 1;4 años (Vorperian et al. [2005]) coincide según los autores con la diferencia entre los valores del F2 para las vocales anteriores [i, e], las vocales centrales [a] y [u] (la vocal [u] a diferencia de en español o en inglés es una vocal central) y la vocal posterior [o].

A partir de las cartas de formantes trazadas se observó una superposición de las vocales [a] e [i], [o] e [i], y [o] y [e] que desapareció alrededor de la edad de 1;6 años. Se percibió también una expansión del espacio vocálico entre los 0;4 y los 1;6 años de edad, expansión coincidente con la rápida maduración de los órganos articulatorios hasta aproximadamente los 1;6 años, tiempo a partir del cual se produjo una maduración más gradual (Vorperian et al. [2005]).

Diferencias inter-individuales fueron advertidas en la expansión del espacio vocálico después de los 2;0 años ya que a partir de esta edad el espacio vocálico de la niña continuó expandiéndose frente al del niño que se contrajo una vez pasados los 2;0 años de edad, para después expandirse de nuevo. Esto puede deberse según Ishizuka et al. (2007:2281) a la influencia de la lengua que rodea al niño en su articulación vocálica:

“[...] Although more work is needed in this area, it can be speculated that, as an infant becomes a proficient speaker of his/her native language, the articulation tends to be influenced by the ambient speech provided by caregivers. Such influences must be present even in earlier periods. However, the influences of anatomical developments are presumably predominant until 24 months of age, and they conceal the influence of ambient speech. Given the fact that the manner of vowel articulation in adults has wide inter-individual variability, it is plausible that each infant develops a particular speech style according to his/her individual environment.

Sin embargo, aunque hay diferencias entre los sujetos después de los 2;0 años de edad, los procesos de desarrollo que tienen lugar hasta esa edad son muy similares. Esto sugiere que los procesos anteriores a esta edad son causados por cambios de desarrollo comunes a los niños y a las niñas sin que el sexo pueda tener influencia (recuerdan los autores que este periodo, el que transcurre entre los 0;4 y los 1;8 años de edad, coincide como ya hemos comentado, y según Vorperian et al. [2005], con un rápido crecimiento de las estructuras del tracto vocal).

Una conclusión ya observada en otros estudios es el hecho de que las frecuencias de los formantes de los niños de este trabajo son más altas que los valores dados para los adultos debido a las características anatómicas del tracto vocal de los infantes.

En cuanto al timbre vocálico que predomina en cada edad, los resultados fueron los siguientes: a los 0;4 años la frecuencia de aparición de las distintas vocales para el sujeto de sexo masculino fue por este orden [i], [a], [o], [e] y [u]; para el sujeto de sexo femenino, la frecuencia fue de [i], [o], [a], [u] y [e]. Al año de edad el orden de predominio vocálico para el niño fue de [a], [i] [o], [e] y [u], y para la niña fue de [a], [o], [i], [e] y [u]. A los 2;0 años, las vocales más frecuentes en el niño fueron por este orden [a], [e], [u], [o], e [i] y en la niña [a], [u], [o], [e] e [i]. Un mes antes de cumplir los 3;0 años de edad (y penúltimo mes en el que los autores ofrecen datos de ambos niños) las vocales que dominaron el espacio vocálico en el niño fueron por este orden [a], [u], [e], [i] e [o], y en la niña fueron [a], [u], [o], [e] e [i]. A partir de estos datos pueden percibirse ciertas tendencias en ambos niños: entre el primer y el segundo año de vida la vocal alta anterior [i] pasó de ser una vocal bastante frecuente a ocupar los últimos puestos. En ese mismo periodo la frecuencia de la vocal alta posterior [u] cambió y de ser una vocal muy poco frecuente pasó a ocupar el tercer y el segundo puesto de aparición en el niño y en la niña respectivamente. A los 3;0 años la vocal [u] fue la más producida por ambos sujetos. Finalmente, y excepto a lo largo del primer año de vida, la vocal [a] fue la vocal más articulada por ambos niños durante todo el periodo de estudio.

Para determinar si existían diferencias entre las lenguas en cuanto a la frecuencia de aparición de alguna de las vocales de los extremos, Rvachew et al. (2008) llevaron a cabo un estudio longitudinal en el que compararon la adquisición de la lengua materna por niños hablantes de inglés canadiense y por niños que hablaban francés de Canadá. Para ello, los autores examinaron el habla de un total de cincuenta y un niños (veinticuatro niños adquirían inglés y veintisiete francés) entre los 0;8 y los 1;6 años de vida, expuestos a una sola de las lenguas y cuyos padres eran hablantes nativos bien de inglés, bien de francés. En las sesiones de grabación se les pidió a las madres que hablaran con sus hijos y que dejaran de hacerlo una vez que el niño comenzara a vocalizar para evitar la superposición del habla. Las vocales procedentes de estas muestras de habla espontáneas fueron seleccionadas si aparecían en la sílaba canónica (Oller [2000]) o si eran producidas de manera aislada como núcleos con resonancia completa y con una duración menor de 500 milisegundos. Se realizaron espectrogramas de FFT de banda estrecha y análisis de LPC para extraer el F1 y el F2. A continuación, cinco hablantes nativos de inglés de Canadá y cinco de francés canadiense con experiencia en la transcripción fonética realizaron un test de percepción en el que los sonidos fueron presentados de forma aleatoria sin que estos supieran la edad del niño o la comunidad lingüística de la que provenían.

Los resultados obtenidos variaron según la vocal del extremo examinado. En el extremo difuso, el ANOVA realizado sobre la vocal [i] reveló un efecto significativo de la edad pero no de la lengua del niño. Antes del 1;0 año, solo el 5% de las vocales producidas por los niños hablantes de francés y el 12% de las producidas por niños hablantes de inglés fueron percibidas como [i]. Sin embargo, con el aumento de la edad el porcentaje de vocales percibidas como [i] aumentó un 33% en siete de los oyentes en las emisiones de los niños de mayor edad hablantes de francés, y un 29% en los niños mayores hablantes de inglés.

En el extremo grave y para la vocal [u], el análisis mostró un efecto significativo tanto de la edad, puesto que el número de vocales percibidas como [u] aumentó a lo largo del periodo de estudio para los niños de ambas lenguas, como un efecto significativo de la lengua materna, puesto que una mayor producción de estas vocales fue percibida en los niños hablantes de inglés. Así, antes del 1;0 año solo el 12% de las vocales emitidas por los niños que hablaban francés fueron percibidas como [u], mientras que un 30% de las vocales articuladas por los niños hablantes de inglés fueron juzgadas como [u] por al menos siete de los diez oyentes. Estos porcentajes aumentaron con la edad en un 31% y un 42% para los niños mayores hablantes de francés e inglés respectivamente. Estas diferencias cuantitativas en cuanto al número de vocales percibidas como [u] fueron encontradas tanto por los oyentes nativos de francés como por los oyentes nativos de inglés. Por otro lado, según Rvachew et al. (2008) la mayor emisión de úes

por parte de los niños de habla inglesa indica que la producción del habla del niño está influenciada por el input auditivo y por los objetivos o metas auditivas a una edad muy temprana (influencia de la lengua ambiente).

En cuanto a la vocal [æ] del extremo agudo, el análisis de varianza reveló un efecto significativo de la edad pero no de la lengua del niño. Antes del 1;0 año el 37% de las vocales producidas por los niños hablantes de francés y el 50% de las producidas por los niños hablantes de inglés, fueron percibidas como [æ] por al menos siete oyentes. Sin embargo, según aumentaba la edad de los sujetos estos porcentajes decrecieron hasta convertirse a los 1;6 años en un 23% para los niños hablantes de francés y en un 32% para los sujetos hablantes de inglés. Como habrá podido observar el lector, al mismo tiempo que se producía un descenso en la frecuencia de aparición de esta vocal, las vocales de los extremos [i] e [u] aumentaban su presencia a lo largo del periodo de estudio. Estos cambios son según los autores consistentes con los cambios de desarrollo y parecen reflejar mejoras en el desarrollo del control motor del habla según Rvachew et al. (2006). Así, es lógico que [æ] sea una de las vocales más frecuentes en las primeras etapas debido a la poca precisión articulatoria que se necesita para su producción. Con la edad, el niño fue mejorando en el control del gesto de apertura de la mandíbula y esto favoreció el aumento de otras vocales de los extremos como [i] o [u].

No se encontró ningún efecto significativo de la edad ni de la lengua en la vocal del extremo compacto [ɑ].

Los datos de este estudio confirman que una expansión del espacio vocálico del niño tiene lugar según aumenta la edad de este, expansión por otra parte ya documentada en otros estudios (Buhr [1980], Kent y Murray [1982], Kuhl y Meltzoff [1996], Robb et al. [1997] y Rvachew et al. [2006]. Esta expansión está asociada según los autores al incremento de las vocales de los extremos [i] e [u] producidas por los niños de mayor edad, en comparación con los sujetos más jóvenes.

En la adquisición de la lengua húngara resulta interesante el trabajo de Zajdó et al. (2011) por el estudio que se hace en este del control articulatorio de los niños. En dicho estudio se analizaron las emisiones vocálicas de cinco niños de 2;0 años y de cinco de 4;0 años en palabras que presentaron la siguiente estructura: dos sílabas formadas por consonante más la vocal breve [i] y dos sílabas formadas por consonante más la vocal larga [i:] (ejemplo /pipi/ vs pi:pi:/). El objetivo era el de determinar si los niños en ese rango de edad diferenciaban los índices de duración en las vocales (estudios previos como el de Zajdó y Powell [2008] indicaron que los sujetos eran capaces de diferenciarlos). Las muestras de habla no fueron por tanto

totalmente espontáneas sino que fueron guiadas con ayuda de juguetes que respondían a la estructura ya mencionada (CVCV o CV(:)CV(:)).

Según indicaron los resultados, la vocal breve [i] estaba en una posición más centralizada que su par largo [i:] lo que apuntaba a una menor estabilidad de los patrones de producción en las vocales breves de manera opuesta a las largas. En contraste, la mayoría de las vocales breves fueron articuladas con la mandíbula situada en una posición baja. Otra conclusión de interés es que los niños de 4;0 años de edad tendían a producir más vocales periféricas que los niños de 2;0 años debido en parte a las siguientes razones: a una mejora en la habilidad de producir vocales con un posición de la mandíbula más alta, a un aumento del control de la posición de los labios, a una mejora en la coordinación de la posición de los labios con las posiciones más altas y más anteriores de la lengua, a un incremento en la habilidad de mantener las posiciones de los labios y de la lengua durante un mayor tiempo, y a una menor variabilidad de los movimientos articulatorios durante la producción de una vocal.

Para acabar con esta revisión bibliográfica comentaremos el artículo de Chi Lin que trata sobre la adquisición del taiwanés por un niño que fue estudiado en dos etapas diferentes, y que hemos dejado para el final debido a la imposibilidad de identificar la fecha en que dicho artículo se publicó. Los datos que debían ser estudiados fueron extraídos de un corpus de habla infantil de taiwanés (TAIRCORP) formado por catorce niños de las edades comprendidas entre los 1;0 y los 5;0 años y del que se seleccionaron las emisiones de un solo sujeto en dos etapas: entre los 1;5 y los 1;6 años (etapa I) y a los 3;0 años de edad (etapa II). Un total de siete sesiones de grabación fueron realizadas y el tiempo total de duración fue casi de unas seis horas (entre todas las sesiones). De estas grabaciones se extrajeron sílabas aisladas no nasales del tipo *li*, *te* o *koh* que contenían una de las seis vocales del taiwanés (/i/, /e/, /a/, /ə/, /ɔ/ y /u/).

Después de medir el primer y el segundo formante de dichas vocales con ayuda de Praat, los resultados fueron los siguientes: la media del F1 de la [i] en la etapa II fue significativamente más baja que la media de esta vocal en la etapa I, mientras que la media del F2 de esta misma vocal fue más alta en la etapa II que en la etapa I. El mismo patrón del F1 mostró la media de la vocal [ə] aunque en el F2 la media formántica de esta vocal en la etapa II fue muy similar a la de la etapa I. Para la [u], la media del F1 no varió demasiado entre ambas etapas (no fue significativamente más baja en la etapa I que en la II) como tampoco lo hizo la media del F2. La media del F1 de la [a] no mostró diferencia entre ambas etapas pero la media del F2 de esta vocal en la etapa II fue significativamente más alta que la de la etapa I. Respecto a la [e], tanto la media del F1 como la del F2 fueron significativamente más altas en la etapa II que en la etapa I. El mismo patrón presentó la [ɔ]. De estos datos se desprende que tan solo para

la [i] se produjo un descenso significativo y únicamente para el F1 entre los 1;8 y los 3;0 años de edad, mientras que en general en el resto de las vocales el F1 y el F2 aumentaron ligeramente con la edad.

Estos resultados no coinciden con los ofrecidos en el trabajo de Gilbert et al. (1997) en el que la media del F1 y del F2 descendía significativamente entre los 1;3 y los 3;0 años de vida, ni con los ofrecidos por autores como Levitt y Auydelott Utman (1992), Clement y Wijnen (1994), Sze (2000) y Chen (2007) en los que la media del F1 y F2 descendió significativamente entre los 2;0 y los 3;0 años de edad. De estos datos puede desprenderse según los autores que excepto para la [u], la altura de la vocal se hace más baja y que la posterioridad de la misma se hace más anterior en la etapa II que en la etapa I (probablemente debido a cambios fisiológicos). Se puede inferir por tanto que el espacio vocálico se desplaza hacia la parte anterior y baja y ello es inconsistente con los estudios previos ya mencionados en los que el espacio vocálico se movió hacia la parte más alta y más posterior (debido a que los valores del F1 y del F2 descendieron significativamente entre los 2;0 y los 3;0 años como ya hemos comentado).

En cuanto a las diferencias en el rango de de las frecuencias de los formantes de las dos etapas, el del F1 y el del F2 de las vocales [i], [a] y [ə] fue significativamente más estrecho en la etapa II que el de la etapa I. Para la [e], el rango del F1 en la etapa II aumentó con respecto al de la etapa I mientras que en el F2 el rango de valores de la etapa II fue más estrecho que el de la etapa I. En cuanto a la [ɔ], el rango de valores del F1 y del F2 aumentó en la etapa II con respecto a la etapa I. La misma evolución del rango formántico presentó la vocal [u] tanto para el F1 como para el F2. De nuevo, no se aportan más datos sobre el resto de vocales.

Una superposición entre las diferentes categorías en el espacio vocálico fue más notoria a los 1;6 años que a los 3;0 años de edad. No obstante, las vocales de los extremos estuvieron perfectamente separadas en el espacio vocálico a la edad más temprana, esto es a los 1;6 años. En esta misma edad las vocales se agruparon mayoritariamente en el área posterior.

Un aspecto que nos gustaría aclarar antes de finalizar el estado de la cuestión, es el de la existencia de ciertos estudios cuya finalidad es la de establecer la correcta articulación de la vocal por el niño. Para cumplir este objetivo, estos estudios no suelen fijarse en el habla espontánea del bebé sino que usan ítems establecidos previamente e intentan que el niño los repita (por ejemplo se pretende que el niño nombre palabras concretas, que repita un sonido un número determinado de veces, etc). Además, la mayoría de estos trabajos, según afirman Selby et al. (2000), se interesan únicamente por las producciones inteligibles de los bebés desechando el resto de vocalizaciones que no solo son muy comunes hasta los 3;0 años de edad (según los autores antes de los 1;7 años las emisiones no inteligibles de los niños pueden dar cuenta de casi

el 50% del comportamiento vocal de estos), sino que pueden transcribirse fonéticamente y además aportan una información muy valiosa sobre las habilidades articulatorias del niño. Consecuentemente al centrarse estos trabajos en la competencia fonológica del niño dejando de lado su dominio fonético, no profundizaremos en el contenido de los mismos y nos limitaremos simplemente a mencionar algunos de ellos.

Entre los primeros estudios encontramos el de Wellman et al. (1931) que fue uno de los que más niños analizó (doscientos cuatro) y cuyo estudio longitudinal comprendía las edades de entre 0;2 y 6;0 años. Algunos meses más abarcó el estudio sobre la articulación vocálica que llevaron a cabo Irwin y Wong (1983) ya que la edad de los niños observados iba desde los 1;6 a los 6;8 años. El periodo de estudio de Hare (1983) es más reducido si bien este incluye a niños de menor edad (de entre 0;2 y 2;0 años). Un espacio de tiempo intermedio es el escogido por Otomo y Stoel-Gammon (1992) que observaron las producciones de seis niños que adquirían el inglés americano entre los 1;10 y los 2;6 años de edad. Paschall (1983) por su parte prefirió fijarse en un momento determinado y analizó la producción vocálica del niño a los 1;6 años. Un estudio que se aparta de la lengua inglesa es el de Stoel-Gammon y Herrington (1990) que analizó las producciones de cuarenta niños hablantes de cantonés desde los 0;10 a los 2;3 años de edad. También en la adquisición de la lengua cantonesa se centra el ya comentado trabajo de Stokes y Wong (2002) quienes descartaron el grupo de sujetos más jóvenes, los que comprendían las edades de entre 0;10 y 1;1 años, precisamente porque la ausencia de palabras inteligibles en estos les impedía establecer el grado de correcta articulación vocálica.

### 3. OBJETIVOS E HIPÓTESIS DE PARTIDA.

Una vez que hemos revisado trabajos de tan distinta índole como los estudios que determinaban las características anatómicas y el desarrollo de las habilidades de control motor del bebé y del niño pequeño a lo largo de su crecimiento, aquellos que indagaban en las capacidades de percepción con las que nace el bebé y su evolución a lo largo del tiempo, los que establecían el inventario de sonidos vocálicos en los primeros años de vida y los que además analizaban acústicamente dichos segmentos, estudios todos que pueden ayudarnos a explicar los resultados que se obtengan en esta investigación, nos disponemos a especificar los objetivos que se pretenden alcanzar a partir de esta tesis doctoral:

- a) Averiguar cuáles son los sonidos vocálicos que producen dos niños que adquieren la lengua española entre los 0;4 y los 3;0 años de vida. Elaboraremos un registro o inventario vocálico con todos los segmentos vocálicos encontrados en las emisiones de los niños entre esas edades.
- b) Clasificar las vocales extraídas de las emisiones de los niños según las categorías vocálicas del español mediante un test de percepción.
- c) Hallar qué vocales emiten los bebés y los niños pequeños con mayor frecuencia a lo largo del periodo de estudio. Intentaremos determinar si las mismas vocales se repiten durante los tres primeros años de vida o si por el contrario la presencia de un determinado timbre vocálico varía a lo largo de las distintas etapas de desarrollo. En relación a este último punto, intentaremos explicar la menor y la mayor presencia de las distintas vocales desde un punto de vista anatómico y perceptivo, y teniendo en cuenta el factor de la influencia de la lengua ambiente.
- d) Analizar acústicamente el F1 y el F2 de las emisiones vocálicas aisladas de los niños entre los 0;4 y los 3;0 años de edad. Observaremos su evolución en el tiempo y comprobaremos si estos formantes descienden, aumentan o se mantienen estables en las distintas fases de crecimiento.
- e) Descubrir a qué edad aparece un espacio vocálico definido. Trataremos de determinar en qué momento el triángulo vocálico del español delimitado por las vocales [a, i, u] puede percibirse.



- f) Averiguar si existen diferencias en el número de emisiones y en la calidad vocálica de estas entre ambos sujetos.

Las preguntas a las que pretendemos dar respuesta con esta tesis doctoral son: ¿qué vocales de la lengua española están presentes a los cuatro meses de vida? ¿Aparecen más vocales según se van desarrollando las estructuras del tracto vocal o su presencia no está relacionada con el crecimiento anatómico? ¿Cuáles son los sonidos vocálicos más comunes en los bebés y en los niños de corta edad? ¿Coinciden los sonidos vocálicos del español observados en las primeras etapas de desarrollo con los encontrados en los bebés y en los niños pequeños que adquieren otras lenguas? ¿Hay una evolución de las frecuencias de los formantes o estas permanecen invariables durante los tres primeros años de vida? Y, en el caso de producirse algún cambio, ¿cuánto y por qué cambian los valores formánticos a lo largo del periodo estudiado?

Las hipótesis de partida son las ya comprobadas en otros estudios que se han centrado en la transcripción y el análisis acústico de los segmentos vocálicos de las emisiones de los niños que adquieren una lengua distinta a la española, y son las que nosotros pretendemos corroborar con este trabajo:

- 1) Los valores formánticos de los bebés superan con creces los valores de los formantes de los adultos (Irwin y Chen [1946], Irwin [1948], Peterson y Barney [1952], Fant [1960], Stevens y House [1961], Eguchi y Hirsh [1969], Lindblom y Sundberg [1969], Cruttenden [1970], Boysson-Bardies et al. [1980], Lieberman [1980], Kent y Bauer [1985], Landberg y Lundberg [1989], Stoel-Gammon y Herrington [1990], Davis y MacNeilage [1995], Kuhl y Meltzoff [1996], Feld [1999], Selby et al. [2000], Sze [2000], Ménard et al. [2004], Dawn Warner-Czyz [2005], Van der Stelt et al. [2006], Kern y Davis [2009], Chen y Kent [2010] y Chi-Lin, en prensa).
- 2) Las primeras vocales en aparecer y las vocales más frecuentes en los primeros meses de vida, y al menos durante el primer año, son las vocales centrales medias y bajas y las vocales anteriores (Irwin y Chen [1946], Irwin [1948], Lindblom y Sundberg [1969], Cruttenden [1970], Boysson-Bardies et al. [1980], Lieberman [1980], Kent y Murray [1982], Kent y Bauer [1985], Landberg y Lundberg [1989], Stoel-Gammon y Herrington [1990], Davis y MacNeilage [1995], MacNeilage [1996], Feld [1999], Selby et al. [2000], Sze

[2000], Ménard et al. [2004], Dawn Warner-Czyz [2005], Van der Stelt et al. [2006], Kern y Davis [2009], Chen y Kent [2010] y Chi-Lin en prensa).

- 3) Algunos sonidos vocálicos pueden articularse al poco de nacer porque el bebé posee las herramientas necesarias para ello, pero otros como [u] necesitan de una maduración anatómica y de una mejora de las habilidades de control motor para poder ser articulados (Kent [1976], Buhr [1980], Tse [1991], Green et al. [2002], Vorperian et al. [2005] y Van der Stelt et al. [2006]).
- 4) En los primeros meses de vida se produce una superposición de las vocales en el espacio vocálico que irá desapareciendo según aumente la edad del niño (Buhr [1980], Lieberman [1980], Kent y Murray [1982], Bond et al. [1982], Kuhl y Meltzoff [1996], Rvachew et al. [2006], Ishizuka et al. [2007], Rvachew et al. [2008] y Chi Lin en prensa).
- 5) A pesar de la inversa relación establecida por Fant (1960) entre la longitud del tracto vocal y las frecuencias de los formantes (de forma que según aumenta la longitud de aquel las frecuencias disminuyen), los cambios dramáticos que sufren las estructuras del tracto vocal durante el primer año de vida no provocan que los valores del F1 y del F2 desciendan de manera significativa (Buhr [1980], Kent y Murray [1982], Kuhl y Meltzoff [1996], Gilbert et al. [1997], Robb et al. [1997] y Chen [2006]).
- 6) Aunque los niños que adquieren la misma lengua meta tienen inventarios vocálicos parecidos, pueden encontrarse diferencias individuales incluso en las primeras etapas del desarrollo (Ferguson y Farwell [1975], Ferguson [1979], Lieberman [1980], Buhr [1980], Kent y Bauer [1985], Ferguson [1986], Camp et al. [1987], Davis y MacNeilage [1995], Sherkane et al. [2002], Chen [2006] y Van der Stelt et al. [2006]).

#### 4. CARACTERIZACIÓN ARTICULATORIA Y ACÚSTICA DE LOS SONIDOS VOCÁLICOS DEL ESPAÑOL.

A pesar de que la investigación presente se centra en el análisis de los dos primeros formantes de las vocales de dos bebés que adquieren el español como lengua materna, y de que puede considerarse por tanto este trabajo un estudio sobre las propiedades acústicas del habla, es necesario comentar una vez más y de manera breve los requerimientos anatómicos y articulatorios que subyacen a dichas propiedades ya que tal y como hemos justificado en la introducción y como afirma Perrier (2005:111): “[...] In natural speech, articulatory movements and the acoustic signal are obviously strongly coupled, since articulatory movements are the source of the acoustics and determine its spectro-temporal characteristics”.

Para la producción del sonido el ser humano cuenta con un conjunto de órganos, cartílagos, tendones, músculos y cavidades (lo que se conoce como *tracto vocal*) cuya función primaria se alejaba de dicho objetivo. En el caso de los pulmones, su cometido principal era el de la respiración, y entre los de la laringe se encontraban el de proteger las vías respiratorias, el de generar una presión torácica necesaria para realizar un esfuerzo como por ejemplo el que conlleva levantar una mesa, o el de ayudar según Ramón y Cajal (1996) a que el bolo alimenticio prosiga su camino al ser tragado. Estas adaptaciones filogenéticas, ya que el ser humano ha transformado estas funciones primarias para la producción de la voz, llevan a Bosma (1975) a afirmar que el habla es una adaptación respiratoria como lo es el llorar o el toser y que esta envuelve además un proceso de espiración.

Además de la obligada adaptación del tracto vocal, el habla humana requiere de tres fases para producirse: respiración, fonación y articulación. El tan citado Kent (1984), que se refiere a estas fases como los movimientos de los mecanismos del habla los cuales consisten en tres grandes componentes funcionales: el sistema respiratorio, el sistema laríngeo y el tracto vocal, advierte de que su tratamiento es necesario en cualquier estudio de adquisición de la lengua. Por ello, hemos decidido incluirlos en nuestra investigación aunque debido a que son muchos y didácticos los manuales en los que se explican estos tres procesos (Jackson-Menaldi [1992], Titze [1994], Baken y Orlikoff [2000], Le Huche y Allali [2004]) nos limitaremos aquí a sintetizarlos y a redirigirlos a la producción de las vocales.

La respiración es la encargada de crear la corriente de aire que constituye la materia inicial de todo sonido. En la inspiración, el diafragma desciende permitiendo que la caja torácica se expanda y el aire entre en los pulmones. En la espiración, el diafragma vuelve a su posición y

al presionar el tórax provoca que el aire se escape por los alveolos a los bronquios y de estos vaya a la laringe.

Una vez que el flujo de aire llega a la laringe y justo antes de que se inicie la emisión, tal y como se explica en Juana Gil (1988), las cuerdas vocales se unen de forma que la presión del aire situado por debajo de la glotis aumenta y cuando esta presión excede a la que mantiene unidas a las cuerdas vocales, estas se separan generándose entre los pliegues vocálicos una fuerza de succión que es conocida como *efecto Bernoulli* y que hace vibrar dicha corriente creándose así un sonido sonoro como en el caso de las vocales. A esta fase se le denomina fonación. Por el contrario, cuando el flujo de aire pasa por entre las cuerdas vocales sin rozarlas, porque estas se encuentran abiertas, el sonido es sordo.

Finalmente, la articulación no es sino la modificación que sufre la columna de aire a su paso por las cavidades faríngea, oral y nasal. En ella intervienen la acción de la úvula, el modo y el punto de articulación. En cuanto a la primera, esta puede permanecer adherida a la pared faríngea para emitir sonidos orales como en el caso de las vocales del español, o puede separarse y quedar suspendida permitiendo el paso de parte del flujo de aire a las fosas nasales como ocurre en los sonidos nasales. La manera de articular el sonido o modo de articulación tiene que ver con si los órganos articulatorios posibilitan la salida del aire sin oponer dificultad, como ocurre en el caso de las vocales, o si por el contrario obstaculizan dicha salida como ocurre con las consonantes (el distinto grado de obstrucción dará lugar a los distintos tipos de consonantes: oclusivas, fricativas, africadas, etc.). Por último, el punto de articulación hace referencia al lugar concreto en el que se produce el sonido. Es este estadio según MacNeilage (1998) en el que reside la principal diferencia entre el habla humana y los sistemas de llamada de los distintos mamíferos (sistemas que no envuelven el componente articulatorio).

En cuanto a esta última fase, las vocales pueden clasificarse desde el punto de vista articulatorio según la posición de la lengua en el eje vertical y en el horizontal, y según el abocinamiento de los labios. Con respecto a la posición de la lengua en el eje vertical, y dependiendo de lo más o menos alejada que se encuentre del paladar la zona más alta de la lengua, las vocales pueden clasificarse en vocales altas, como las vocales españolas [i, u], medias, como [e, o], y bajas como [a] (Gil Fernández [1988]). En la dimensión horizontal, o según la posición más o menos adelantada en que se encuentre la lengua en el eje anteroposterior de la cavidad bucal (Gil Fernández [1988]), los sonidos vocálicos pueden ser anteriores o palatales, si la zona más elevada de la lengua se sitúa en la parte anterior como ocurre en la [i] y la [e], posteriores o velares, si la parte más alta de la lengua está situada en la zona posterior como se da con la [u] y la [o] (término este último que ha sido objeto de

discusión por poder considerarse esta última vocal más uvular que velar), y centrales, cuando la lengua mantiene una posición central como sucede en la [a]. En ambos ejes el cuerpo de la lengua posee un papel más activo que la punta y es probable que en la articulación de estas vocales la intervención de la musculatura extrínseca de la lengua (más lenta y menos concisa) sea mayor que la de la musculatura intrínseca, más precisa y rápida (Perkell [1969]). Finalmente, si un redondeamiento de los labios se produce durante la emisión, las vocales serán redondeadas como [o] y [u], y no redondeadas si dicho abocinamiento no se produce, como en [i], [e] y [a].

Años antes, Tomás Navarro Tomás (1918) había especificado la posición que presentaban los articuladores envueltos en la producción de la vocal española en el momento de su emisión. A continuación ofrecemos la configuración articulatoria para cada una de las vocales:

Para la pronunciación de la [i] (1918:46-47):

“[...] La punta de la lengua se apoya contra los incisivos inferiores; el dorso se eleva contra el paladar duro, tocándolo ampliamente a ambos lados y dejando en el centro una abertura estrecha; este contacto alcanza generalmente por delante hasta los dientes caninos; abertura de las mandíbulas, unos 4 mm entre los incisivos; abertura labial alargada, con las comisuras de los labios un poco retiradas hacia atrás, tensión muscular, media”.

En la emisión de la [e] (1918:50):

“[...] La articulación de esta vocal se forma sobre el paladar duro, correspondiendo a un punto algo más interior que el de la i (aclaración: la i es la i abierta de silba o mirra); la punta de la lengua se apoya contra los incisivos inferiores; el dorso se eleva contra el paladar, tocándolo a ambos lados hasta la mitad aproximadamente de los segundos molares, y dejando en el centro, entre el paladar y la lengua, una abertura mayor que la de la i; la abertura de los labios es asimismo algo mayor que la de la i; abertura de las mandíbulas, entre los incisivos, 6 mm aproximadamente; tensión muscular, media”.

En la articulación de la [a] (1918:54):

“[...] La a requiere una abertura de los labios mayor que la que representan las demás vocales; abertura de las mandíbulas, unos 10 mm, entre los incisivos; la lengua, suavemente extendida en el hueco de la mandíbula inferior, toca con los bordes, a ambos lados, la línea de los molares inferiores, elevando su dorso un poco hacia la parte media de la boca; la punta de la lengua, algo más baja que el borde de los incisivos inferiores, roza la cara interior de éstos hacia las encías; el punto de

articulación determinado por la pequeña elevación del dorso de la lengua corresponde, aproximadamente, al límite entre el paladar duro y el velo del paladar, a igual distancia de los puntos correspondientes a las vocales i, u”.

Para la producción de la [o] (1918:57):

“[...] Los labios avanzan un poco hacia fuera, abocinándose y dando a su abertura una forma ovalada; abertura de las mandíbulas, unos 6 mm, entre los incisivos; la lengua se recoge hacia el fondo de la boca, elevándose por la parte posterior contra el velo del paladar; la punta de la lengua desciende hasta tocar los alvéolos inferiores; tensión media”.

En la realización de la [u] (1918:60):

“[...] Los labios, más avanzados y abocinados que en la o, forman una abertura ovalada pequeña; separación de las mandíbulas, unos 4 mm, entre los incisivos; la lengua se recoge hacia el fondo de la boca, elevándose más que en la o, por su parte posterior, contra el velo del paladar; la punta de la lengua, al nivel de los alvéolos inferiores se separa un poco de estos o solo los roza suavemente, manteniéndose como suspendida en el hueco de la mandíbula inferior; tensión muscular, media”.

En cuanto a la articulación de la [i], cabe matizar que según el análisis de las vocales del español que llevaron a cabo Martínez Celdrán y Fernández Planas (2007) con ayuda de palatogramas, es esta vocal, de todos los sonidos vocálicos, la que mayores requisitos articulatorios exige debido a que es la que conlleva una mayor superficie de contacto lingual con los bordes de la zona palatina.

Tomás Navarro Tomás (1918) precisa también cuáles son las diferentes variantes que estas vocales pueden presentar en función del contexto en el que aparecen y distingue tres tipos de aces (según el eje horizontal de la boca) y variedades abiertas y cerradas para las vocales [i, e, o, u]. Nosotros no incluimos esta clasificación puesto que las vocales que van a ser analizadas aparecen aisladas en la cadena fónica. Por otro lado, resulta innecesario advertir que el bebé recién nacido tendrá dificultades o simplemente será incapaz de pronunciar estos sonidos colocando los órganos articulatorios en las posiciones exactas indicadas por Navarro Tomás. Sin embargo, con el paso del tiempo el bebé irá ganando en desarrollo motor y sus habilidades de control motor aumentarán facilitándole la articulación de los distintos tipos de sonidos.

Cabe mencionar que esta descripción tradicional de la articulación vocálica (basada en la abertura de la boca, el eje antero-posterior y el abocinamiento de los labios) no ha estado exenta de polémica debido tal y como recuerda Iribar (2012) a una fundamentación basada

únicamente en la consideración de la parte más elevada del dorso de la lengua, teniendo en cuenta los ejes ya citados.

Este autor cree que una caracterización articulatoria más acertada del vocalismo debería tener en cuenta otros factores tales como la zona en la que se produce la mayor constricción en la cavidad buco-faríngea. El estudio de esta nueva variable se presenta en su tesis doctoral y forma parte de una investigación en fonética articulatoria experimental, llevada a cabo por el Laboratorio de Fonética de la Universidad de Deusto (proyecto DAELPACE), en la que se tomaron imágenes de resonancia magnética de la articulación de los sonidos vocálicos del español y del euskera precedidos de la consonante bilabial sorda ([p]) de ocho informantes (abreviados en C01-C04 y E01-E04) que presentaron los requisitos necesarios desde el punto de vista lingüístico, articulatorio y fisiológico.

Tras analizar la colección de imágenes de resonancia magnética, Iribar observó que podían distinguirse cuatro parámetros en cuanto a la constricción articulatoria: (1) Zona de mayor constricción (ZMC) o Zona de constricción primaria (ZCP); (2) Zona o zonas de constricción secundaria (ZCS); (3) Punto central de la constricción (PCC); y Punto de mayor constricción (PMC). A partir de estos cuatro parámetros Iribar (2012:50-52) llevó a cabo la descripción de las distintas vocales:

“La vocal [i] presenta una ZMC amplia y anterior, en la que es difícil distinguir un PCC. El PMC varía considerablemente según los informantes [...] El resto del tracto es muy abierto, por lo que no cabe hablar de ZCS.

La vocal [e] presenta una ZMC más reducida que la anterior, y de localización variable: es habitualmente palatal (con diferencias notables en la extensión), pero puede ser postpalatal [...] También hay diferencias en el grado de la constricción. [...] Las diferentes localizaciones de la ZMC suponen que tanto el PCC como el PMC también varíen [...]. El resto de tracto está en general abierto, pero en alguna ocasión (E04) parece advertirse una ligera ZCS faríngea.

La ZMC más indeterminada corresponde a la vocal [a]. No parece claro que, como supone la teoría, la ZMC esté siempre en la parte baja de la faringe. En no pocos casos (C03, E02, etc.), el acercamiento articulatorio que se observa en la faringe se debe especialmente a la acción de la epiglotis. Pero la epiglotis también parecía acercarse a la pared faríngea en algunos casos de [i] y de [e], y sin embargo entonces no lo habíamos considerado como una posible ZCS. No hay que olvidar, además, que un corte mediosagital muestra el punto de mayor acercamiento de la epiglotis: en los sucesivos cortes sagitales que podrían realizarse, la separación de la epiglotis con la

pared faríngea sería cada vez mayor. En cualquier caso, considerar o no a este elemento como un articulador supone situar la constricción en distintas alturas de la faringe.

Sea considerada de una manera u otra, no siempre la constricción faríngea parece con claridad la ZMC de la vocal [a], puesto que en ocasiones (E01, E04) parece situarse en la úvula. El caso es que, además de en la faringe, la vocal [a] presenta otra constricción que consideraremos de todos modos como ZCS– situada en la cavidad oral.

La vocal [o] comparte buena parte de la indeterminación que presenta la vocal [a] en lo referente a su ZMC, teóricamente velo-faríngea. En algunos casos (C02, E04) parece encontrarse, efectivamente, en la zona alta de la faringe, aunque también se registra un contacto de la úvula con la lengua que confunde un tanto la articulación<sup>45</sup>. En otros casos (E03, C02), la ZMC se sitúa directamente sobre la úvula (que sería, entonces, el PMC). Finalmente, hay articulaciones (E01, E02) cuya ZMC se sitúa en el paladar blando. En resumen, la vocal [o] presenta una constricción variable, que puede localizarse desde el paladar blando hasta la zona alta de la faringe.

La vocal [u] presenta una ZMC de extensión generalmente reducida y de localización variable, aunque no tanto como en los dos casos anteriores. Efectivamente, la ZMC puede adelantarse incluso al postaladar (E02), centrarse en el paladar blando (C04) o situarse en la úvula (C02). De este modo, tanto el PMC como el PCC pueden ocupar algún punto a lo largo de dicho espacio. La raíz de la lengua atrasada parece sugerir en ocasiones (E01, E04) la presencia de una ZCS faríngea”.

En cuanto al abocinamiento de los labios, las imágenes muestran con claridad la protusión labial en [u] y no tanto en [o].

No obstante y a pesar de las ventajas de esta caracterización, la descripción articulatoria de las vocales según los ejes tradicionales sigue vigente puesto que tal y como afirma el propio Iribar, recordando a Catford (1981), si bien esta última concepción no es exacta, tampoco es falsa y resulta además apoyada por datos perceptivos y acústicos.

Los procesos de fonación y articulación del sonido pueden explicarse también desde un punto de vista acústico si nos referimos a las diferentes estructuras del tracto vocal como aquellos agentes que modifican la columna de aire procedente de los pulmones para crear una onda sonora. Esta modificación acarrea una perturbación de las moléculas que forman dicha columna de manera que estas abandonan su estado de reposo y transmiten el movimiento a otras moléculas. Este movimiento produce una vibración y se crea la ya mencionada onda sonora. Se trata de una onda compleja porque a la onda procedente de la vibración completa de las cuerdas vocales o ciclo, se le unen las ondas producidas por la vibración de estas en cada una de sus



partes (armónicos) creando esta superposición de ondas simples una onda compleja y periódica o repetitiva. Dicha onda se considera periódica porque aunque no se repita de forma exacta el mismo perfil de ciclo en intervalos de tiempo regulares, puesto que a medida que los sonidos se van amortiguando la amplitud de estos disminuye y los ciclos no son por tanto idénticos, puede considerarse como tal ya que según Gil Fernández (1988) estos matices no ocasionan errores importantes y sin ellos la descripción acústica resulta más sencilla.

A esta onda sonora que consta de una cresta o estado de compresión, que no es sino el acercamiento más próximo de las partículas de aire perturbadas por el movimiento, y de un valle o fase de rarefacción, en la que las partículas se encuentran más separadas de lo normal, también se le denomina frecuencia fundamental. Esta frecuencia posee unos armónicos que son múltiplos de dicha frecuencia fundamental y que según Gil Fernández (1988) no podríamos escuchar si no fuera por la actividad de la resonancia que permite que estos se amplifiquen a lo largo del tracto vocal. Esto es, en su camino hacia el exterior el tono laríngeo puede poner en vibración las distintas cavidades supraglóticas produciéndose así el efecto de la resonancia que puede resultar a su vez en el reforzamiento de determinados armónicos de la onda sonora original siempre que su frecuencia natural de vibración coincida con la de dichas cavidades. Gil Fernández (1988:26-27) explica el fenómeno de la resonancia con ayuda de dos diapasones situados uno junto al otro de forma que si hacemos vibrar uno de ellos, los movimientos de esa vibración actuarán como “pequeños golpes” en el segundo diapason y lo pondrán en movimiento:

“[...] Pero esto solo se producirá si los golpes de las moléculas llegan al segundo diapason en el momento adecuado, eso es, si la frecuencia natural de ambos cuerpos coincide. [...] En otras palabras, la velocidad en que se suceden las compresiones y rarefacciones de la onda incidente (la onda que produce el primer diapason), debe corresponderse con la frecuencia natural de resonador (nuestro segundo diapason) y entonces, no solo el primero resuena en el segundo, es decir, le hace moverse, sino que las amplitudes de ambos movimientos se suman y el sonido que obtenemos se refuerza”.

El tracto vocal actúa por tanto como un resonador (o cuerpo que se ve afectado por las vibraciones de otro) con distintas frecuencias naturales de vibración debido a las múltiples cavidades que lo componen. Tal y como explica Gil Fernández (1988) en su manual sobre los sonidos del lenguaje, los valores de las frecuencias naturales del tracto vocal dependen de la configuración que este adopte de forma que cualquier modificación del tracto vocal implicará un cambio en sus frecuencias naturales de vibración y por tanto el reforzamiento de diferentes armónicos de la onda sonora. De este modo, la amplificación de los armónicos de dos vocales

emitidas por la misma persona y con una misma frecuencia fundamental, será distinta debido a que la forma que presenta el tracto vocal en la articulación de cada una de ellas es diferente. Para ejemplificar estas diferencias acústicas Gil Fernández (1988) toma las vocales [a] e [i] y explica que el hecho de que la [a] del español se articule con la cavidad bucal totalmente abierta, con una separación de unos 10 mm entre los incisivos y con la cavidad faríngea disminuida por el retroceso de la raíz de la lengua, provoca que se observen picos espectrales (reforzamiento de frecuencias o frecuencias de mayor amplitud) a los 700 Hz y a los 1200 Hz, mientras que para la [i], cuya configuración articulatoria requiere que el dorso de la lengua esté elevado hacia el paladar duro dejando un canal estrecho para la salida del aire y que además los labios estén estirados, las frecuencias reforzadas se observan a los 300 Hz y a los 2200 Hz aproximadamente (valores dados para un hablante adulto).

A esta amplificación o refuerzo de las frecuencias, o resonancias, se le denomina *formante*, término usado por primera vez por el físico Ludimar Hermann a finales del siglo XIX. Son estos formantes los que definen el timbre del sonido y esta cualidad acústica la que diferencia unos sonidos de otros (de este modo según Gil Fernández [1988], dos sonidos se percibirán con timbre diferente si el número y la amplitud de los armónicos que componen sus ondas, son igualmente, distintos). En cuanto al número de formantes necesarios para identificar una vocal, desde los años cincuenta hasta la actualidad los experimentos con síntesis de habla, mencionados en la introducción, han venido demostrando que para reconocer una vocal y diferenciarla de otra basta con especificar los valores de sus dos primeros formantes, esto es, los del F1 y los del F2. Frente a esta función identificativa, el F3 aportaría datos acerca del hablante, de su sexo, edad, etc. El profesor Martínez Celdrán (1998:41) a partir de un análisis espectrográfico de las vocales del español matiza esta afirmación y señala que:

“[...] solo las lenguas que distinguen fonológicamente entre anteriores redondeadas/no redondeadas y/o posteriores redondeadas/no redondeadas necesitan de verdad la referencia del tercer formante. A las lenguas, como el español, cuyas vocales anteriores no son redondeadas y las posteriores son redondeadas solo les basta con mencionar los dos primeros formantes”

Las características resonadoras de las cavidades supraglóticas llevaron a muchos investigadores a considerar el tracto vocal como un filtro resonador que deja pasar ciertas frecuencias, amplificándolas, y atenúa otras. Es lo que se conoce como teoría de la fuente y el filtro. Dicha teoría fue propuesta por primera vez en el siglo XVIII por el investigador Kratzenstein (1782) quien construyó un tracto vocal con un conjunto de tubos de tamaños similares a los de las cavidades que lo componen para la producción de las vocales del ruso. Según Kratzenstein (1782) la relación que existe entre los sonidos del habla y el tracto vocal

supralaríngeo es en parte análoga a la que hay entre los tubos de un órgano y las notas musicales puesto que en el órgano de caños la longitud y la forma de cada tubo determinan la cualidad musical de la nota producida por el tubo de la misma manera que en la producción de la voz los cambios en la forma y en la longitud del conducto de aire o tracto vocal humano supralaríngeo determinan que se produzca un sonido u otro. De este modo, los tubos del órgano y del tracto vocal actúan como filtros acústicos.

En el siglo XX Delattre (1948) defendió la consideración del tracto vocal humano como un filtro acústico y añadió la concepción de la glotis como fuente del sonido en la teoría de la fuente y el filtro que sería apoyada de forma más científica por Fant (1960) a comienzos de la década de los sesenta. Según Delattre (1948), la glotis o fuente da una frecuencia que será filtrada por las cavidades de resonancia, o filtro, produciendo así los distintos sonidos del habla humana. Para describir esta teoría, Delattre (1948:239) utiliza un símil entre las cavidades de resonancia y una botella. Para él, el cuerpo de la botella se correspondería con las cavidades laringo-faríngeas mientras que el cuello de la misma constituiría la cavidad bucal<sup>3</sup>. Así explica él mismo el funcionamiento de dicha teoría:

“[...] Les résonances qui caractérisent le timbre d’une voyelle orale résultent du filtrage que subit le ton glottal (la pure vibration des cordes vocales) en passant para la bouche (et par les cavités guturales que nous sous-entendrons ici). La bouche se comporte comme un filtre (ou un résonateur, ce qui revient au même) qui ne laisse passer que certaines des vibrations issues de la glotte. Les fréquences que la bouche laisse passer sont différentes pour chaque voyelle; et si elles sont différentes c’est principalement que les cavités de résonance qui les filtrent changent de forme et/ou de dimensions. [...] des résonances buccaux et gutturaux se comporte sensiblement comme un résonateur dont la forme serait celle d’une bouteille à long goulot; les résonances du goulot seraient celles des cavités qui se trouvent entre les lèvres et le point le plus élevé de la langue, et les résonances de la bouteille celles qui se trouvent derrière le point le plus élevé de la langue. Pratiquement, donc, il semble qu’on puisse parler de deux résonateurs: un résonateur antérieur (le goulot) que nous appellerons résonateur buccal, et un résonateur postérieur (la bouteille) que nous appellerons résonateur gutural”.

---

<sup>3</sup> Para otros autores como Martínez Celdrán la elevación de la lengua es el punto clave de forma que en la articulación de la [i] el cuello estaría conformado por los alveolos, los dientes y los labios mientras que la botella sería el velo y la cavidad faríngea quedando fuera la laringe ya que esta es la fuente.

Stevens y House (1955:485) elaboraron la teoría de la fuente y el filtro como un modelo para la producción vocálica. En este modelo el tracto vocal se concebía como un tubo acústico que poseía dimensiones transversales controladas por la posición de la lengua:

“[...] In order to analyze and evaluate the acoustic performance of the human vocal mechanism during the production of vowels it is convenient to view the vocal tract as an acoustic tube of variable cross-sectional dimensions. The tube is excited at one end (the glottis) by an acoustic source, and sound energy is radiated from the mouth opening at the opposite end”.

Según este artículo de Stevens y House (1955:485) sobre la descripción cuantitativa de la articulación de la vocal, durante la articulación de esta última, la lengua forma una constricción o región de mínima área transversal. Por otro lado,

“[...] As the constriction moves forward, the dimensions of the pharynx increase appreciably; as the constriction moves back, the dimensions of the oral portion of the tract increase. [...] In effect, the x-ray studies indicate that during the articulation of vowels the dimensions of the vocal tract along the length of the tongue are controlled primarily by the position of the tongue constriction and by the degree of tongue constriction”.

En cuanto a los formantes, los autores creen que (1955:488-489):

“[...] Formant 1. In general, high first formants are associated with a narrow tongue constriction near the glottis and an unrounded, large mouth opening. The first formant is low when the mouth opening is small and rounded or when there is a narrow tongue constriction near the mouth opening.

Formant 2. Formant 2 generally increases in frequency as the point of constriction moves forward from the glottis...”

En *Acoustic theory of the speech production*, Gunar Fant (1960) relacionó la configuración anatómica del tracto vocal con el output acústico para definir la teoría de la fuente y el filtro. Para Fant (1960), el tracto vocal actúa como un filtro resonador cuya configuración y fisiología afecta a la frecuencia de la corriente de aire que ha sido modificada por la fuente o glottis creándose así los formantes. En relación a estos, según Fant (1960) son las dos primeras resonancias frecuenciales, esto es el F1 y el F2, las que determinan la cualidad de las vocales.

El filtro es independiente por tanto de la fuente y todo ello puede considerarse como un solo tubo resonador o bien como dos tubos resonadores “the twin-tube resonator model” correspondientes a las distintas partes de las cavidades supraglóticas. En esta última concepción,

el primer tubo comprendería desde la glotis hasta la úvula, y el segundo iría desde la úvula hasta los labios sin contar la radiación de estos (la radiación producida por los labios es independiente del filtro).

Según Gunar Fant (1960) los correlatos articulatorios de los sonidos vocálicos son tres: el punto de constricción de la lengua en la cavidad oral incluyendo el eje vertical o altura de la lengua y el eje horizontal o grado de anterioridad y posterioridad de la misma; el área transversal o *cross-sectional area* de la constricción de la lengua que no es sino una descripción de las dimensiones de las sucesivas partes de la corriente de aire desde la glotis hasta los labios que se logra estableciendo líneas perpendiculares a dicha corriente que permiten establecer el área de resonancia en cada uno de esos puntos o líneas en las cavidades supraglóticas (Stevens y House [1955]); y por último, el redondeamiento de los labios.

Pero quizás el símil utilizado en el manual sobre la fisiología, la percepción y la acústica del habla de Lieberman y Blumstein (1988), esté entre los que ejemplifiquen de forma más clara la función del filtro y los distintos cálculos acústicos. En él, el tracto vocal se compara con un colador de arena y un cubo de playa de un niño. El colador o criba, está repleto de agujeros que no poseen exactamente el mismo diámetro puesto que no se trata de un colador hecho con precisión. Dichos agujeros dejarán pasar los granos de arena cuyo diámetro sea menor que el de ellos, y estos caerán al cubo. Se define así la función de transferencia del colador (tracto vocal) puesto que las partículas de diámetro más pequeño se transferirán o pasarán por el colador al cubo de la misma forma que el tracto vocal transfiere o deja pasar algunas frecuencias y no otras. Así, las cavidades del tracto vocal supralaríngeo actúan como un colador o filtro acústico ajustable que permite a ciertas bandas de longitud de onda del sonido pasar a través de ellas (los filtros sensibles a la longitud de la onda son análogos al filtro de diámetro de la partícula de arena). Puede describirse por tanto la función de transferencia del filtro supralaríngeo del tracto vocal en términos de frecuencia: las frecuencias centrales a las cuales la máxima energía de sonido pase a través del filtro.

La función de transferencia del tracto vocal cambia con la vocal puesto que este transferirá o realzará determinadas frecuencias según la vocal de la que se trate. Así, las diferencias entre los sonidos vocálicos, o cualidad vocálica, son independientes de la actividad de la laringe y son la consecuencia de los cambios en la forma del tracto vocal supralaríngeo que suprime la transferencia de la energía del sonido a ciertas frecuencias y deja pasar la máxima energía a través de él a otras frecuencias. Las frecuencias a las cuales la energía puede pasar a través del conducto vocal supralaríngeo son denominadas frecuencias formánticas (1988:36):

“[...] The frequencies at which local energy maxima may pass through the supralaryngeal air passages are called formant frequencies. The formant frequencies are determined by the length and shape of the supralaryngeal vocal tract, which acts as an acoustic filter”.

Acercándose más a la definición dada por Ludimar Hermann, los investigadores Lieberman y Blumstein (1988) describen los formantes como aquellas frecuencias a las cuales un modo normal de vibración existe en el sistema, es decir, cuando la frecuencia natural de vibración de las distintas cavidades del tracto vocal es igual a la de la frecuencia que lo atraviesa (también para estos investigadores los dos primeros formantes son los necesarios para la identificación de la vocal). Estos formantes están determinados según Lieberman y Blumstein (1988) por las variaciones del área transversal a lo largo de la longitud del tracto vocal (a estas variaciones se le denominan *función de área*).

Otras definiciones del concepto de formante han sido recogidas en distintos estudios acústicos pero todas ellas comparten el significado original. Entre ellas encontramos la de Monsen y Engebretson (1983) según la cual un formante además de ser la propiedad de resonancia del tracto vocal (Stevens y House [1955], Fant [1960] y Pickett [1980]), es la propiedad de la señal acústica en sí misma: concentración de energía a lo largo de la escala de frecuencia definida por la prominencia de varios armónicos. Pero también podemos referirnos al formante según el aspecto o características que este presenta en la herramienta que lo analiza. En este sentido, todos los autores están de acuerdo con la definición de Delattre (1948) de que en un espectrograma tridimensional (gráfico que representan la frecuencia, duración e intensidad en el tiempo) un formante es reconocible como una concentración de oscuridad ascendente o descendente en el tiempo, según cambian sus frecuencias. Semejante descripción es la que dan Denes y Pinson (1963) cuando aluden a los formantes como a las bandas de energía estables localizadas en el espectro frecuencial, y la que ofrecen los autores Monsen y Engebretson (1983) cuando comentan que en un espectrograma el formante se muestra como bandas negras con ascenso y descenso en la frecuencia con el tiempo, o como picos y valles creados por la prominencia de determinados armónicos. Si en vez de un espectrograma utilizamos un espectro para analizar el sonido (gráfico en el que se representan la amplitud y la frecuencia de un sonido), Joos (1948) aclara que los formantes se corresponden con los picos principales del perfil o espectro.

Una vez identificado el formante como uno de los constituyentes acústicos de la vocal y subrayada la importancia de los dos primeros en el reconocimiento de la misma, sería interesante establecer qué correlatos articulatorios son los responsables de estos dos primeros formantes.

Uno de los primeros autores en establecer la relación articuladora-acústica de los formantes fue Delattre (1948) quien al utilizar el símil ya comentado entre una botella y el tracto vocal, añadió que las cavidades guturales<sup>4</sup> del tracto vocal representadas en el cuerpo de la botella serían las responsables del F1 de las vocales, mientras que el cuello de la botella, que comprendería la cavidad bucal desde la úvula hasta los dientes, se correspondería con el F2 de estos sonidos. Por otro lado y en relación a las dimensiones de la cavidad bucal en las que se desplaza la lengua, Delattre (1948) cree que cuanto más se aleja la lengua del paladar, en el eje vertical, más se elevan los formantes, mientras que en el eje horizontal se da una relación inversa entre la frecuencia y la longitud de la cavidad bucal de forma que a medida que aumenta la frecuencia, la longitud de la cavidad bucal disminuye y viceversa (1948:239-240):

“[...] dans le sens vertical... nous constatons qu’il existe une relation constante et directe entre la fréquence et l’ouverture buccale: plus la langue s’écarte du palais, plus les fréquences s’élèvent [...] dans le sens horizontal, nous constatons qu’il existe une relation constante et inverse entre la fréquence et la longueur de la cavité buccale: á mesure que la fréquence augmente, la longueur de la cavité buccale diminue; et inversement, á mesure que la fréquence diminue, la longueur de la cavité buccale augmente.

[...] il existe une relation constante et inverse entre la hauteur de la formante 2 et la longueur de la cavité de résonance buccale”.

Tres años más tarde Delattre et al. (1951) matizaron que el término *altura de la lengua* puede no ser el más apropiado a la hora de definir el F1 y que quizás sí lo sea el término más general de “abertura” del tracto oral. Sus conclusiones con respecto a los correlatos articulatorios responsables de los dos primeros formantes fueron las siguientes (1951:229):

“[...] Conclusion for formant 1. The relation between formant 1 position and articulatory position should be stated in the following terms: There is a direct relation between formant 1 frequency rising and overall opening of the vocal tract. The higher the frequency of formant 1, the wider the overall opening; and inversely.<sup>5</sup>

---

<sup>4</sup> Con el término *gutural*, término que actualmente no se usa, Delattre se refiere a lo que hay inmediatamente después del punto más elevado de la lengua que llega hasta la faringe incluida esta en las cuerdas vocales.

<sup>5</sup> Actualmente se ha comprobado que es la abertura faríngea lo que lo determina de forma que si esta es amplia el F1 es bajo y viceversa.

[...] For formant 2 [...] We shall see here that there is a direct relation between tongue backing and formant 2 frequency lowering [...] In relation to formant 2 lowering, tongue backing is not measured by how far back the highest point of the tongue is (as in the phonetic triangle) but by how far back-and up the tongue as a whole is retracted”.

Con respecto al redondeamiento de los labios, Delattre (1951:231) añadió:

“[...] We may conclude that there is a direct relation between lip rounding and frequency lowering of formant 2: as the lips are rounding formant 2 is lowering, and inversely”.

También para Fant (1960:210) el F1 se corresponde con la cavidad faríngea del modelo del tracto vocal, frente al F2 que depende del grado de anterioridad o posterioridad de la lengua en la cavidad bucal:

“[...] the optimal place of the pass for a high F1 is in the back part of the model, i.e., in the pharynx, and the optimal place for a high F2 is in the middle of the front half of the model, i.e., in the palatal region. F1 decreases and F2 increases if a tongue pass located in the front half of the model is narrowed, and the effect is the reverse within certain limits when the tongue pass is located in the posterior half of the model”.

Para Stevens y House (1995) los primeros formantes son altos cuando hay una constricción de la lengua estrecha cerca de la glotis, los labios no están redondeados y la apertura de la boca es grande. El F2, por otro lado, aumenta su frecuencia cuando la constricción de la lengua se mueve hacia delante en la cavidad oral.

Muchos otros investigadores como Stevens y House (1961), Lindblom y Sundberg (1971), Pickett (1980), etc. han corroborado la relación entre la frecuencia del F1 y la altura de la lengua y apertura de la boca, y la correspondencia del F2 con la anterioridad del cuerpo de dicha lengua, pero solo Ladefoged (1982) sugirió que un indicador más preciso de la ubicación de la lengua en la dimensión horizontal es la diferencia entre el F1 y el F2, más que el grado de anterioridad o posterioridad de la lengua.

En lo que respecta a la especificación articulatorio-acústica en los formantes de las vocales del español, es necesario que acudamos de nuevo a Gil Fernández (1988). Según la autora, la relación entre el aumento del F1 y la apertura de la boca está presente en la [a] cuyo primer formante es el más alto de todas las vocales del español debido a que es también el sonido más abierto de esta lengua. Por el contrario, la [i] y la [u] al ser las vocales más cerradas, poseen el F1 menor de los sonidos vocálicos. Por lo que respecta a la [e] y a la [o] y al articularse estas con un grado medio de apertura de la boca, sus F1 poseen valores medios. En



cuanto al segundo formante y puesto que este depende del grado de anteriorización del cuerpo de la lengua, las vocales anteriores [i] y [e] son las que presentan un F2 más alto frente a la [o] y a la [u] que al articularse en la parte posterior de la cavidad oral poseen el F2 más bajo. La [a], al pronunciarse en la parte central de la boca, posee un F2 medio. Finalmente, en cuanto a la relación entre el descenso del F2 y el abocinamiento de los labios, esta queda manifestada en las vocales [u] y [o] cuyos segundos formantes son los más bajos de todos los sonidos vocálicos debido a que el redondeamiento labial alarga la cavidad de resonancia oral produciendo un descenso en la frecuencia (según Gil Fernández algunos autores piensan que en realidad este abocinamiento incorpora una nueva cavidad).

No obstante, respecto a la relación entre la cavidad y el formante y en cuanto a la especificación del valor formántico en la identificación de la vocal, algunos investigadores creen necesario puntualizar varios aspectos. Por un lado, Delattre (1948) opina que ninguna cavidad es directamente la responsable de la frecuencia de un formante sino que un formante depende, más o menos, de la suma de cavidades, especialmente cuando dicha suma alcanza la forma de un *tube of uniform cross-section*. Por otro lado, son muchos los autores que abogan por la relación entre los formantes en la identificación de una vocal frente a la consideración de sus valores absolutos tal y como señalábamos en el punto 2.2.1 cuando hablábamos de las habilidades perceptivas de los bebés. De esta forma, según Stevens (1983) un índice acústico que permite diferenciar unas vocales de otras es la proximidad entre el primer y el segundo formante. Cuando el F1 y el F2 están próximos, forman una concentración de energía dentro de la región de baja frecuencia mientras que la amplitud frecuencial en las altas frecuencias es relativamente baja. Las vocales que presentan esta estructura formántica son las vocales posteriores. Cuando por el contrario el F1 y el F2 están considerablemente separados, no se produce una concentración de energía espectral en la parte media del espectrograma y en vez de eso el F2 se sitúa más próximo al F3. Las vocales que muestran estas características son las vocales anteriores.

Una vez caracterizadas articulatoria y acústicamente las vocales del español, y a pesar de no tratarse este de un estudio acústico de las emisiones adultas, creemos conveniente indicar los valores de los formantes que el profesor Martínez Celdrán (1995b) obtuvo para cada una de las vocales del español puesto que una de nuestras hipótesis de partida consiste precisamente en demostrar que los valores de los bebés son más altos que aquellos de los adultos. Para extraer los datos, Martínez Celdrán (1995b) pidió a cinco hablantes adultos masculinos y a cinco femeninos, de entre 20 y 30 años, que pronunciaran cada uno cinco veces las vocales del español precedidas de una consonante oclusiva sorda-sonora en un contexto tónico del tipo pam-p/b-V-na, tan-t/d-V-na, kan-k/g-V-na.

La media de los valores formánticos para cada una de las vocales de los hablantes masculinos fue la siguiente: la [i] presentó un F1 de 313 Hz y un F2 de 2200 Hz; la [e] registró un F1 de 457 Hz y un F2 de 1926 Hz; la [a] obtuvo una media de 699 Hz para el F1 y de 1471 Hz para el F2; la [o] presentó un F1 de 495 Hz y un F2 de 1070; y la [u] registró un F1 de 349 Hz y un F2 de 877 Hz.

En cuanto a las medias de los formantes para las vocales emitidas por los cinco hablantes femeninos, estas fueron: el F1 de la [i] presentó unos 369 Hz de media mientras que el F2 alcanzó los 2685 Hz; el F1 de la [e] fue de 576 Hz y el F2 de 2367 Hz; el F1 de la [a] registró unos 886 Hz y el F2 unos 1712 Hz; la [o] presentó un F1 de 586 Hz y un F2 de 1201 Hz; en cuanto a la [u], la media del F1 fue de 390 Hz y la del F2 de 937 Hz.

Los análisis estadísticos realizados por Martínez Celdrán (1995b) demostraron que existen diferencias significativas entre la voz masculina y la femenina. Las vocales femeninas poseen valores más altos en general tanto en el F1 como en el F2. Estos datos, comenta Celdrán, coinciden con los datos para el español por otros autores como Delattre (1965), Quilis y Esgueva (1983) y por el propio Martínez Celdrán (1984).

Para finalizar este apartado sobre la caracterización articulatoria y acústica de las vocales del español y a pesar de que tal y como se verá en el capítulo de la metodología se han desechado los sonidos vocálicos nasales emitidos por los niños, creemos necesario describir los correlatos acústicos de dichos sonidos pues han sido estos los que nos han llevado a identificar las vocales como tales y a excluirlas del análisis.

Desde el punto de vista articulatorio, los sonidos nasales se producen cuando la úvula queda suspendida tras separarse de la pared faríngea permitiendo que parte de la corriente de aire procedente de los pulmones ascienda a la cavidad nasal y sea expulsada por las fosas nasales (lo que se conoce como el puerto velo-faríngeo o v-p). En el caso del hablante adulto y del niño mayor, son estos los que controlan voluntariamente la separación del velo del paladar de la pared faríngea pero en el bebé, tal y como hemos estudiado, la situación es bien distinta debido a su configuración anatómica (ver punto 2.1.1). A pesar del acuerdo general en lo que respecta a los correlatos articulatorios de la nasalidad, los distintos estudiosos no se ponen de acuerdo en cuanto a los índices acústicos que la manifiestan. No obstante, ofrecemos a continuación una revisión de los principales trabajos cuyos experimentos han intentado determinar los efectos acústicos de la nasalidad.

Los primeros índices acústicos de la nasalidad fueron observados en los experimentos con síntesis de habla llevados a cabo por Delattre et al. (1952) y Delattre (1954), y más tarde

fueron confirmados por House y Stevens (1956) y por Hattori et al. (1956), autores estos últimos que analizaron la lengua japonesa. Según estos experimentos, el primer índice en importancia, ya que es el que puede transformar una vocal oral en una nasal, es la reducción de la intensidad en el F1. Para Delattre (1958), es probable que esta reducción en la intensidad del F1 pueda deberse bien a la atenuación de las cavidades fibrosas de la nariz, que según House y Stevens (1956) solo actúan en las ondas de baja frecuencia en el nivel del F1, bien a las antirresonancias que podrían suprimir algunos tonos del F1 según Hattori et al. (1956). De manera aclaradora Gil Fernández (1988) explica que la reducción de la intensidad se comprende mejor si se tiene en cuenta que las frecuencias que se producen en las fosas nasales se oponen a las que tienen lugar en la boca, por lo que la energía de estas últimas se debilita en gran medida.

El segundo lugar en relevancia lo ocupa la presencia de un formante nasal situado alrededor de los 250 cps y al que se le denomina FN1 o primer formante nasal (este índice acústico, que fluctúa entre los 250 Hz y los 450 Hz, fue también encontrado en los estudios de Hattori et al. [1958], Fujimura y Lindqvist [1971], Lindqvist-Gauffin y Sundberg [1976] y Maeda [1982a]). Según Delattre (1948), este segundo correlato contribuye en gran medida a la nasalidad porque cuando está presente solo es necesario un pequeño descenso en la intensidad del F1 para que la vocal sea identificada como nasal. Respecto a este segundo correlato, Matyear et al. (1998) introducen una matización: cuando el F1 oral es alto, como en las vocales centrales bajas del español, el formante nasal está por debajo del oral mientras que si el F1 oral es bajo, como ocurre en las vocales anteriores y posteriores altas del español, la frecuencia del formante nasal está por encima de la del primer formante oral. Respecto a lo que ocurre con el F2, no parece haber un acuerdo en cuanto al efecto de la nasalización.

En algunas ocasiones, apuntan Hawkins y Stevens (1985), podemos observar más de una resonancia adicional en la vecindad del primer formante y dichas resonancias pueden ser atribuidas a la intervención de los senos nasales (Lindqvist y Sundberg [1976] y Maeda [1982a]). En cuanto a la abertura velo-faríngea, los resultados de Maeda (1982a) y los de Hawkins y Stevens (1985) mostraron que dicha abertura debía ser mayor en las vocales bajas que en las altas para producir la nasalidad. Según Maeda (1982a), este hecho indica que la región del primer formante es principalmente la responsable de la percepción de la nasalidad ya que para las vocales bajas, las cuales poseen un F1 con una frecuencia más alta, se requiere un mayor acoplamiento a la cavidad nasal (aumento de la abertura velo-faríngea). Estudios cine-radiográficos, como los desarrollados por Delattre (1968), y técnicas de transmisión por fibra óptica, como las utilizadas por Benguerel y Lafargue (1981), han corroborado estas diferencias en el área del puerto velo-faríngeo dependiendo de la altura vocálica.

Otro índice acústico de la nasalidad es según Chen (1997) el aumento del ancho de banda del primer formante por la implicación de los senos paranasales (que provocan asimismo el descenso de su amplitud). Por su parte, Stevens (1997) cree que este aumento del ancho de banda se produce por una mayor pérdida acústica en la cavidad nasal (seguramente por la actuación de los senos paranasales de la que hablaba Chen). Según este último, puede observarse en el espectro un pico adicional debido a la introducción del par polo-cero alrededor de la frecuencias de 800-1100 Hz, además del pico espectral que según Delattre (1954) puede observarse para las vocales nasales del francés por debajo del primer formante y que puede estar causado según Maeda (1982a) por una resonancia de los senos maxiliares.

Desde el punto de vista perceptivo, Wright (1980) señala que las vocales nasales podrían ser distinguidas de las vocales orales incluso por hablantes de lenguas que no presentan dicha distinción. Esto fue corroborado por Hawkins y Stevens (1985) en cuyo trabajo descubrieron que los oyentes americanos, quienes no distinguían entre vocales nasales y orales en su lengua, fueron mejores en la distinción de la nasalidad que los oyentes de gujarati, hindi y bengali cuyas lenguas sí poseían esa distinción. Por otro lado, diferentes estudios han mostrado que el contexto vocálico puede influir en la percepción de la nasalidad. Krakow y Beddor (1991) encontraron que los hablantes de inglés americano percibían mejor la nasalidad cuando las vocales nasales aparecían aisladas o en contextos orales que cuando estas aparecían en contextos nasales. También la cualidad de la vocal influye en la percepción de la nasalidad ya que tal y como hemos comentado las vocales altas requieren menos acoplamiento nasal que las vocales bajas. No obstante, según Matyear et al. (1998) sigue habiendo muchos transcritores que tienden a no transcribir las vocales nasales y esto se debe a que no son sensibles a ello.

## 5. METODOLOGÍA.

### 5.1. Sujetos.

La selección de sujetos fue una tarea ardua y complicada debido por un lado al largo periodo temporal que comprendía la investigación, tres años de estudio, y por otro a que la frecuencia de grabación era semanal. Estas fueron las razones que nos obligaron a descartar desde un principio a la mitad de los sujetos seleccionados puesto que los padres no aceptaron que sus hijos participaran en un estudio de tan amplia duración y cuyas sesiones de grabación semanales les resultaban pesadas e incompatibles con su vida cotidiana. Estas dificultades hicieron que tan solo pudiésemos contar con dos sujetos, un niño y una niña, para estudiar las vocalizaciones del habla infantil en el periodo de tiempo que va desde los 0;4 a los 3;0 años de vida.

No obstante, un año después de haber estado grabando semanalmente las emisiones de los dos bebés, nos encontramos con un escollo importante ya que uno de los sujetos (el niño varón) abandonó el estudio debido a que sus padres tuvieron que marcharse del país por motivos laborales. Este contratiempo nos obligó a buscar a otro sujeto que debía contar con las mismas características que el anterior: tener menos de 0;4 años y ser además varón. Ocho meses más tarde nuestra búsqueda dio sus frutos y encontramos a una nueva pareja que aceptó que su hijo participara en el estudio.

En cuanto a las características de los niños, se trata de dos bebés españoles, concretamente madrileños, cuya lengua materna es el español y cuya variedad dialectal es la norteña-central que comprende desde Asturias hasta Aragón, incluyendo la zona central peninsular de Castilla la Vieja y Castilla-La Mancha. Esta es también la variedad dialectal de sus padres y de la mayoría de las personas que los rodean. Por tanto, y en cuanto a la homogeneidad en las características de los sujetos que van a ser estudiados, los dos niños comparten la lengua materna, son hablantes de la misma variedad lingüística, han nacido en la misma ciudad y viven en la misma provincia (a unos escasos cuarenta kilómetros).

La coincidencia dialectal y de residencia de los sujetos es recomendada a la hora de estudiar el habla de los bebés tal y como defienden los autores Clopper et al. (2005) y Vorperian y Kent (2007:1514): “[...] Ideally dialect should be taken into consideration in the interpretation of data from any particular study and, more specifically, the place of birth and childhood residence for the characteristics of low vowels and high back vowels”.

### **5.1.1. Nacimiento y entorno de la niña.**

La niña nació el 30 de junio de 2007 en Madrid, cumplidas 40 semanas y 3 días del embarazo. Tanto el periodo de gestación como el estado de la niña después de nacer fueron normales.

La niña fue la primogénita de la pareja. Su padre nació en Madrid el 26 de septiembre de 1972 y vivió allí hasta el nacimiento de la niña. Los padres de este, abuelos paternos de la niña, son emigrantes conquenses (concretamente de la localidad de Mota del Cuervo) que se marcharon a Madrid en busca de oportunidades. La madre de la niña, nacida el 10 de marzo de 1972, es natural de Cabañas de la Sagra, un pueblo situado al noreste de Toledo donde vivió hasta los 18 años. A esta edad se instaló en Madrid donde residió hasta el nacimiento de la niña. Sus padres, abuelos maternos de la niña, siguen viviendo en este pueblo de Toledo.

Unos meses antes de que naciera la niña, los padres se mudaron a San Sebastián de los Reyes, una localidad situada a 19 km al norte de Madrid. En cuanto al nivel económico y académico de estos, ambos poseen títulos universitarios y pertenecen a la clase social media.

A lo largo de los cuatro primeros meses de vida el cuidador que más tiempo pasaba con la niña era la madre, que permanecía con ella durante las horas de luz. Por la tarde, el padre se les unía. Después de estos cuatro meses la niña permanecía 8 horas en la guardería, cuyas cuidadoras habían nacido en distintas localidades de Madrid, y pasaba el resto del día con ambos cuidadores aunque el cuidador más presente era el padre ya que la madre llegaba a casa a última hora de la tarde.

### **5.1.2. Nacimiento y entorno del niño.**

El niño, también primogénito de la familia, nació en Madrid el 29 de abril de 2008 a las 37 semanas de gestación. Su desarrollo en el útero materno y su estado después del nacimiento fueron normales.

Su padre nació en Madrid el 11 de septiembre de 1975 y vivió allí hasta el nacimiento del niño. Los padres de este, abuelos paternos del pequeño, son naturales de la comunidad de Madrid, su padre de Madrid capital y su madre de Cenicientos, un pueblecito al suroeste de Madrid. La madre del bebé nació en Madrid el 4 de octubre de 1975 y vivió allí hasta el nacimiento del niño. Los padres de esta, abuelos maternos del niño, tienen diferente procedencia: su madre procede de Segura de León, un pueblo situado al sureste de Badajoz, y su padre procede de Marazoleja, un pueblecito que está pasada la sierra de Guadarrama, a unos 20 km de Segovia capital.

Un año antes de que el niño naciera, sus padres se instalaron en Rivas-Vaciamadrid, una localidad situada a 15 km al sureste de Madrid. Ambos progenitores tienen estudios universitarios y pertenecen a la clase social media.

En lo que respecta al cuidado del niño, a lo largo de los cuatro primeros meses de vida, y como en el caso de la niña, el cuidador que más tiempo pasó con el niño fue la madre, que permaneció con él durante las horas de luz acompañándolos el padre por la tarde. Desde el cuarto mes de vida y hasta el año y medio, los cuidadores fueron los abuelos maternos ya que los padres no obtuvieron plaza en ninguna guardería. A partir del año y medio el niño acudió a la guardería por la mañana donde era atendido por profesionales madrileñas. Por la tarde, era su padre el que lo acompañaba ya que la madre regresaba del trabajo al anochecer.

### **5.1.3. Evolución de los niños a lo largo del estudio.**

En cuanto al diagnóstico pediátrico, nada más nacer y durante los tres años estudio, ambos sujetos fueron sometidos a distintas pruebas médicas que confirmaron un desarrollo motor, visual, perceptivo, etc. normal. A pesar de este crecimiento normal y en relación al estado de salud, ambos bebés experimentaron los malestares e indisposiciones típicos de los bebés y niños pequeños debidos tanto a factores intrínsecos a su crecimiento, como el dolor ocasionado por la aparición de los primeros dientes, los cambios de temperatura debidos a las vacunas, etc., como a factores extrínsecos a este crecimiento, como el del contagio de enfermedades por la permanencia diaria de los bebés en la guardería. Dichas alteraciones en la salud provocaron que en algunos meses las sesiones de grabación no se realizaran con la misma frecuencia.

Respecto a la niña, el número de emisiones se vio reducido a lo largo del segundo y del tercer año de vida debido a que la incorporación de la pequeña a la guardería conllevó el aumento de enfermedades contagiosas que en algunas ocasiones impidieron que las sesiones de grabación pudieran realizarse, y que en otras mermaron drásticamente el número de sonidos de voz modal emitidos por la pequeña.

En el caso del niño, y además de estas afecciones características de los bebés y de los niños de corta edad, este sufrió en el segundo año de vida un descenso de las defensas que resultó en una mayor propensión al estado de enfermedad tanto por el contagio con otros niños de la guardería como por el desarrollo de enfermedades propias como la de la inflamación de las amígdalas (muy frecuente) o la de la infección de distintas partes del cuerpo. Las consecuencias directas fueron por un lado la reducción del número de sesiones de grabación debido a que el

niño no se encontraba en condiciones aptas para realizarlas, y por otro la disminución de los sonidos válidos para el análisis acústico no solo por el menor número de grabaciones realizadas sino por el hecho de que incluso cuando estas pudieron realizarse, las emisiones del niño se veían afectadas por los taponamientos nasales, la abundancia de mucosidad, la ronquera de la voz, la gangosidad, etc. Esto hizo que la cantidad de sonidos fuera escasa e incluso inexistente en muchos de los meses de este segundo año.

Las causas de que en el tercer año de vida volvieran a reducirse las sesiones de grabación y disminuyeran los sonidos vocálicos válidos para el análisis acústico fueron dos. En primer lugar, se dio la circunstancia de que la madre estuvo hospitalizada por riesgo a perder al segundo bebé que esperaba y ello limitó las sesiones de grabación. Las sesiones volvieron a reducirse cuando desafortunadamente la madre abortó y ello nos impidió grabar al niño durante el periodo recuperación. En segundo lugar, y una vez que las defensas del niño habían mejorado según los informes médicos, este comenzó a sufrir picor de ojos y taponamientos nasales constantes sin que las pruebas médicas señalaran una enfermedad concreta. Pasado un tiempo, los distintos pediatras concluyeron que se trataba de una alergia. Estas circunstancias provocaron un nuevo descenso de las sesiones de grabación de las emisiones del niño y de las muestras válidas para el análisis acústico (por ejemplo entre los 2;2 y 2;3 años, y entre los 2;3 y los 2;4 años, ninguna sesión de grabación tuvo lugar).

## **5.2. Características del estudio.**

### **5.2.1. Sesiones de grabación.**

Se trata de un estudio longitudinal en el que se analizó el habla de dos bebés cuyas emisiones fueron grabadas semanalmente en el periodo de tiempo comprendido entre los 0;4 y los 3;0 años de edad.

Frente a los estudios de carácter transversal en los que tal y como advierten Smith y Kenney (1998) no se reflejan los patrones de desarrollo de los niños de forma individual, los estudios que recogen las emisiones de los niños a lo largo de un periodo de tiempo permiten analizar a fondo el aprendizaje del habla sin que se pierda ningún punto del continuum evolutivo (López Ornat [1994b]). Por otro lado, un estudio longitudinal es preferible a la hora de estudiar la adquisición de cualquier componente del habla del niño (ya sea este morfológico, sintáctico, fonético, etc.) al tratarse dicha adquisición de un proceso que se desarrolla en un periodo de tiempo en el que es importante contar con datos recogidos en diferentes momentos para realizar un análisis exhaustivo de la evolución de los mismos.



A pesar de que se estipuló que las sesiones de grabación tuvieran lugar una vez por semana, la frecuencia de grabación varió debido a la indisposición o enfermedad del niño, de los padres o de cualquier otro familiar cercano (causa que imposibilitó que estas se produjeran con la misma asiduidad). De este modo, en algunos meses las sesiones se redujeron a tres, a dos y a una, o incluso a ninguna en el caso del niño, por los motivos ya explicados en el apartado anterior. A continuación, presentamos dos tablas que especifican el número de sesiones de grabación realizadas en cada uno de los meses de los tres años de vida en los que el habla de los infantes fue registrada. En estas, la columna de la izquierda indica la edad del sujeto y la de la derecha el número de sesiones que tuvieron lugar en ese mes:

**Tabla 5.1. Sesiones de grabación de la niña.**

Sesiones de grabación de la niña					
0;4-1;0 año		1;0-2;0 años		2;0-3;0 años	
0;4-0;5	1	1;0-1;1	4	2;0-2;1	2
0;5-0;6	2	1;1-1;2	4	2;1-2;2	4
0;6-0;7	4	1;2-1;3	4	2;2-2;3	3
0;7-0;8	4	1;3-1;4	3	2;3-2;4	4
0;8-0;9	4	1;4-1;5	4	2;4-2;5	3
0;9-0;10	4	1;5-1;6	4	2;5-2;6	4
0;10-0;11	4	1;6-1;7	4	2;6-2;7	3
0;11-1;0	4	1;7-1;8	4	2;7-2;8	4
		1;8-1;9	3	2;8-2;9	3
		1;9-1;10	4	2;9-2;10	2
		1;10-1;11	3	2;10-2;11	3
		1;11-2;0	4	2;11-3;0	2

**Tabla 5.2. Sesiones de grabación del niño**

Sesiones de grabación del niño					
0;4-1;0 año		1;0-2;0 años		2;0-3;0 años	
0;4-0;5	4	1;0-1;1	3	2;0-2;1	1
0;5-0;6	4	1;1-1;2	4	2;1-2;2	2
0;6-0;7	3	1;2-1;3	4	2;2-2;3	0
0;7-0;8	4	1;3-1;4	0	2;3-2;4	0
0;8-0;9	4	1;4-1;5	2	2;4-2;5	3
0;9-0;10	4	1;5-1;6	4	2;5-2;6	2
0;10-0;11	4	1;6-1;7	4	2;6-2;7	3
0;11-1;0	2	1;7-1;8	4	2;7-2;8	2
		1;8-1;9	3	2;8-2;9	3
		1;9-1;10	2	2;9-2;10	4
		1;10-1;11	4	2;10-2;11	2
		1;11-2;0	2	2;11-3;0	2

Respecto a estas tablas, es necesario matizar algunas cuestiones. En primer lugar, observamos dos diferencias en cuanto a la frecuencia de grabación de las emisiones de ambos niños. La primera es que las sesiones de grabación de la niña durante los dos primeros años de vida cumplieron en su mayoría con el criterio de grabación semanal, esto es cuatro veces al mes, aunque esta frecuencia decayó en el tercer año. Por el contrario, y tal y como se aprecia en la tabla que recoge las sesiones de grabación del niño, tan solo doce meses de los veinticuatro primeros cumplieron con la frecuencia de grabación fijada, reduciéndose esta aún más en el tercer año de vida. La segunda diferencia es que la niña fue grabada en cada uno de los meses de los tres primeros años de vida, mientras que en el niño encontramos meses en los que por las circunstancias explicadas en el apartado anterior, no fue posible grabar sus emisiones.

En segundo lugar, el hecho de que la frecuencia de grabación sea mayor en ambos niños en los dos primeros años de vida no garantiza que en todas las sesiones se encuentren sonidos aptos para un análisis acústico (debido tal y como se ha explicado a los cambios en la salud del pequeño), ni que pueda registrarse el habla del bebé durante el mismo periodo de tiempo en cada una de las sesiones (debido a diferentes contratiempos que aparecieron durante la grabación y que se comentarán en el apartado siguiente). La elevada frecuencia de grabación

por otra parte, y aunque no sea relevante para el objetivo de esta tesis, permitió que el investigador comprendiera cualquier intento de comunicación de los niños.

En cuanto a la obtención de los datos, seguimos un método observacional con el que registramos el comportamiento verbal del niño en situaciones naturales de la vida cotidiana tales como el momento de las comidas, la hora del baño, el tiempo de juego (sin que una selección previa de juguetes fuera realizada para las sesiones de grabación), etc. Se trataba de recoger el habla del niño en distintas situaciones diarias reales tal y como proponía Bruner (1973:13) que frente al lenguaje artificial del laboratorio se posicionaba “a favor del desorden de la vida del hogar” razón por la que el autor comentaba “fuimos hacia los niños en lugar de hacerlos venir hacia nosotros”.

Las sesiones de grabación, cuya duración era de una hora, tenían lugar en la casa de los niños frente a otros estudios en los que las grabaciones se realizaban en el laboratorio o en los que se alternaban ambos lugares como en el trabajo de Zajdó (2006). Una vez en la casa familiar, el investigador registraba el habla del niño en contextos naturales de interacción entre el niño y alguno de sus progenitores o entre el niño y cualquier otro familiar. No obstante, y debido a la elevada frecuencia de grabación, el bebé pronto se familiarizó con el investigador acostumbrándose a su presencia de modo que este pudo participar en las conversaciones espontáneas surgidas en las diferentes situaciones cotidianas (la familiarización fue tal que según los padres de los dos sujetos los bebés preguntaban por el investigador o incluso simulaban llamarle por teléfono cuando este no estaba en la casa familiar). A lo largo de las sesiones, se les pidió a los familiares que interactuaran con el bebé y que cesaran de hablarle una vez que el niño empezara a vocalizar para evitar superposiciones de habla.

Durante los primeros meses de grabación y debido a la corta edad de los bebés, la interacción conversacional no era tan frecuente como en las etapas posteriores de forma que parte de las emisiones del bebé se recogieron en las siguientes circunstancias: cuando alguno de sus padres charlaba con el investigador estando el bebé presente pero sin que ninguno de ellos le prestara atención, y mientras el investigador y el progenitor permanecían en silencio bajo la atenta mirada del bebé. En todas ellas, las emisiones del bebé parecían responder a un interés por llamar la atención.

En esta observación natural de las muestras de habla espontáneas (metodología bautizada como RETAMHE -Registro y posterior Transcripción y Análisis de las Muestras de Habla Espontánea- por Diez-Itza [1992]), se consideraron espontáneas como tal también aquellas vocalizaciones en las que el bebé imitó alguna emisión del familiar o del investigador porque según Stokes y Wong (2002:606): “[...] No attempt was made to exclude imitated

vocalizations from spontaneous production, as this procedure does not cause significant differences in the types and tokens of the sounds produced by typically developing children” (los autores citan a la vez a Ferguson y Farwell [1975] y a Selby et al. [2000]).

La grabación se ejecutó con ayuda de dos grabadoras: una grabadora Sony ICD-P520 que grabó en mono a una frecuencia de muestreo de 16.000 Hz, con una respuesta en frecuencia de 8000 Hz, válida para analizar los dos primeros formantes y cuyo formato de grabación era PCDM o \*wav; y una grabadora Marantz PMD 620 que grabó a una frecuencia de muestreo de 44.000 Hz siendo la frecuencia de muestreo de 22.000 Hz y cuyo formato de grabación también fue \*wav. Los sonidos analizados en el primer año de vida se extrajeron de las emisiones grabadas con la grabadora Sony para ambos bebés. A partir del segundo año de vida, se utilizó la grabadora Marantz.

Se recopilaron un total de 198 horas de grabación que el investigador escuchó con el programa Adobe Audition (versión 3.0) para extraer los sonidos vocálicos que cumplían con los criterios establecidos. De estas 198 horas, 109 horas pertenecieron a las sesiones de grabación de la niña quien fue grabada 27 horas en el primer año de vida, 45 horas en el segundo y 37 en el tercero. En cuanto al niño, se reunieron un total de 89 horas de grabación de las que 29 horas fueron recopiladas durante el primer año, 36 horas durante el segundo y 24 durante el tercero. Como podemos observar, el número de horas de grabación es mayor en la niña que en el niño y en ambos el segundo año de vida es el que comprendió un número mayor de horas de grabación.

### **5.2.2. Criterios de selección.**

Una vez reunidas las sesiones de grabación, el investigador escuchó las emisiones de los bebés a lo largo del proceso de adquisición y seleccionó de ellas los sonidos que reunían las siguientes características:

- Se extrajeron los sonidos vocálicos con plena resonancia formántica siguiendo el criterio articulatorio de Oller (1986). Estos sonidos aparecen en la etapa de expansión (a partir de los 0;4 años) y son sonidos que se articulan con fonación normal y con el tracto vocal abierto.
- Se seleccionaron los sonidos vocálicos aislados no insertos en sílabas y por tanto no precedidos ni seguidos de consonante. A pesar de la controversia que encontraremos si nos asomamos a la literatura en cuanto a la influencia de la consonante y del tipo de consonante en la vocal a la que acompaña en contextos del tipo CV o CVC, decidimos excluir las consonantes porque tal y como

demonstró Mermelstein (1978) incluso las consonantes oclusivas labiales (aquellas en las que no interviene la acción de la lengua) pueden influir y determinar las características la vocal que las acompaña.

Este criterio ha sido el principal causante de que el número de sonidos vocálicos válidos para ser analizados acústicamente no fuera tan elevado como podría esperarse teniendo en cuenta la frecuencia con la que se produjeron las sesiones de grabación. La disminución de las vocales se debe tal y como documentaron Zmarich y Lanni (1998), Zmarich y Miotti (2003) y Zmarich y Bonifacio (2004) a que una vez que comienza el balbuceo, el número de vocales que aparecen de manera aislada se reduce drásticamente en beneficio de las construcciones de consonante más vocal o de consonante vocal consonante (CV o CVC). El estudio de Zmarich y Miotti (2003) aportó datos concretos y nos mostró como una vez que el niño pasaba a la etapa del primer balbuceo (entre los 0;10 y los 1;2 años), el tanto por ciento de sonidos vocálicos aislados emitidos por el pequeño era del 20,2% mientras que el de construcciones del tipo CV era del 67,5%. Estos valores se mantuvieron sin cambios importantes durante el segundo balbuceo (entre los 1;4 y los 1;6 años) pero el tanto por ciento de vocales aisladas volvió a reducirse una vez que comenzó la etapa de la primera palabra donde tan solo se observó un 14,8% de vocales aisladas frente al 74,9 % de estructuras del tipo CV.

Es necesario que destaquemos que ha sido la elección de este criterio, el que establece que la vocal debe aparecer de manera aislada, la que ha provocado que el número de sonidos extraídos y analizados sea muy reducido en comparación con la cantidad de vocales examinadas en los trabajos revisados en el estado de la cuestión, en los que el número de muestras obtenidas fue considerablemente más elevado por la selección y el análisis de todas las vocales emitidas por el niño sin necesidad de que estas tuvieran que aparecer aisladas. Desde nuestro punto de vista, la consideración de la vocal aislada debe prevalecer aunque la consecuencia sea la reducción del número de muestras que pueden ser analizadas.

- Siguiendo los criterios de Buhr (1980), se excluyeron todas aquellas emisiones consideradas como voz no modal, esto es, risas, gritos, lloros, aspiraciones y sonidos vegetativos, así como la voz *creaky*, *breathy* y *whisper* porque a pesar de que estas tienen un condicionamiento físico, bien no presentan formantes

claros, bien presentan formantes distorsionados que no pueden ser tomados como modelos vocálicos.

Es importante aclarar que en muchas de las emisiones de los bebés, y debido al periodo de experimentación por el que estos atraviesan (experimentación con sus órganos, con las posturas, etc.), es frecuente que un sonido sea voz modal al principio de la emisión pero luego presente rasgos no modales si se extiende en el tiempo. Esta peculiaridad contribuyó a la reducción del número de sonidos vocálicos válidos para ser analizados acústicamente, sobre todo en los primeros meses de vida.

- Se descartaron las emisiones que los niños produjeron con un juguete en la boca y aquellas que tuvieron lugar cuando se produjeron de manera simultánea otros sonidos o ruidos (como el de un juguete con sonido, el sonido del teléfono, etc.).
- Se tomaron los sonidos vocálicos cuya  $f_0$  se mantuvo más o menos estable durante toda la emisión y no presentó cambios bruscos. Resulta común en el habla del bebé y del niño pequeño que este comience la emisión con una determinada  $f_0$  y que después vaya modificándola a lo largo de la misma. Estos sonidos no se seleccionaron.
- Los sonidos nasales se excluyeron del análisis acústico, no del test de percepción, debido a las distintas dificultades que planteaba su análisis acústico (dificultades comentadas en el capítulo cuarto que trataba de la caracterización articulatoria y acústica de los sonidos vocálicos del español).

Nos gustaría incidir en el hecho de que este criterio, el de la exclusión de las vocales nasales, ha sido otro de los factores que ha contribuido de manera considerable a la reducción del número de vocales válidas para ser analizadas acústicamente en ambos bebés a lo largo del periodo de estudio. La incidencia de este criterio en la reducción del número de vocales se explica porque debido a los condicionamientos anatómicos que presenta el bebé (tratados en el punto 2.1 del estado de la cuestión), la mayor parte de los sonidos emitidos por este durante los primeros años de vida son nasales de modo que si se excluyen estos, el número de vocales se reduce. Si bien es cierto que la presencia de estos sonidos disminuyó a partir del sexto mes de vida como consecuencia del desacoplamiento de la laringe y naso-faringe, del descenso de la laringe, y de la

separación del velo y la epiglotis (descendiendo esta última), los sonidos nasales no llegaron a desaparecer por completo y estuvieron presentes a lo largo del periodo de tiempo que comprendió el estudio. Además de estos determinantes fisiológicos, factores como la gangosidad o voz con resonancia nasal características no solo de los niños sino de cualquier adulto con síntomas gripales, hicieron que la nasalidad estuviera presente durante todo el periodo estudiado.

Por otro lado, y a pesar de nuestra rigurosidad a la hora de descartar los sonidos nasales, la falta de acuerdo sobre los correlatos acústicos de la nasalidad (expuesta en el capítulo cuarto) y la dificultad añadida que supone el análisis acústico de las vocales de sujetos de tan corta edad han podido favorecer el que se considere como voz modal algún sonido nasal, especialmente en el primer año de vida. Esta circunstancia ha sido señalada en otros trabajos tales como los de Gilbert et al. (1997) en los que los autores advierten de que la dificultad que conlleva el análisis de sonidos en sujetos tan jóvenes ha podido acarrear la inclusión de sonidos nasales descartados en un principio en el estudio.

Finalmente, y debido a que la exclusión de las vocales nasales no solo influyó en la reducción del número de vocales válidas para ser analizadas acústicamente sino en la disminución de las vocales que debían ser escuchadas y clasificadas en las categorías vocálicas del español, decidimos incluir los sonidos vocálicos nasales en el test de percepción para establecer qué cualidad vocálica se repetía más en cada uno de los meses estudiados. Esta inclusión nos ayudó a determinar qué timbre vocálico predominó en cada uno de los periodos analizados incluso en aquellos meses que presentaron un número muy reducido de sonidos vocálicos modales.

- Se descartaron los sonidos vocálicos saturados por causar estos una distorsión de la onda sonora y por ende, de la voz. Se considera sonido saturado aquel que presenta una distorsión de la señal que se produce cuando la amplitud de la onda es mayor que la permitida, lo que conlleva un achatamiento o recorte de la forma de onda. Este criterio contribuyó también, aunque en menor medida que los ya explicados de vocal aislada y de vocal no nasal, a la reducción del número de vocales válidas para ser analizadas acústicamente puesto que era frecuente encontrar sonidos saturados en el habla de los sujetos debido a

factores tales como las fuertes subidas y bajadas de la  $f_0$  típicas de las emisiones de los bebés.

La duración de los sonidos vocálicos extraídos varió de una señal a otra porque el bebé, a estas edades tan tempranas, tiene dificultades para controlar el tiempo de emisión de los sonidos y no repite los mismos esquemas. No obstante, no tuvimos en cuenta dicho correlato acústico a la hora de seleccionar las vocales porque este reducía el corpus ya mermado por el criterio de vocal aislada y de vocal no nasal. Por otro lado, el hecho de que el bebé comenzara la emisión con una vocal perfectamente modal y la acabara con una pedorreta u otras voces no modales (*creacky*, *breathy*, etc.), así como la circunstancia de que el final de la emisión coincidiera con la presencia de cualquier otro sonido o ruido externo a esta, nos llevó a cortar la vocal justo antes de que empezara la voz no modal o el ruido externo contribuyendo también este hecho a que la duración de las vocales no fuera la misma.

Para asegurarnos de que los sonidos extraídos respondían a estos criterios, dos investigadores dedicados al análisis acústico del habla infantil pertenecientes al instituto de investigación alemán “Research Center on Multilingualism” de la Universidad de Hamburgo, seleccionaron al azar un 10% del total de los sonidos del niño y un 10% de los de la niña y comprobaron que estos cumplían con los criterios establecidos.

Siguiendo estos criterios, se extrajeron un total de 1124 sonidos vocálicos de los cuales 300 fueron sonidos modales o válidos para el análisis acústico y 824 fueron nasales. En cuanto a las vocales modales, 188 pertenecieron a la niña y pueden desglosarse como sigue: 106 vocales fueron extraídas en el primer año de vida, 30 en el segundo y 52 en el tercero. En el niño, 112 vocales modales fueron seleccionadas y de estas 54 se extrajeron del primer año de vida, 34 del segundo y 24 del tercero. Respecto a los sonidos vocálicos nasales, se recopilaron 824 de los que 489 pertenecieron a la niña y 335 al niño. De las 489 vocales nasales de la niña, 94 fueron seleccionadas en el primer año de vida, 194 en el segundo y 201 en el tercero. En cuanto al niño, 74 vocales nasales fueron extraídas en el primer año de vida, 175 en el segundo y 86 en el tercero.

Como puede observar el lector, durante el primer año de vida los niños produjeron una mayor cantidad de sonidos vocálicos aislados lo que se tradujo en un mayor número de sonidos válidos para analizar acústicamente. De este modo, durante el primer año de vida fueron 106 los sonidos vocálicos extraídos en la niña y 54 los seleccionados en el niño, frente a las 52 vocales en la niña y las 24 en el niño que fueron extraídas en el tercer año. Esta reducción coincide con las etapas de adquisición fonológica en las que las emisiones de vocal aislada disminuyen en



beneficio de las del tipo CV tal y como mostraron los trabajos de Zmarich y Lanni (1998), Zmarich y Miotti (2003) y Zmarich y Bonifacio (2004).

Se empleará el término de *vocal* para referirnos a los sonidos vocálicos extraídos de las emisiones de los bebés a pesar de la falta de acuerdo en cuanto a la terminología que debe emplearse para aludir a las producciones vocálicas de estos. Autores como Kent y Bauer (1985) creen que la distinta configuración anatómica que presenta el bebé en comparación con la del adulto y los constantes cambios a los que aquella está sometida, favorecen el empleo de los términos usados por primera vez por Martin (1981) de *vocant* y *closant* para referirse a las vocales y a las consonantes precursoras de las vocales y consonantes de la lengua adulta. Otros autores prefieren usar los términos de Grunwell (1987) de *vocoid* y *contoid*, otros los de *vowel-like* y *consonant-like* y otros los de *vocal* y *consonante* para referirse a los sonidos del habla de los niños.

### **5.2.3. Dificultades encontradas en la grabación.**

A la hora de llevar a cabo las grabaciones, nos hemos encontrado con una serie de dificultades debidas a la corta edad de los sujetos, cuyo estado y comportamiento eran impredecibles en cada una de las sesiones, y a que estas grabaciones tenían lugar en la casa familiar donde podía surgir cualquier imprevisto que ni el investigador ni los padres podían controlar y que afectaba a la sesión de grabación de esa semana.

Entre las primeras dificultades, encontramos la de la actitud de los niños la cual resultaba muy cambiante en las distintas sesiones de grabación e incluso a lo largo de una misma sesión. De este modo, uno de los problemas encontrados en los primeros meses de grabación fue el de la reacción que provocaba en los niños la hora de la merienda. En dicha situación, el rechazo del bebé a la fruta se manifestaba con gritos y llantos que traían consigo mucosidad y taponamiento de las fosas nasales que modificaban la cualidad de la voz de los infantes dejando esta de ser modal. Esta circunstancia hacía que se desecharan todas las emisiones producidas por el bebé en lo que restaba de grabación. Por otra parte, según el niño iba creciendo, las rabietas iban aumentando y daban paso a enfados que iban acompañados de llantos y de pataletas que de nuevo modificaban la cualidad de la voz al tiempo que añadían todo tipo de ruidos a las grabaciones (los ocasionados por el lanzamiento de juguetes, las patadas contra el suelo, etc.), obligándonos estos comportamientos a desechar gran parte de estas emisiones.

De manera extrínseca al sujeto, encontramos otros impedimentos que dificultaron el transcurso de la sesión de grabación. Entre ellos destacó la interrupción de la misma por la visita inesperada de algún familiar o de un vecino que podía detener la sesión de grabación durante algún tiempo o incluso durante horas teniendo esta que suspenderse. Por otro lado, imprevistos tales como el pinchazo de una rueda o el ingreso hospitalario de algún familiar redujeron el tiempo de la grabación en algunas ocasiones o incluso cancelaron la misma.

En cuanto a las dificultades técnicas, cabe destacar el hecho de que tuviera que desecharse la primera selección de los sonidos extraídos en el primer año de vida de ambos bebés porque el programa utilizado para su extracción recortó las frecuencias de los mismos y las redujo casi a la mitad. Esto hizo que tuviéramos que comenzar de nuevo a escuchar las grabaciones y a cortar los sonidos que cumplieran con los criterios ya indicados.

### 5.3. Análisis realizados.

#### 5.3.1. Análisis acústicos.

Según Vorperian y Kent (2007) el análisis acústico de las emisiones del bebé y del niño joven es una buena herramienta para reflejar diferentes procesos de desarrollo como el crecimiento de las estructuras del tracto vocal, las diferencias de sexo en esos patrones de crecimiento o la maduración del control motor del habla. Por otra parte, los análisis acústicos son no invasivos, pueden ser realizados con modernos programas y sistemas de ordenador y son aplicables a distintos tipos de emisiones grabadas tanto en el laboratorio como en ambientes naturales. El hecho de que este tipo de análisis supere además las limitaciones de los test de percepción que se valen de la transcripción fonética es otra de las muchas ventajas para McAllister (2009). Para este autor, la transcripción fonética que sigue al test de percepción y que usa segmentos y contrastes de la fonología adulta, puede contener distorsiones e incluso omitir aspectos que oscurezcan detalles importantes del habla del niño. De ahí la necesidad de completar dicha transcripción con un análisis acústico.

Para el análisis acústico se usó un programa informático de análisis de voz llamado *Ancalvoz* en la versión inicial (versión 1.1 de 2009) que pasó a denominarse *PangurBan* en una versión posterior (versión 2.4 de 2011). Dicho programa muestra el espectrograma del sonido (frecuencia e intensidad en el tiempo) junto con un análisis de Fourier (*Fast Fourier Transform*) y un análisis lineal o LPC (*Linear Predictive Coding*), ofreciéndonos la posición y el valor de los formantes sin que puedan estos confundirse con los armónicos. El programa *Ancalvoz-PangurBan* es un subconjunto del programa *Glottex* que se traduce en la “Aplicación para la

evaluación y la calidad de la biometría de la voz”. Ambos programas han sido diseñados por el “Grupo de investigación en Informática Aplicada al Procesado de Señal e Imagen” de la Universidad Politécnica de Madrid (GIAPSI). Tanto *Ancalvoz* como *PangurBan* se utilizan como una aplicación del programa matemático *Matlab*. La función del programa es analizar la onda glótica modificada por el tracto vocal mostrando las frecuencias de los formantes que aparezcan hasta los 8000 Hz.

Las opciones y ventanas de este programa que hemos usado para el análisis de los formantes han sido:

- Ventana editable ‘Ruta Registro’: permite la especificación de la ruta de la carpeta donde se desee registrar un audio nuevo o desde donde se desea abrir el audio. En nuestro caso es la carpeta que contiene los audios.
- Ventana editable ‘Archivo Registro’: permite la especificación del nombre del archivo que se desea abrir, es decir, del audio específico para analizar.
- Ventana editable ‘Duración (seg.)’: permite especificar la duración del registro de audio que vamos a analizar.
- Ventana editable ‘Frecuencia muestreo (Hz)’: especifica la frecuencia a la que se desea muestrear la señal que se va a grabar o la señal ya introducida por nosotros.
- Ventana editable: ‘Tiempo ini.’: determina el instante de inicio de un registro del audio abierto. Por otro lado, la ventana editable ‘Tiempo fin.’ determina el segundo final de la señal. Una vez seleccionadas estas ventanas pulsamos el icono de ‘oír’ para escuchar solamente la señal comprendida en el tiempo que hemos determinado con las opciones anteriores.
- Botón ‘Abrir’: permite abrir el archivo de audio y presentarlo gráficamente en la ventana gráfica superior para su procesamiento.
- Ventana ‘Nivel’: durante la pulsación de los botones ‘test’ y ‘grabar’ mide el nivel promedio de la señal.
- Ventana gráfica superior: presenta el registro de audio activo.
- Botón ‘Oír’: permite reproducir y escuchar el registro de audio abierto.
- Botón ‘Borrar’: permite eliminar el registro activo.

- Botón ‘Evaluar’: permite determinar los intervalos de silencio y de sonido mediante el umbral especificado en la ventana ‘Factor En’ señalando los intervalos de silencio en color rojo. También mide el nivel máximo absoluto alcanzado por la señal grabada en unidades relativas entre 0 y 1 por medio de la ventana ‘Nivel Máximo’. Se presenta en color rojo cuando dicho nivel supera el valor de 0.5 lo que indica que está próximo a la saturación pero no está saturado.
- Ventana desplegable ‘Número de Muestras’: permite determinar el número de muestras que se utilizarán en las ventanas de tiempo de la transformada rápida de Fourier (FFT) en la estimación tiempo-frecuencia.
- Botón ‘Formantes’: permite representar el espectrograma tiempo-frecuencia en la ventana gráfica inferior, y detectar la presencia de los primeros formantes.
- Botón ‘Figura resumen’: una vez que hemos pulsado en ‘Formantes’ pulsamos en ‘Figura resumen’ y este botón nos ofrece en una ventana independiente una figura que contiene dos paneles: serie temporal y análisis LPC. Esta opción nos sirve para corroborar los valores formánticos del espectrograma obtenido en ‘Formantes’ y comprobar que los valores formánticos del espectrograma son los del LPC.
- Ventana gráfica inferior: se presenta en ella el espectrograma del sonido (tiempo-frecuencia) con los formantes de este.
- Ventana informativa ‘Mensaje en proceso’: a través de la misma se emiten unos mensajes de error y de advertencia sobre actuaciones previas imprescindibles.
- Ventana informativa ‘Nivel Máximo’: ofrece el valor máximo absoluto de la señal respecto a 1. Si se supera esta cifra la señal está saturada.

Una vez obtenidas las frecuencias de los dos primeros formantes de los sonidos vocálicos extraídos en cada uno de los meses, procederemos a realizar cartas de formantes en las que quedará reflejada la posición de dichos sonidos según el triángulo vocálico del español. Se trata de gráficos que representan el espacio vocálico concebido por Disner (1986) como aquel gráfico que muestra dónde se localiza una vocal en ambos espacios, el espacio acústico y el articulatorio. Para ello utilizaremos el programa informático *Plotformants* (versión 3.0) que es un programa de acceso gratuito creado por la Universidad de California Los Ángeles que permite representar las vocales en un gráfico a partir de su F1 y su F2 como si fuera una carta de formantes. Para crear estas cartas de formantes hemos seguido los siguientes pasos:

- Hemos creado un archivo de *Excel* con dos columnas en las que hemos recogido los valores del F1 y del F2 de las vocales, y hemos grabado dicho archivo con el formato *.txt* para que el programa pudiera abrirlo.
- A continuación hemos abierto el programa y hemos seleccionado la opción ‘Open’ para abrir el archivo de *.txt*.
- Una vez abierto el archivo hemos pulsado la pestaña ‘Options’ y en la opción ‘Calculation/Display options’ hemos delimitado el valor mínimo y máximo del F1 y del F2 que mostrarán los dos ejes de la carta de formantes (F1: de 200 Hz a 1700 Hz; F2: de 500 Hz a 4000 Hz).
- Finalmente, hemos hecho clic en la pestaña ‘Plot’ y dentro de esta hemos seleccionado la opción ‘F1 vs F2’ que nos ha llevado directamente a la carta de formantes.

### **5.3.2. Test de percepción.**

A pesar de que la mayor parte de los trabajos que han tenido por objetivo la identificación del timbre vocálico o consonántico de las emisiones de habla infantil se han valido de un test de percepción para lograr este cometido, algunos estudiosos han señalado las limitaciones que conlleva la transcripción de los segmentos vocálicos de los bebés y niños pequeños a partir de un test de percepción. Entre ellos encontramos a Oller (1986) y a Koopmans-van Beinum y Van der Stelt (1986) que como comentamos en el apartado 2.3.1 decidieron diseñar ante estas limitaciones sus propios sistemas de codificación, y también a autores como Kent (1984) y Kent y Bauer (1985:492) quienes subrayaron estas limitaciones y las achacaron a la distinta configuración anatómica entre el bebé y el adulto, a las diferencias cuantitativas y cualitativas entre ambos sistemas y a los constantes cambios que sufre el primero: “[...] These anatomical differences may limit the degree to which the units and dimensions of adult language can be used to characterize the vocalizations of infants”.

No obstante, nosotros hemos decidido incluir esta prueba en nuestra tesis doctoral por dos motivos: por un lado, porque como comentábamos en las líneas anteriores el test de percepción sigue siendo la herramienta más utilizada en los estudios de adquisición para determinar el timbre del sonido; por otro, porque este análisis nos permitirá determinar cuál es el timbre vocálico más común en aquellos periodos temporales en los que debido a la corta edad de los sujetos el número de sonidos modales para analizar acústicamente es muy reducido.

En cuanto a las características del test de percepción realizado, estas son las siguientes: cinco sujetos hablantes nativos de español de entre treinta y cuarenta años realizaron la audición e identificación de los sonidos vocálicos por separado, es decir, sin haber tenido acceso previo a los resultados de los otros sujetos, y los etiquetaron según las cinco vocales del español. Los oyentes podían escuchar el sonido tantas veces como les fuera necesario para identificar el timbre vocálico y como ninguno de ellos presentaba conocimientos en Fonética Acústica, no podían apoyarse en el espectrograma vocálico a la hora de identificar la vocal que estaban escuchando de forma que tan solo podían confiar en su audición.

Esta forma de trabajar dificulta en gran medida la identificación de los sonidos de sujetos de tan corta edad y por ello gran parte de los estudios en los que se ha realizado un test de percepción permiten al investigador tener acceso al espectrograma del mismo (es el caso de los trabajos de Gilbert et al. [1997] y de Selby et al. [2000]). Por otro lado, la dificultad que la transcripción del habla de un bebé conlleva fue señalada por Lieberman (1980) quien expuso que incluso para transcripores bien entrenados resulta difícil esta labor de etiquetado. Otro factor que cabe añadir es según Ishizuka et al. (2007) la subjetividad de los transcripores a la hora de transcribir un sonido. Por todo ello, se utilizarán para este estudio las señales modales y nasales a las que cuatro de los cinco oyentes hayan adjudicado el mismo timbre vocálico.

### **5.3.3. Análisis estadísticos.**

Finalmente y para darle a nuestro trabajo una mayor fiabilidad, realizamos un análisis del promedio y de la desviación estándar de las frecuencias del F1 y del F2 de las vocales modales extraídas en las emisiones de ambos bebés a lo largo de los tres años de vida. Se elaboraron gráficos que reflejaron la evolución de ambas medidas a lo largo de los tres periodos estudiados, y tablas que recogieron los valores exactos obtenidos en cada uno de los meses analizados (así como las medias anuales de los tres periodos). Por otro lado, y para comprobar si un descenso significativo de los valores del F1 y del F2 se produjo con el paso del tiempo, se realizó un análisis de varianza sobre el conjunto de valores (ANOVA). Todos estos análisis estadísticos quedan explicados en el capítulo séptimo.

## 6. RESULTADOS.

### 6.1. Resultados del análisis acústico.

Presentaremos en este apartado los valores máximos y mínimos de los dos primeros formantes de los sonidos vocálicos extraídos en cada uno de los meses de los tres primeros años de vida, junto con una carta de formantes en la que estos sonidos quedan representados según las coordenadas del F1 y F2. Esas cartas de formantes nos ayudarán a establecer la posición que ocupa cada sonido en el espacio vocálico y nos permitirán observar la evolución de dicho espacio vocálico a lo largo del periodo de estudio. A partir de las mismas, nos aventuraremos a describir los tipos de vocales que predominan en cada mes según la altura de la lengua en el eje vertical (sonidos altos, medios y bajos), y según su desplazamiento en la dimensión anterior-posterior (sonidos anteriores, centrales y posteriores). Esto nos permitirá determinar las preferencias vocálicas de los niños y su habilidad o falta de ella para producir nuevas vocales a lo largo de su desarrollo. Debido a la escasez de sonidos vocálicos en determinados meses de análisis dicha preferencia será más visible una vez obtenidos los resultados del test de percepción (apartado 6.2), al contar este con un mayor número de señales gracias a la inclusión de las vocales nasales.

Creemos conveniente agrupar los resultados del análisis formántico de los sonidos por sujeto en vez de por edad para que pueda observarse con mayor claridad la evolución del espacio vocálico de un mismo sujeto a lo largo del tiempo, quedando el apartado 6.3 para una futura comparación de los resultados obtenidos en ambos infantes. Arrojaremos en primer lugar los datos obtenidos en la niña y después los del niño.

Finalmente y en relación a los periodos temporales, creemos también necesario hacer una aclaración en cuanto a la terminología empleada para referirnos a los mismos. Si bien hemos utilizado los términos de *primer periodo temporal* para hacer referencia al que transcurre entre los 0;4 años y el 1;0 año de vida, *segundo periodo temporal* para aludir al que avanza entre los 1;0 y los 2;0 años y *tercer periodo temporal* para señalar al que se mueve entre los 2;0 y los 3;0 años, para equiparar nuestro estudio con las distintas investigaciones internacionales nos hemos referido también a dichos periodos con los términos de *primer año de vida*, *segundo año de vida* y *tercer año de vida*. Al aludir por tanto al segundo año de vida, estaríamos apuntando a los meses en los que el niño aún no ha cumplido dicho segundo año pero que forman parte del transcurso temporal existente entre el primer año ya cumplido, el 1;0 año, y el segundo, los 2;0 años (es decir, estaríamos hablando de por ejemplo los 1;4 años). Esta

terminología nos parece también adecuada puesto que una vez cumplido el primer año de vida, el niño se encuentra en el segundo año de su vida aunque aún no posea los dos años completos.

Hecha esta aclaración y aunque hemos utilizado indistintamente ambas expresiones temporales a lo largo del estudio, hemos empleado la terminología en años (y no en periodos) en la denominación de los distintos apartados de los capítulos de esta tesis tras concluir que resultaría más aclarador a la hora de interpretar y comparar los resultados obtenidos. No obstante, el lector sabrá de qué punto temporal se habla en cada momento ya que tanto las cartas de formantes que presentan la posición de los sonidos según su F1 y F2, los gráficos que muestran la evolución de los análisis estadísticos en los distintos periodos, las tablas que recogen los valores exactos del promedio y la desviación estándar en las distintas etapas y las que presentan los resultados del test de percepción, incluyen además de la terminología general de *segundo año de vida*, por ejemplo, los meses exactos en los que se presentan los sonidos/valores según la terminología usada internacionalmente en los trabajos de adquisición: 1;4 años, 1;5 años, 1;6 años etc.

### **6.1.1. Datos obtenidos en la niña.**

#### **6.1.1.1. Primer año de vida.**

Durante el primer año de vida pudieron extraerse de las emisiones de la niña sonidos vocálicos modales en cada uno de los meses analizados.

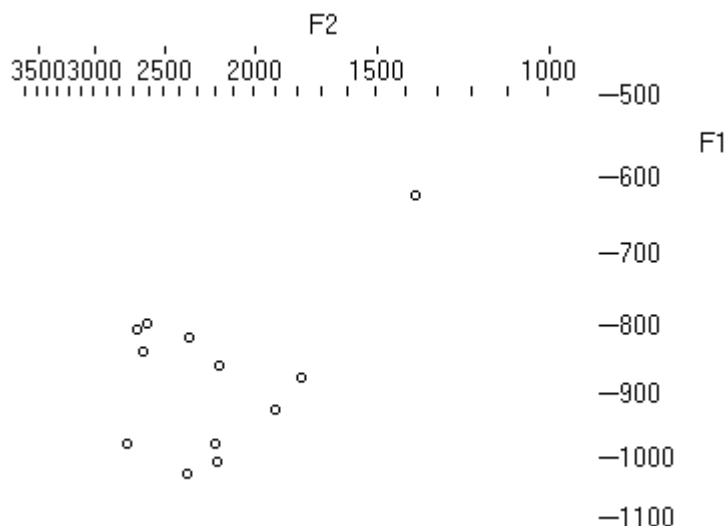
En el primer mes de análisis, esto es entre los 0;4 y los 0;5 años, hemos extraído y analizado un total de trece sonidos de los que cabe destacar:

En cuanto a los valores de los formantes y tal y como se observa en el gráfico 6.1, el F1 se situó entre los 620 Hz y los 1020 Hz por lo que hubo una fluctuación de 400 Hz. El F2 se movió entre los 1380 y los 2770 Hz por lo que la oscilación en este caso fue de 1390 Hz. En lo que respecta al timbre vocálico con el que pueden corresponderse estos valores, gran parte de los sonidos parece ocupar la zona central e izquierda del gráfico con lo que podría tratarse de vocales anteriores y de vocales centrales según la dimensión de antero-posterioridad de la cavidad bucal. Por otro lado, algún sonido posterior puede visualizarse también a la derecha del gráfico. En cuanto a la altura vocálica, se trata de vocales anteriores medias (los valores del F2 por debajo a los 3000 Hz indican una ausencia de vocales altas), de vocales centrales medias y bajas y de posteriores medias. La ausencia de vocales anteriores y posteriores altas impide que



la diferencia entre los valores mínimos y máximos del F2 sea elevada. Llama la atención la superposición de sonidos en la zona central del gráfico.

**Gráfico 6.1. Carta de formantes de la niña 0;4-0;5 años<sup>6</sup>.**

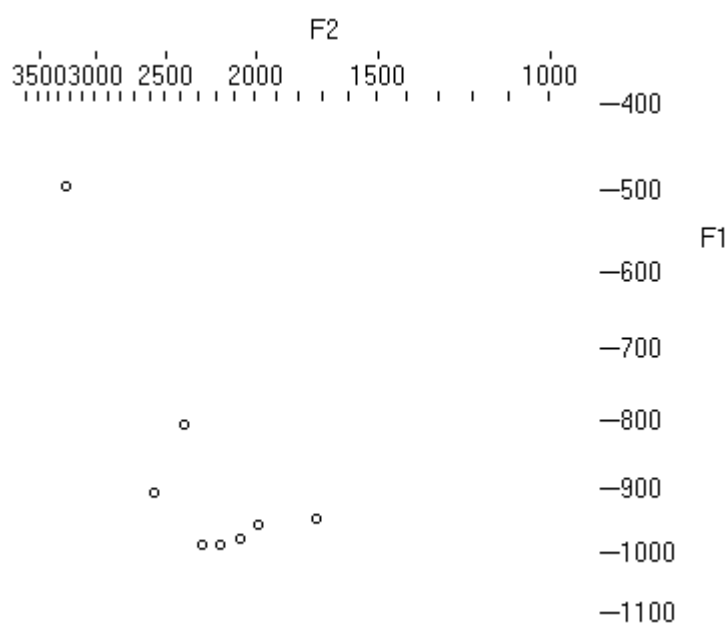


Entre los 0;5 y los 0;6 años se extrajeron ocho vocales modales y tal como refleja el gráfico 6.2 el F1 osciló entre los 490 Hz y 980 Hz, con una diferencia de 490 Hz. El F2 estuvo entre los 1740 Hz y 3280 Hz, con una variación de 1540 Hz. Respecto a la posición que ocuparon estos sonidos según los ejes vertical y horizontal en los que se mueve la lengua, las vocales centrales son las más numerosas estando las vocales anteriores y posteriores mínimamente representadas. La variedad se encuentra en este mes en la altura vocálica puesto que además de vocales centrales medias y bajas encontramos al menos una señal anterior alta debido al elevado F2 que esta presenta. Ello provocó que la diferencia entre los valores mínimos y máximos del F2 fuera más notable que en el mes anterior.

---

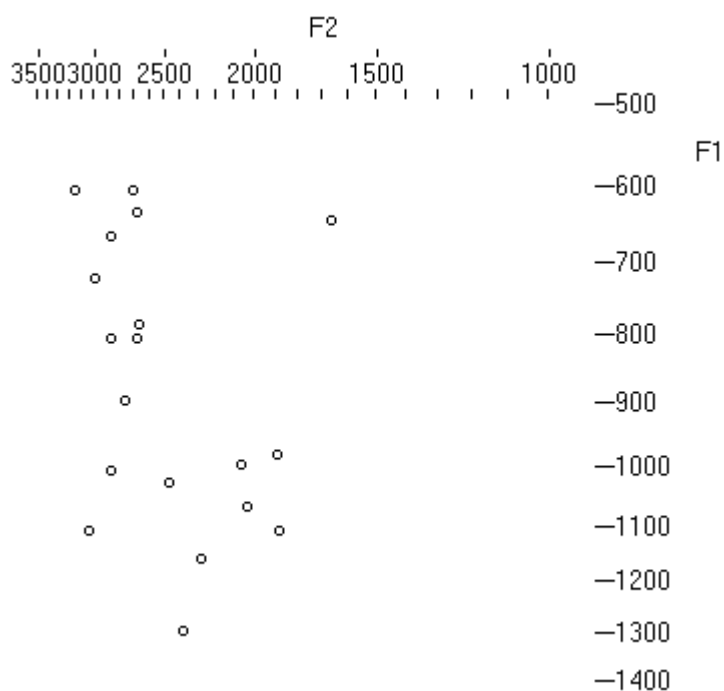
<sup>6</sup> Estos gráficos son en realidad cartas de formantes realizadas con el programa Plotformants que representan a escala logarítmica los valores del primer y del segundo formante de cada una de las vocales con ayuda de dos ejes, el de las abcisas o eje horizontal que se corresponde con el F2 y el de las ordenadas o eje vertical que se corresponde con el F1 (para cualquier duda puede consultarse el *Manual de fonética española* de Martínez Celdrán y Fernández Planas).

**Gráfico 6.2. Carta de formantes de la niña 0;5-0;6 años.**



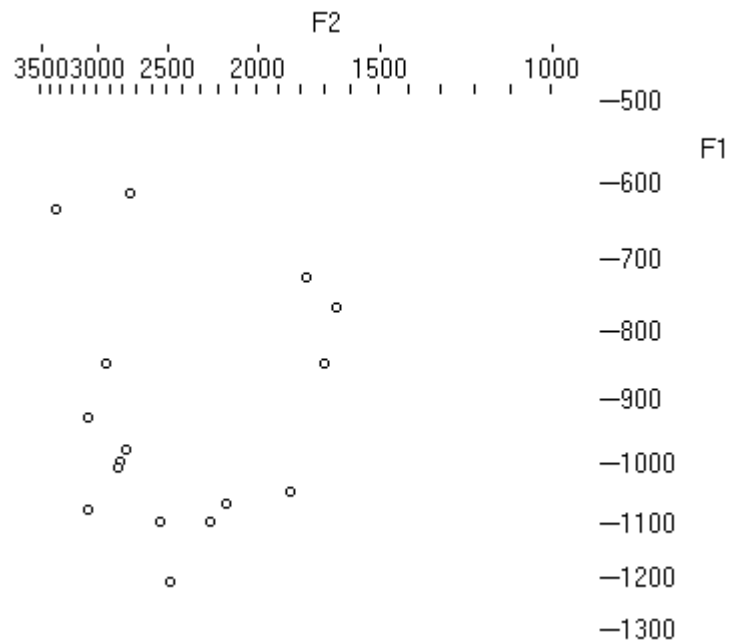
En el periodo de tiempo comprendido entre los 0;6 y los 0;7 años se extrajeron diecinueve sonidos vocálicos de las emisiones de la niña (gráfico 6.3). Los valores del F1 se movieron entre los 600 Hz, valor mínimo, y los 1290 Hz, valor máximo. El F2 por su parte fluctuó entre los 1680 Hz, valor mínimo, y los 3190 Hz de valor máximo. En cuanto a la posición que ocuparon estos sonidos en el gráfico, se trató en su mayoría de vocales anteriores seguidas de vocales centrales y finalmente de posteriores. Respecto a la ubicación de los mismos según la posición de la lengua en el eje vertical, podemos hablar de vocales anteriores medias (y de alguna alta), de vocales centrales bajas y medias, y de vocales posteriores medias. Quizás haya entre estos últimos una vocal posterior ligeramente alta cuyos formantes no distarían mucho de los de la vocal posterior media. El hecho de que puedan estar presentes los cinco timbres vocálicos del español, o al menos cuatro de ellos, favorece que las diferencias entre los valores máximos y mínimos tanto del F1 como del F2 sean importantes (690 Hz en el F1 y 1510 Hz en el F2). Por otra parte, la presencia de estos timbres ayuda a que un tímido esbozo del triángulo vocálico del español pueda observarse en este periodo de tiempo en las emisiones de la niña.

**Gráfico 6.3. Carta de formantes de la niña 0;6-0;7 años.**



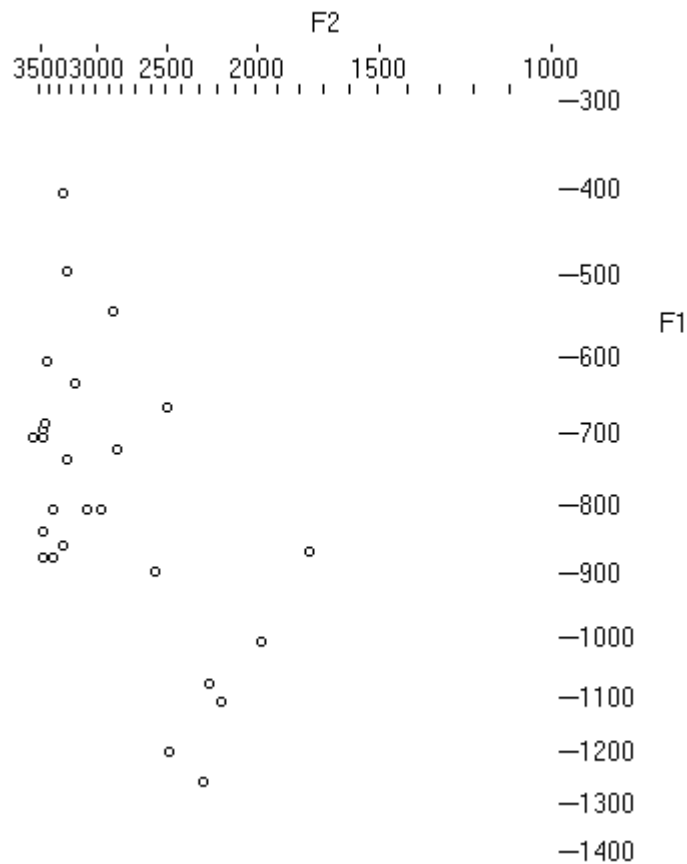
Entre los 0;7 y 0;8 años (gráfico 6.4) se extrajeron dieciséis sonidos vocálicos en los que el valor mínimo del F1 fue de 610 Hz y el máximo de 1200 Hz (diferencia de 590 Hz). En cuanto al F2, este se movió entre los 1670 Hz como valor mínimo y los 3380 Hz como valor máximo (diferencia de 1710 Hz). Puede observarse que la mayoría de los sonidos ocupan la parte izquierda de la carta de formantes, es decir, son anteriores, y que el número de vocales centrales y posteriores es similar. Respecto a la altura vocálica, al menos un sonido anterior alto está presente aunque se trate en su mayor parte de vocales anteriores medias. La presencia de este sonido provoca que la diferencia entre los valores mínimos y máximos del F2 sea mayor. Los sonidos vocálicos centrales son bajos y los posteriores parecen medios. No obstante y como en el mes anterior, quizás alguna vocal posterior alta esté representada.

**Gráfico 6.4. Carta de formantes de la niña 0;7-0;8 años.**



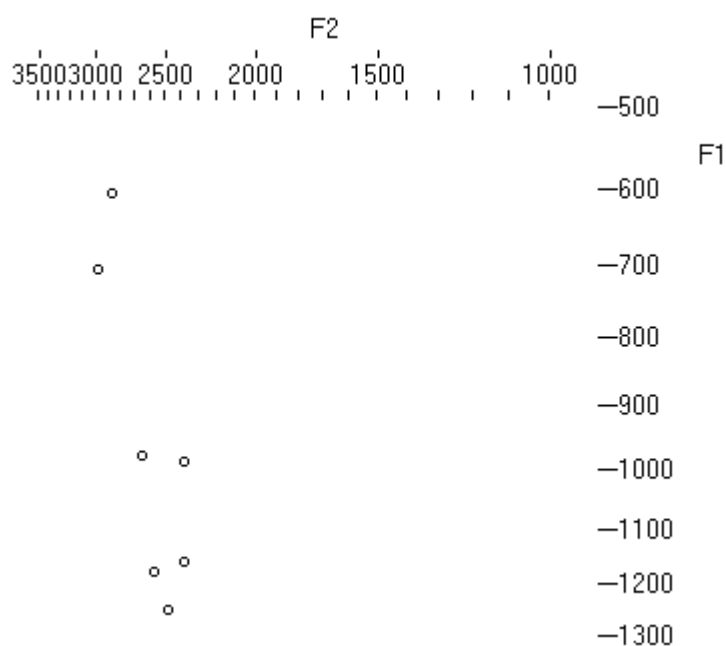
Entre los 0;8 y los 0;9 años los sonidos seleccionados fueron veintiséis tal y como refleja el gráfico 6.5. La diferencia entre el valor mínimo y el máximo en el F1 fue de 850 Hz, siendo el valor mínimo de 400 Hz y el máximo de 1250 Hz. El F2 por su parte, presentó un valor mínimo de 1780 Hz y uno máximo de 3590 Hz manifestando una desigualdad de 1810 Hz. En cuanto al timbre vocálico, la mayoría de las vocales son anteriores altas y medias y una clara superposición entre estas puede observarse en la carta de formantes. Les siguen las centrales bajas y finalmente las posteriores medias. La presencia de vocales anteriores altas y posteriores medias vuelve a ser la causante de la elevada diferencia entre el valor máximo y mínimo del F2 de casi dos mil hertzios. La disposición de estas señales dibuja de nuevo algo similar al triángulo vocálico del español aunque algún ángulo esté ausente.

**Gráfico 6.5. Carta de formantes de la niña 0;8-0;9 años.**



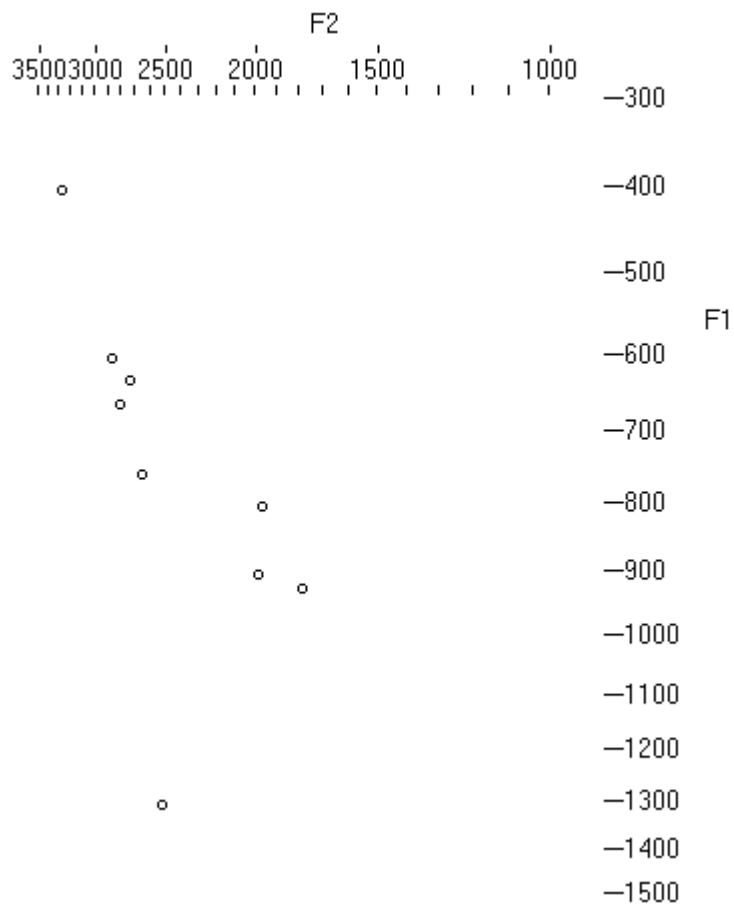
Siete sonidos vocálicos fueron extraídos entre los 0;9 y 0;10 años de vida (gráfico 6.6). En ellos, el valor mínimo del F1 fue de 600 Hz y el valor máximo de 1240 Hz (diferencia de 640 Hz). Por otro lado, el F2 registró un valor mínimo de 2400 Hz y uno máximo de 3000 Hz (diferencia de 600 Hz). En cuanto a la disposición de los mismos, la mayoría de vocales se situó en la parte central e izquierda de la carta de formantes representando por tanto a vocales centrales y anteriores de diferentes alturas. Puede intuirse el triángulo vocálico del español si imaginamos el ángulo posterior a la derecha del gráfico.

**Gráfico 6.6. Carta de formantes de la niña 0;9-0;10 años.**



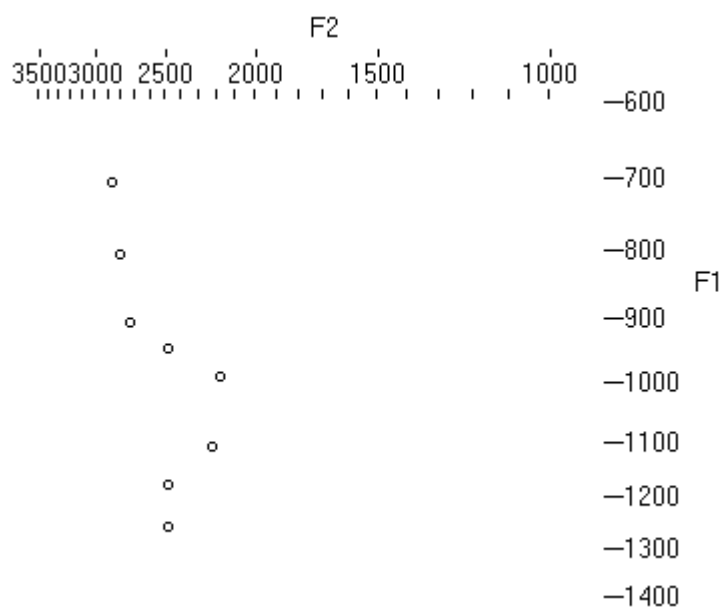
En el penúltimo mes del primer año de vida, es decir, en el que transcurre entre los 0;10 y los 0;11 años, nueve sonidos vocálicos fueron extraídos (gráfico 6.7). Respecto a los valores formánticos, el F1 presentó un valor mínimo de 400 Hz y uno máximo de 1300 Hz (diferencia de 900 Hz), mientras que el F2 registró una valor mínimo de 1800 Hz y uno máximo de 3300 Hz (diferencia de 1500 Hz). Por lo que respecta a la distribución de estos valores en el gráfico podemos observar como la mayoría de los sonidos son anteriores medios aunque parece haber también alguna vocal anterior alta. En la parte central de la carta de formantes encontramos una vocal central baja y en el lado derecho varias vocales posteriores probablemente medias. Los distintos timbres vocálicos identificados en este mes acentúan la diferencia entre los valores mínimos y máximos tanto del F1 como del F2.

**Gráfico 6.7. Carta de formantes de la niña 0;10-0;11 años.**



En último mes del primer año de vida ocho sonidos vocálicos fueron seleccionados (gráfico 6.8). El F1 presentó un valor mínimo de 700 Hz y uno máximo de 1250 Hz con una diferencia de 550 Hz. El F2, por otro lado, registró un valor mínimo de 2200 Hz y uno máximo de 2900 Hz con una diferencia de 700 Hz. En cuanto a las vocales que representan estos valores, se trata de vocales centrales y anteriores.

**Gráfico 6.8. Carta de formantes de la niña 0;11-1;0 año.**



#### **6.1.1.2. Segundo año de vida.**

Debemos aclarar dos puntos antes de comenzar con el análisis acústico de las señales de este segundo año de vida. En primer lugar, el número de vocales extraídas en las emisiones de la niña a lo largo del segundo año de vida es mucho más reducido que el del primero, y entre las causas de este descenso encontramos dos razones principales: por una parte, la disminución del número de vocales que aparecen aisladas en beneficio de aquellas que aparecen en el contexto consonántico (con el comienzo del balbuceo aumenta el número de vocales precedidas y seguidas de consonantes y esta cantidad se ve incrementada a partir del primer año de vida); por otra, los frecuentes resfriados de la niña durante este segundo año de vida (que imprimen a la voz resonancias nasales características de la gangosidad) mermaron la cantidad de sonidos válidos para analizar llegando incluso algunos meses a presentar solo sonidos nasales. No obstante y a pesar de esta disminución de vocales que cumplan con los criterios establecidos para su selección y análisis, proseguiremos nuestra tarea omitiendo aquellos meses en los que no se pudo extraer ninguna vocal modal (como entre los 1;0 y 1;1 años, los 1;1 y los 1;2 años, y los 1;3 y 1;4 años) e incluyendo aquellos otros en los que algunos sonidos vocálicos fueron seleccionados (aunque en algunos de ellos únicamente sean una o dos señales las válidas).

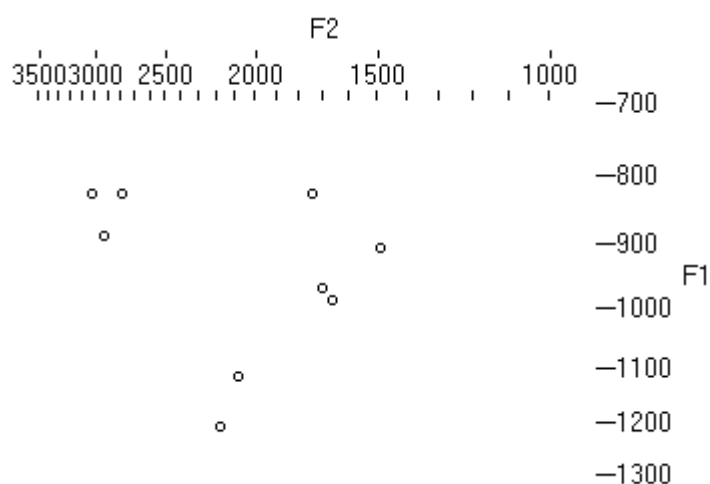
En segundo lugar y como consecuencia de la circunstancia anterior, hemos decidido agrupar las señales modales seleccionadas en este segundo año de vida en dos periodos temporales que reflejen todas aquellas vocales analizadas en la primera mitad de este segundo periodo y todas las examinadas en la segunda. El motivo radica en que debido a la reducción de



vocales por los motivos arriba indicados, en muchas de las cartas de formantes tan solo aparecían una o dos señales hecho que ocasionaba una sensación de poca representatividad. La escasez de vocales en el tercer año de vida de la niña así como en el segundo y en el tercero del niño, hace que repitamos esta misma agrupación en esos periodos temporales.

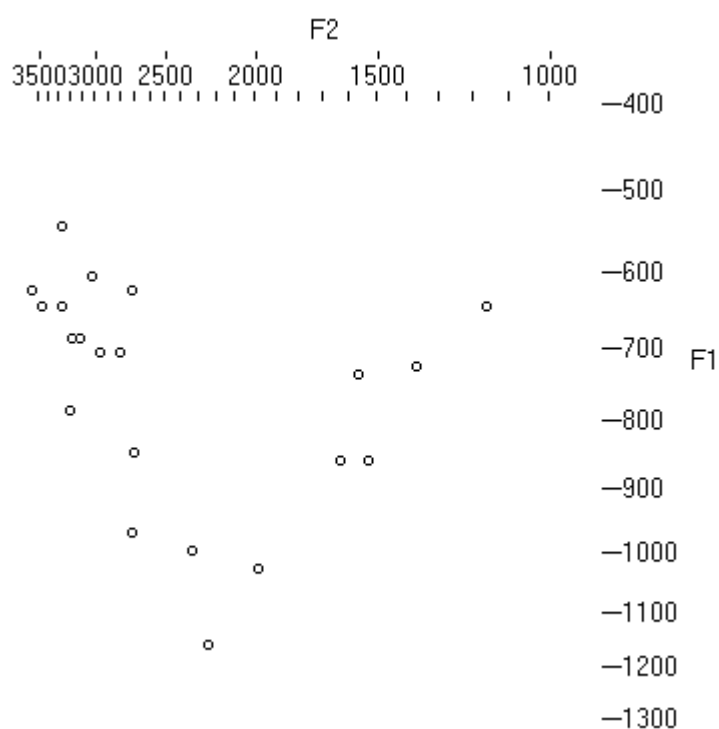
Entre los 1;0 y los 1;6 años nueve sonidos vocálicos fueron analizados (gráfico 6.9). El F1 se movió entre los 820 Hz y los 1200 Hz (diferencia de 380 Hz) mientras que el F2 estuvo entre los 1500 Hz y los 3060 Hz (diferencia de 1560 Hz). A pesar del reducido número de señales los tres ángulos del triángulo vocálico estuvieron representados en proporciones similares. De esta forma, encontramos tres vocales anteriores, dos centrales y cuatro vocales posteriores que delimitan el triángulo vocálico del español.

**Gráfico 6.9. Carta de formantes de la niña 1;0-1;6 años.**



Mayor fue el número de vocales seleccionadas entre los 1;6 y los 2;0 años de vida ya que se extrajeron un total de veintiuna señales (gráfico 6.10). Los valores del primer formante estuvieron entre los 540 Hz y los 1150 Hz (diferencia de 610 Hz), y los del segundo se movieron entre los 1170 Hz y los 3590 (diferencia de 2420 Hz). En cuanto a su ubicación, las vocales quedaron distribuidas, como en el periodo anterior, por todos los ángulos del triángulo encontrándose por tanto vocales anteriores y posteriores, altas y medias, y vocales centrales. Llama la atención el hecho de que que la mayor parte de las señales se agrupe en el extremo izquierdo de la carta lo que supone una mayor presencia de vocales anteriores seguidas por las posteriores y quedando en último lugar las vocales centrales.

**Gráfico 6.10. Carta de formantes de la niña 1;6-2;0 años.**



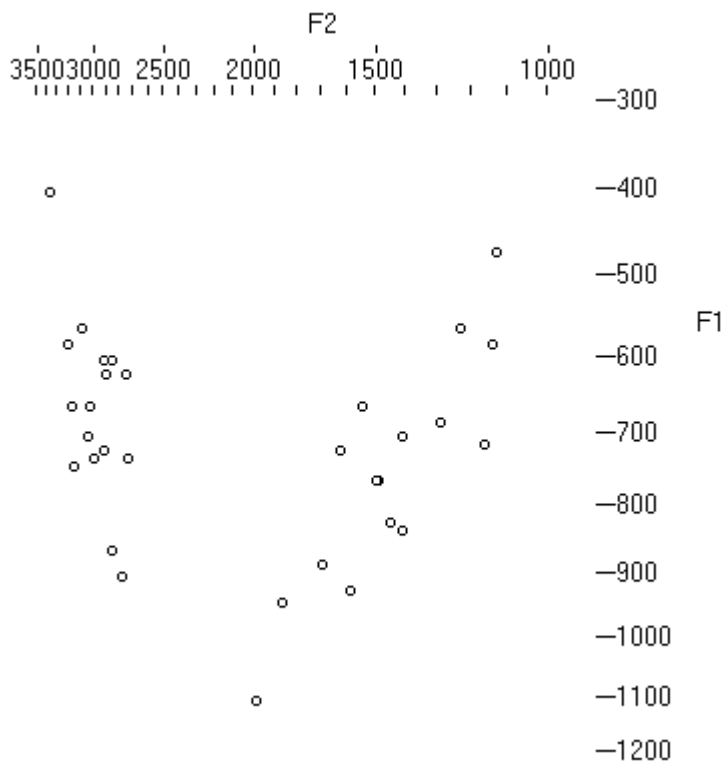
### **6.1.1.3. Tercer año de vida.**

Debido a que el número de vocales extraídas en este tercer año de vida es inferior a la cantidad de vocales seleccionadas durante el primer año por las razones comentadas en el punto 6.1.1.2, y tal y como anunciábamos en dicho apartado, procederemos a agrupar las vocales analizadas en este tercer año de vida en dos periodos temporales que van de los 2;0 a los 2;6 años y de los 2;6 a los 3;0 años. Por otra parte, a lo largo de este tercer año encontramos dos meses en los que ninguna vocal de voz modal pudo ser extraída: entre los 2;0 y 2;1 años y entre los 2;10 y 2;11 años.

En el primer periodo temporal de este tercer año de vida, es decir el comprendido entre los 2;0 y los 2;6 años, treinta y tres vocales fueron seleccionadas (gráfico 6.11). En cuanto a la diferencia de valores, en el F1 esta fue de 700 Hz Hz siendo el valor mínimo de 400 Hz y el máximo de 1100 Hz. Respecto al F2, los valores oscilaron entre los 1140 Hz, valor mínimo, y los 3400 Hz, valor máximo (diferencia de 2260 Hz). La amplia distancia encontrada entre los valores mínimos y máximos del F1 y F2 se debe a la presencia de vocales que ocupan todos los extremos del triángulo vocálico, esto es vocales anteriores, centrales y posteriores, y que además

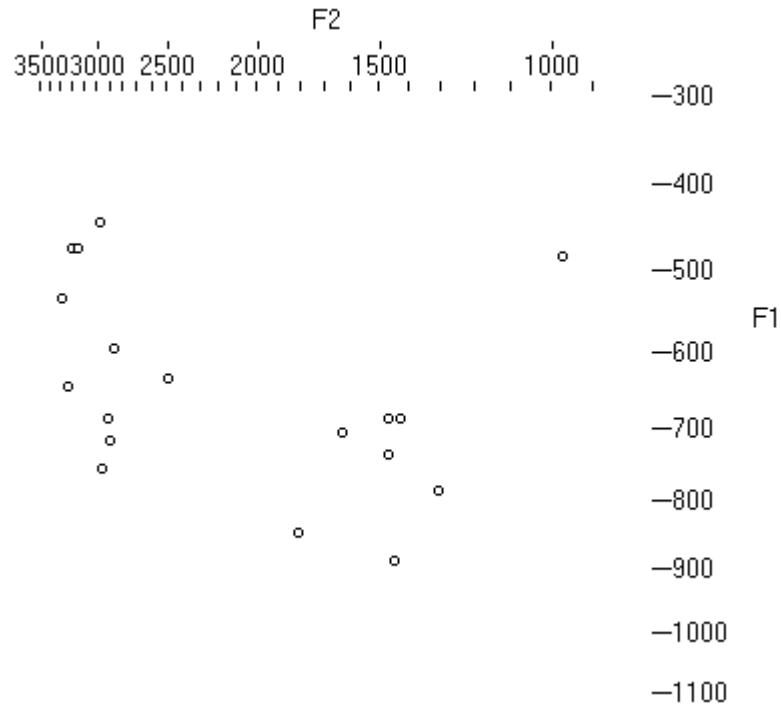
se sitúan a diferentes alturas. Una vez más el número de vocales anteriores supera al de vocales posteriores quedando la cantidad de vocales centrales en último lugar.

**Gráfico 6.11. Carta de formantes de la niña 2;0-2;6 años.**



En la segunda mitad del tercer periodo analizado (gráfico 6.12) se extrajeron un total de diecinueve vocales cuyo F1 se movió entre los 440 Hz y los 880 Hz (diferencia de 440 Hz), y cuyo F2 fluctuó entre los 980 Hz y los 3320 Hz (diferencia de 2340 Hz). Una vez más las vocales situadas en el extremo izquierdo de la carta de formantes, vocales anteriores, superaron en número a las ubicadas en el extremo derecho, vocales posteriores, y en ambos casos encontramos vocales altas y medias. No obstante, llama la atención la escasez de vocales centrales que como ya hemos observado han ido reduciendo su presencia en estos dos últimos años.

**Gráfico 6.12. Carta de formantes de la niña 2;6-3;0 años.**



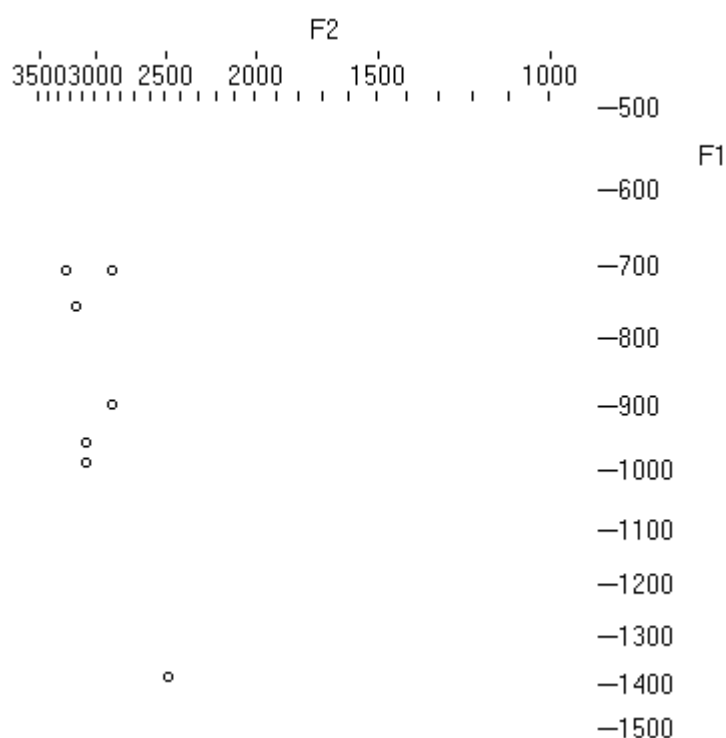
**6.1.2. Datos obtenidos en el niño.**

**6.1.2.1. Primer año de vida.**

Durante el primer año de vida pudieron extraerse vocales de voz modal en cada uno de los meses analizados en el niño. No obstante, el número de señales varió y en algunos meses tan solo encontramos cuatro vocales válidas para realizar el análisis acústico.

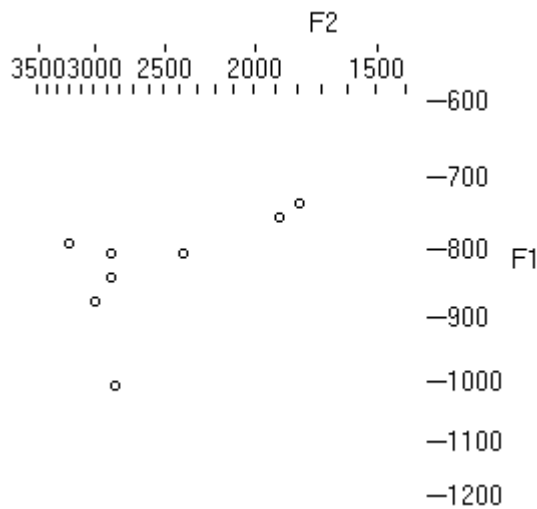
En el primer mes analizado del primer año de vida, esto es entre los 0;4 y los 0;5 años, siete sonidos vocálicos fueron extraídos (gráfico 6.13). El F1 varió entre los 700 Hz y los 1375 Hz mostrando unos 675 Hz de diferencia entre el valor máximo y el mínimo. El F2 por otro lado, presentó un valor mínimo de 2500 Hz y uno máximo de 3265 Hz ofreciendo así una fluctuación de 765 Hz. En cuanto a la posición de estos valores en el espacio vocálico, todas las vocales menos una se situaron en el extremo izquierdo de la carta de formantes, parte anterior de la cavidad bucal, excepto una señal que se ubicó en el centro de la misma. En lo que respecta a la altura vocálica, los sonidos que ocupan la zona anterior pueden corresponderse con vocales medias y altas, mientras que el sonido central parece ser una vocal baja abierta.

**Gráfico 6.13. Carta de formantes del niño 0;4-0;5 años.**



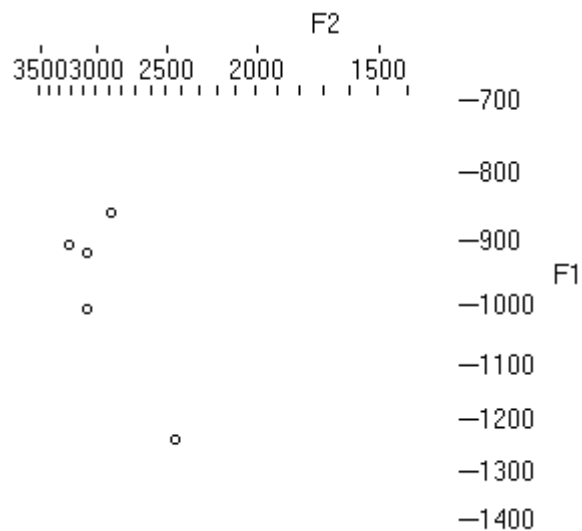
Como podemos observar en el gráfico 6.14, entre los 0;5 y los 0;6 años se extrajeron ocho sonidos vocálicos que presentaron para el primer formante un valor mínimo de 730 Hz y uno máximo de 1000 Hz (diferencia de 270 Hz). En el F2, el valor mínimo fue de 1810 Hz y el máximo de 3230 Hz (diferencia de 1420 Hz). Respecto a la posición de los sonidos en el espacio acústico-articulatorio, la distribución de estos parece dibujar por primera vez el triángulo vocálico del español si bien se trataría de un triángulo escaleno invertido. Se trata pues de vocales anteriores medias, de quizás un sonido vocálico central y de dos vocales posteriores probablemente de altura media.

**Gráfico 6.14. Carta de formantes del niño 0;5-0;6 años.**



Un total de cinco vocales se extrajeron entre los 0;6 y los 0;7 años de vida tal y como refleja el gráfico 6.15. Los valores mínimos y máximos para el F1 en este periodo fueron de 850 Hz y 1230 Hz respectivamente, con una variación de 380 Hz, mientras que para el F2 fueron de 2400 Hz y 3260 Hz, con una oscilación de 860 Hz. Como podemos observar en la carta de formantes, se trata en su mayoría de vocales anteriores altas y medias y de una posiblemente central baja dado su elevando F1.

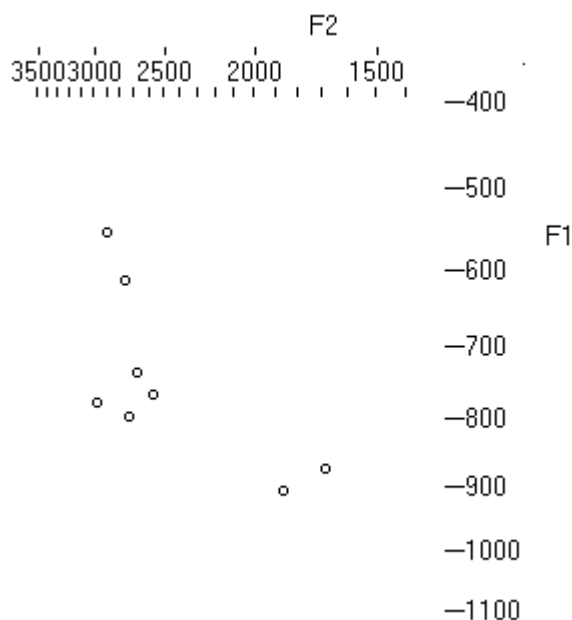
**Gráfico 6.15. Carta de formantes del niño 0;6-0;7 años.**



El número de sonidos vocálicos extraídos entre los 0;7 y los 0;8 años en las emisiones del niño fue de ocho (gráfico 6.16). El F1 varió entre los 550 Hz, valor mínimo, y los 900 Hz, valor máximo. El F2 por su parte, estuvo entre los 1700 Hz y los 3000 Hz (valor mínimo y máximo respectivamente). La notable diferencia entre los valores del F2 (1300 Hz) nos advierte

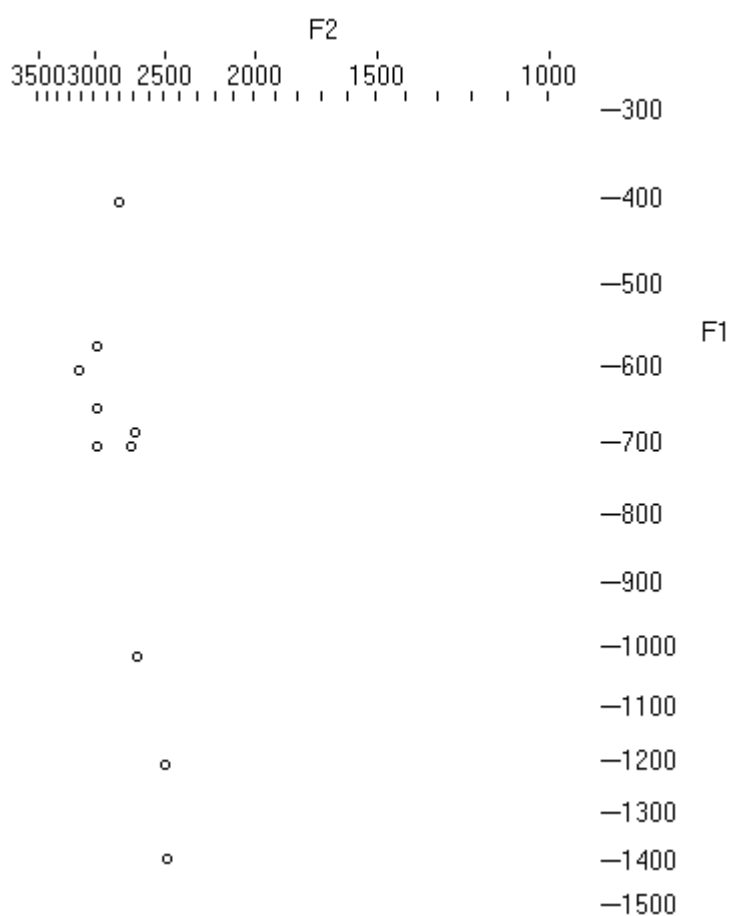
del diferente timbre de las señales incluidas en este mes. De este modo los sonidos vocálicos situados a la izquierda del gráfico apuntan a vocales anteriores, medias y altas, mientras que los ubicados a la derecha del mismo nos remiten a vocales posteriores bajas cuyos valores formánticos parecen muy similares.

**Gráfico 6.16. Carta de formantes del niño 0;7-0;8 años.**



En el periodo comprendido entre los 0;8 y los 0;9 años de vida, un total de diez sonidos vocálicos fueron extraídos (gráfico 6.17). El valor mínimo del F1 fue de 400 Hz y el máximo de 1390 Hz. Los valores del F2 fueron de 2500 Hz y 3160 Hz respectivamente. Comparados con los valores de otros meses, en este periodo las diferencias entre los valores máximos y mínimos tanto del F1 como del F2 no fueron muy elevadas (de 990 Hz y 660 Hz respectivamente). Al contemplar la carta de formantes observamos que la mayor parte de las vocales se agrupan en una columna inclinada hacia la parte izquierda de la misma lo que supone un predominio de vocales anteriores altas y medias. No obstante, el alto F1 de al menos dos de las señales y su posición en la parte central de la carta de formantes parecen indicar la presencia de vocales centrales bajas.

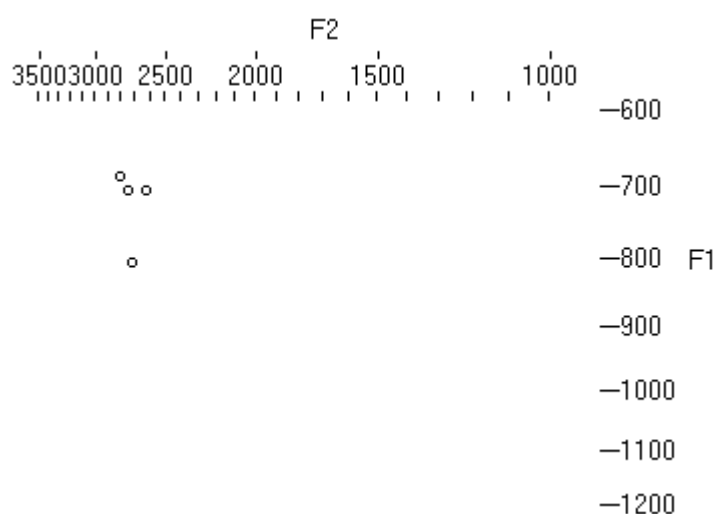
**Gráfico 6.17. Carta de formantes del niño 0;8-0;9 años.**



Cuatro vocales fueron analizadas entre los 0;9 y los 0;10 años (gráfico 6.18). El F1 presentó una variación de 120 Hz siendo el valor mínimo de 680 Hz y el máximo de 800 Hz. El F2 registró unos valores mínimos y máximos de 2650 Hz y 2830 Hz respectivamente habiendo por tanto una diferencia entre ambos de 180 Hz. Una vez más, las señales se agruparon mayoritariamente en la parte izquierda de la carta de formantes lo que nos lleva a señalar la presencia de vocales anteriores probablemente de altura media.

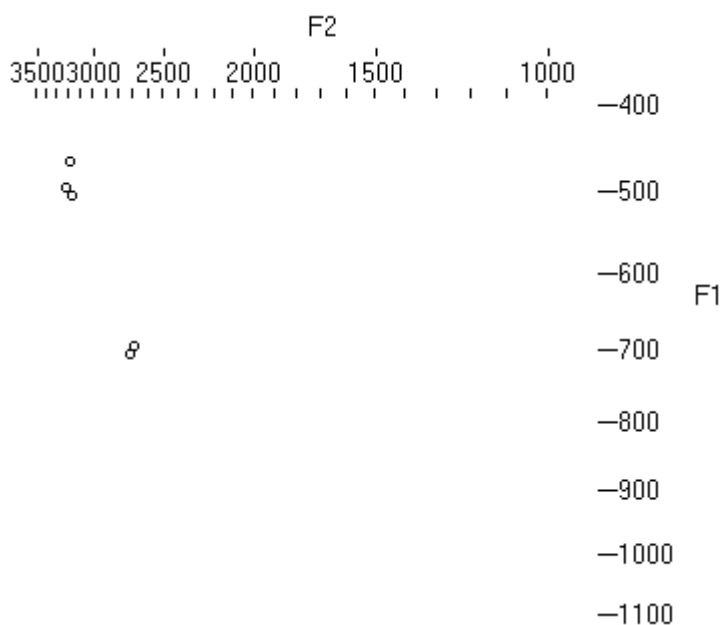


**Gráfico 6.18. Carta de formantes del niño 0;9-0;10 años.**



En el penúltimo mes de análisis del primer año de vida cinco vocales fueron seleccionadas (gráfico 6.19). El F1 se movió entre los 460 Hz y los 700 Hz presentando una diferencia de 240 Hz entre el valor mínimo y el máximo. El F2 registró un valor mínimo de 2720 Hz y uno máximo de 3260 Hz con una diferencia por tanto de 540 Hz. Como podemos observar, las señales se agruparon en la parte izquierda del gráfico lo que nos indica que se trata de vocales anteriores altas y medias.

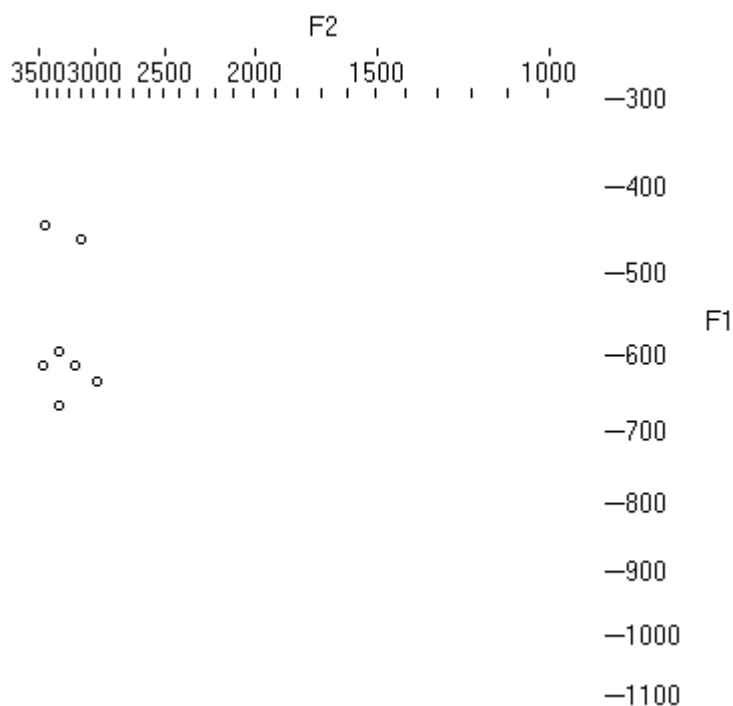
**Gráfico 6.19. Carta de formantes del niño 0;10-0;11 años.**



Siete sonidos vocálicos fueron extraídos en el último mes del primer año de vida, esto es, entre los 0;11 y el 1;0 año, tal y como apreciamos en el gráfico 6.20. El F1 se movió entre

los 440 Hz y los 660 Hz (diferencia de 220 Hz), y el F2 entre los 3000 Hz y los 3470 Hz (diferencia de 470 Hz). En cuanto a la ubicación de las vocales en la carta de formantes estas se situaron en la parte izquierda de la misma por lo que se trata de vocales anteriores de diferentes alturas.

**Gráfico 6.20. Carta de formantes del niño 0;11-1;0 año.**



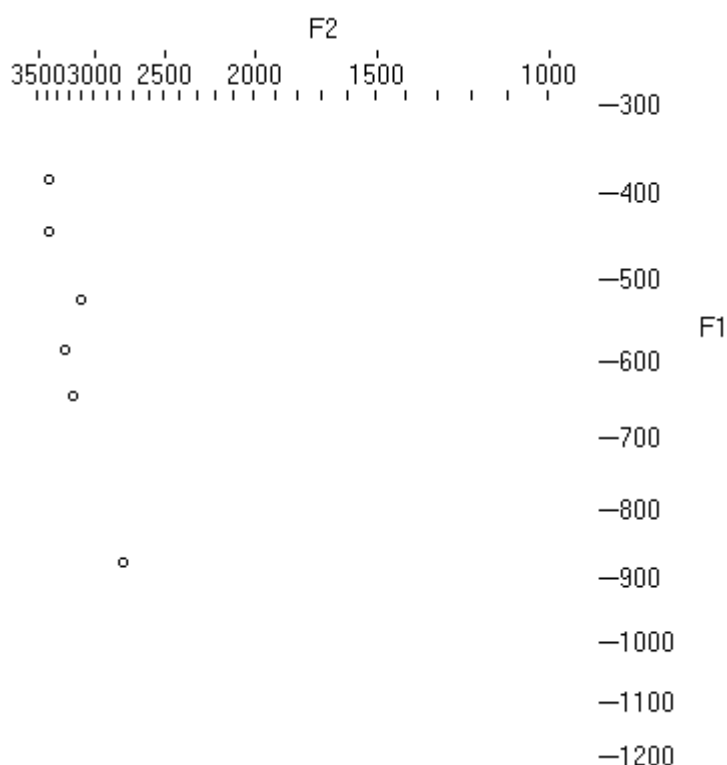
### 6.1.2.2. Segundo año de vida.

El número de vocales que reunió los criterios de selección durante el segundo año de vida fue inferior al número de vocales extraídas en las emisiones del niño a lo largo del primer año. Entre los motivos de esta reducción se encontraron, al igual que en el caso de la niña, la disminución de los sonidos vocálicos aislados a partir del balbuceo, y los problemas de salud diagnosticados en el niño durante este segundo año de vida ya comentados en el apartado 5.1.3 del capítulo de la metodología. Estos factores no solo redujeron drásticamente el número de vocales que debían ser analizadas acústicamente sino que provocaron que en algunos meses de este segundo año de vida ningún sonido vocálico fuera válido para el análisis. Esto último sucedió en los periodos comprendidos entre los 1;2 y 1;3 años, y entre los 1;3 y 1;4 años. En otros meses, tan solo una vocal fue seleccionada (entre los 1;1 y 1;2 años, los 1;4 y 1;5 años, y entre los 1;10 y 1;11 años). Por ello y tal y como procedimos en el análisis acústico de la niña,

agruparemos las vocales seleccionadas en el niño en aquellas emitidas en los seis primeros meses del periodo y las analizadas en los seis segundos.

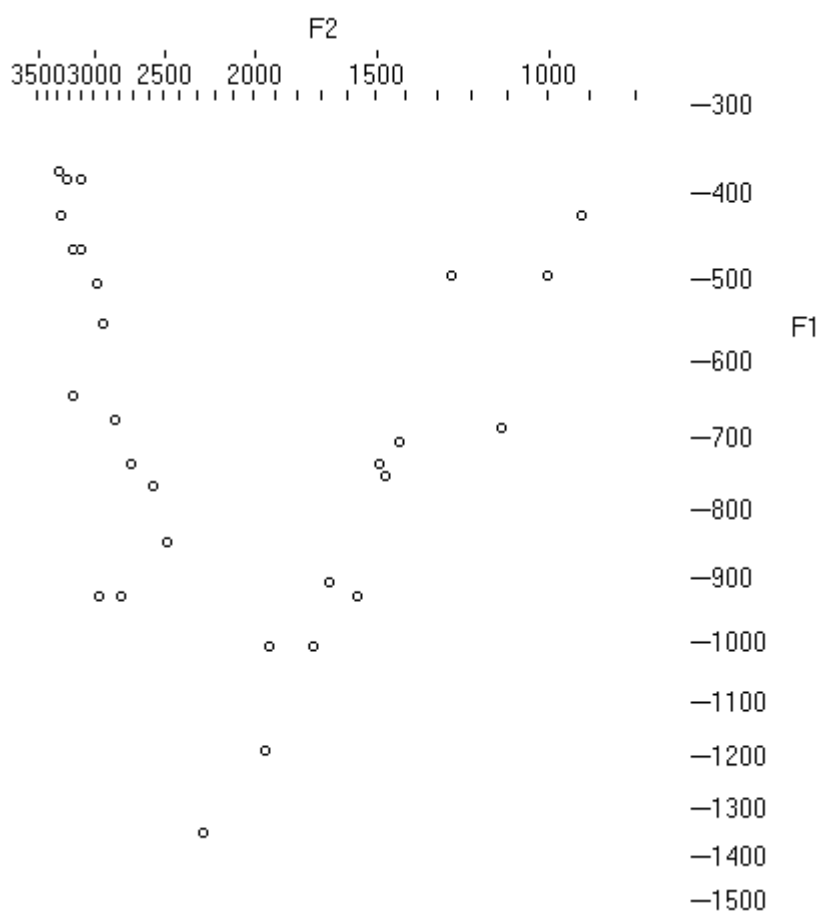
Seis vocales fueron analizadas en el primer periodo de este segundo año de vida tal y como refleja la carta de formantes inferior (gráfico 6.21). El valor mínimo del F1 fue de 380 Hz y el máximo de 870 Hz habiendo una diferencia de 490 Hz entre ambos. En cuanto al F2, este osciló entre los 2800 Hz, valor mínimo, y los 3420 Hz, valor máximo (distancia de 620 Hz). Como podemos observar, todas las señales se agrupan en la parte izquierda del gráfico tratándose por tanto de vocales anteriores de diferentes alturas, esto es, altas y medias.

**Gráfico 6.21. Carta de formantes del niño 1;0-1;6 años.**



En los seis últimos meses del segundo año vida veintiocho vocales fueron seleccionadas (gráfico 6.22). El F1 presentó un valor mínimo de 370 Hz y uno máximo de 1340 Hz (diferencia de 970 Hz). Respecto al F2, el valor más bajo fue de 930 Hz y el máximo de 3320 Hz (diferencia de 2390 Hz). Las vocales se distribuyeron por todos los ángulos de la carta de formantes superando no obstante las vocales anteriores, altas y medias, a las vocales posteriores (también altas y medias), y estas últimas a las vocales centrales.

**Gráfico 6.22. Carta de formantes del niño 1;6-2;0 años.**

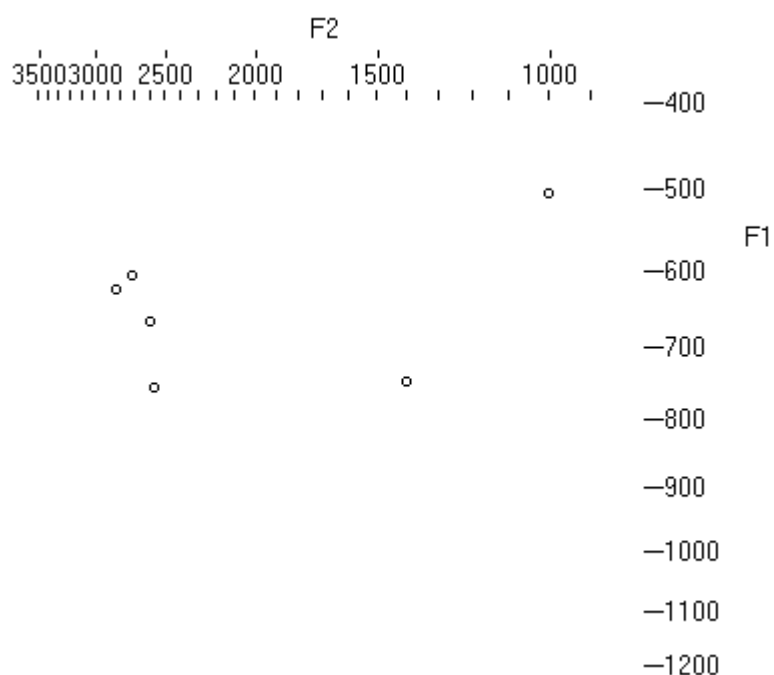


### **6.1.2.3. Tercer año de vida.**

Debido a las circunstancias explicadas en el capítulo de la metodología (punto 5.1.3) el número de vocales seleccionadas y analizadas en las emisiones del niño en el tercer año de vida fue menor que el de los años anteriores. Los meses en los que ninguna vocal modal pudo extraerse fueron los comprendidos entre los 2;2 y 2;3 años, 2;3 y 2;4 años, 2;5 y 2;6 años, y los 2;11 y 3;0 años. Al igual que en el año anterior, también encontramos meses en los que solo un sonido vocálico pudo ser seleccionado: entre los 2;0 y 2;1 años, y entre los 2;1 y 2;2 años. Estas circunstancias nos han llevado una vez más a agrupar las vocales en un primer periodo temporal que comprende los seis primeros meses de este tercer año de vida y en un segundo periodo temporal que abarca los seis últimos.

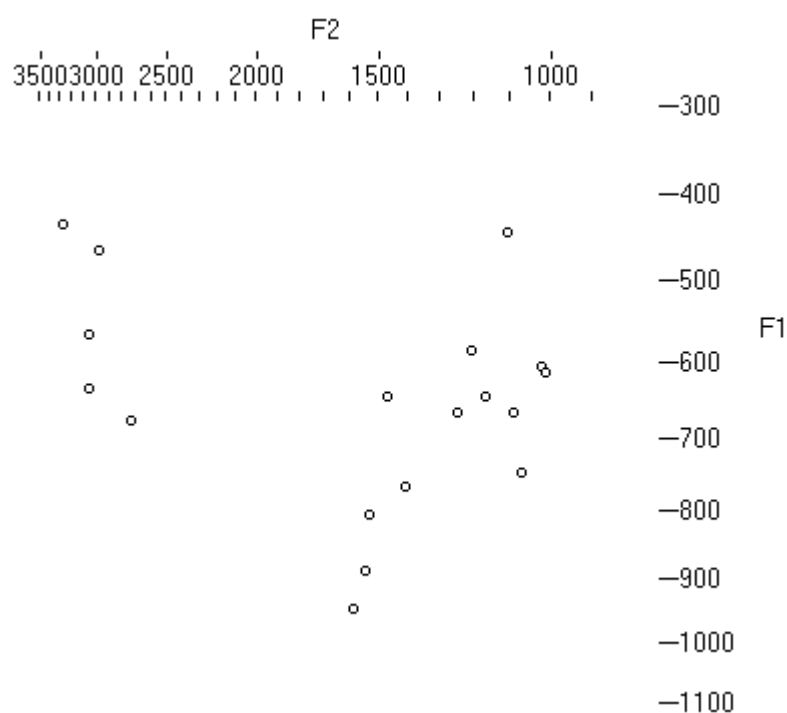
Tan solo seis vocales fueron extraídas y analizadas en los seis primeros meses del tercer año de vida del niño (gráfico 6.23). La distancia entre el valor mínimo y máximo del primer formante fue de 250 Hz, siendo el valor más bajo de 500 Hz y el más alto de 750 Hz. En el F2 la diferencia fue de 1850 Hz registrando el segundo formante un valor mínimo de 1010 Hz y un máximo de 2860 Hz. Tal y como muestra la carta de formantes las vocales anteriores, de altura media, superan a las vocales posteriores entre las que podemos observar una vocal media y otra alta.

**Gráfico 6.23. Carta de formantes del niño 2;0-2;6 años.**



Un total de dieciocho vocales fueron analizadas en el último periodo temporal del tercer año de vida (gráfico 6.24). El primer formante osciló entre los 430 Hz, valor mínimo, y los 940 Hz de valor máximo (diferencia de 510 Hz). El F2 por otro lado, se movió entre los 1030 Hz y los 3300 Hz siendo la diferencia entre el valor más bajo y el más alto de 2270 Hz. Respecto a la ubicación de las señales en la carta de formantes, observamos como por primera vez el número de vocales situadas en el extremo derecho del gráfico, esto es vocales posteriores, supera en casi el doble a las ubicadas en la parte izquierda, vocales anteriores. En cuanto a la altura, hallamos vocales anteriores y posteriores tanto altas como medias predominando más estas últimas.

**Gráfico 6.24. Carta de formantes del niño 2;6-3;0 años.**



## **6.2 Resultados del test de percepción.**

A continuación presentamos los sonidos vocálicos a los que al menos cuatro de los cinco oyentes atribuyeron el mismo timbre después de haber realizado el test de percepción. Estas vocales incluyeron tanto las vocales modales que cumplían los criterios de selección, como las vocales nasales que se incorporaron para ver qué tipo de timbre predominaba a lo largo de los tres años analizados. Debemos subrayar que el número de sonidos (tanto modales como nasales) presentes en cada uno de los meses de los tres años de vida analizados es menor que la cantidad de señales reales que fueron extraídas en cada uno de ellos para el análisis acústico debido a que muchas de estas señales fueron descartadas por la no coincidencia de timbre entre los sujetos que realizaron el test de percepción.

El número de vocales totales a las que cuatro de los cinco transcripores atribuyeron el mismo timbre vocálico fue de 836 de las cuales 501 pertenecieron a la niña y 335 al niño. En cuanto a la niña y de las 501 muestras, 155 fueron modales y 346 nasales. Dentro de las vocales modales, 85 vocales fueron identificadas con el mismo timbre vocálico en el primer año de vida, 25 en el segundo y 45 en el tercero. En las nasales, los datos por años fueron de 58, 126 y 162 respectivamente. Respecto al niño, de las 335 vocales identificadas con el mismo timbre por

cuatro transcritores, 99 fueron modales y 236 nasales. Dentro de las primeras, 45 fueron reconocidas con el mismo timbre vocálico en el primer año, 32 en el segundo y 22 en el tercero. Los sonidos nasales por su parte se distribuyeron de la siguiente forma: 49 fueron identificados de manera común en el primer año de vida, 109 en el segundo y 78 en el tercero.

En cuanto a la presentación de los datos, expondremos en primer lugar y con ayuda de una tabla las vocales modales a las que los transcritores atribuyeron el mismo timbre vocálico en cada uno de los meses así como el número de veces que cada vocal se repitió dentro del periodo. A continuación exhibiremos en unas tablas similares los sonidos nasales. El incluir este último tipo de sonidos en el test de percepción, como ya indicamos en el apartado de la metodología, nos permitirá establecer con una mayor precisión, sobre todo en aquellos meses en los que el número de señales válidas es muy reducido, la cualidad vocálica más frecuente en los distintos periodos analizados.

### **6.2.1. Timbres vocálicos identificados en las emisiones de la niña.**

#### ***6.2.1.1. Primer año de vida.***

De los ciento seis sonidos modales extraídos de las emisiones de la niña durante el primer año de vida, cuatro de los cinco oyentes coincidieron en la identificación de ochenta y cinco. Tal y como se desprende de la tabla posterior (tabla 6.1), las vocales anteriores y la vocal central baja fueron las más frecuentes a lo largo de todo el periodo. La vocal anterior semialta [e] fue la que más veces se repitió a lo largo de todo el año excepto entre los 0;5 y 0;6 años, mes en el que se ausentó (44,7%). Le siguió en frecuencia la vocal central baja [a] presente en todos los meses y repetida hasta en veinte ocasiones (23,5%). Otra vocal anterior, en este caso alta, se encontró entre las más comunes y pese a su ausencia en determinados meses de este primer año esta vocal apareció un total de trece veces lo que supuso un 15,2% de la muestra. Muy cerca se encontró la vocal posterior semialta [o] con una presencia del 14,1% y ausente tan solo entre los 0;9 y 0;10 y entre los 0;11 y 0;1 años. Respecto a la [u], llama la atención su ausencia antes de los 0;7 años (omisión que comentaremos en el apartado de la discusión) frente al resto de vocales del español que estuvieron presentes en el repertorio de la niña a los 0;6 años. Después de esta edad, tan solo apareció una vez entre los 0;6 y 0;7 años y otra vez entre los 0;7 y 0;8 años representando únicamente el 2,3% de las vocales extraídas a lo largo de este primer año.

**Tabla 6.1. Sonidos vocálicos modales identificados en la niña en el primer año de vida (0;4-1;0 año).**

Sonidos vocálicos modales niña. Primer año de vida.					
Edad	a	e	i	o	u
0;4-0;5	5	3		1	
0;5-0;6	1		1	1	
0;6-0;7	2	11		3	1
0;7-0;8	2	7	1	3	1
0;8-0;9	3	9	9	3	
0;9-0;10	3	1	1		
0;10-0;11	1	4	1	1	
0;11-1;0	3	3			
Total	20	38	13	12	2

En cuanto a los sonidos nasales (tabla 6.2), cuatro transcripores adjudicaron el mismo timbre vocálico a cincuenta y ocho señales de las noventa y cuatro extraídas. En esta ocasión el orden se invirtió y la vocal central baja [ã] superó con creces a la vocal anterior semialta [ẽ] (51,7% de [ã] frente al 31% de [ẽ]). La vocal posterior semialta [õ] apareció diez veces (17,2%) y fue llamativa la ausencia de las vocales altas, tanto de la anterior como de la posterior.

**Tabla 6.2. Sonidos vocálicos nasales identificados en la niña en el primer año de vida (0;4-1;0 año).**

Sonidos vocálicos nasales niña. Primer año de vida.			
Edad	ã	ẽ	õ
0;4-0;5			
0;5-0;6			1
0;6-0;7	2		
0;7-0;8	2	1	4
0;8-0;9	7	6	1
0;9-0;10	6	4	1
0;10-0;11	8	5	2
0;11-1;0	5	2	1
Total	30	18	10



### 6.2.1.2. Segundo año de vida.

De treinta vocales modales extraídas en el segundo año de vida de la niña, veinticinco fueron identificadas con el mismo timbre (tabla 6.3). Al igual que en el primer año de vida, la vocal anterior semialta [e] fue la vocal que más se repitió apareciendo hasta en ocho ocasiones y representando así el 32% de la totalidad. Sin embargo, en este segundo periodo la vocal central baja no ocupó la segunda posición ni tampoco la tercera ya que tan solo fue identificada en cuatro ocasiones (16%) siendo superada por lo tanto por la vocal anterior alta [i] que apareció siete veces (28%) y por la vocal posterior semialta [o] que apareció cinco veces y representó el 20% de la muestra total. Finalmente, la vocal posterior alta [u] apareció una vez lo que supuso un 4% de la muestra.

A partir de estos datos se desprende que dentro de los sonidos seleccionados como comunes en este segundo año de vida pareció mantenerse la tendencia observada en el primer año de vida según la cual las vocales anteriores, en concreto la anterior semialta [e] y en este periodo también la anterior alta [i], se repitieron con mayor frecuencia. No obstante, la vocal central baja pasó a ocupar la cuarta posición siendo superada por la vocal posterior semialta [o] si bien mínimamente ya que la diferencia fue tan solo de una muestra. La vocal posterior alta siguió estando en la última posición en cuanto a frecuencia.

**Tabla 6.3. Sonidos vocálicos modales identificados en la niña en el segundo año de vida (1;0-2;0 años).**

Sonidos vocálicos modales niña. Segundo año de vida.					
Edad	a	e	i	o	u
1;0-1;1					
1;1-1;2					
1;2-1;3		1			
1;3-1;4					
1;4-1;5					
1;5-1;6	2	2		1	
1;6-1;7		1			
1;7-1;8	1		1		
1;8-1;9	1	1	1	2	
1;9-1;10			2	1	1
1;10-1;11		2	2	1	

1;11-2;0		1	1		
Total	4	8	7	5	1

Ciento veintiséis sonidos vocálicos de los ciento noventa y cuatro nasales extraídos a lo largo del segundo año de vida fueron reconocidos con el mismo timbre vocálico por cuatro de los cinco oyentes (tabla 6.4). Una vez más, el número de vocales anteriores superó al resto de timbres vocálicos identificados en este periodo representando un 60,2% de la muestra total. La vocal anterior semialta [ẽ] fue de nuevo la más presente (51,5%) seguida de la vocal central baja [ã] que supuso un 26,9% de la muestra. La siguiente vocal más identificada fue la vocal posterior semialta [õ] (10,3%) que superó, aunque tan solo en dos muestras, a la vocal anterior alta [ĩ] (8,7%). La vocal posterior alta fue identificada en tres ocasiones (2,3% del total).

**Tabla 6.4. Sonidos vocálicos nasales identificados en la niña en el segundo año de vida (1;0-2;0 años).**

Sonidos vocálicos nasales niña. Segundo año de vida.					
Edad	ã	ẽ	ĩ	õ	ũ
1;0-1;1	8				1
1;1-1;2	2	2		3	
1;2-1;3	3	1		3	
1;3-1;4	1	3	2		
1;4-1;5	4	7	2	1	
1;5-1;6		6		1	1
1;6-1;7	7	4	1		1
1;7-1;8		14	3	2	
1;8-1;9		8	1	1	
1;9-1;10	4	8	1	1	
1;10-1;11	2	7	1	1	
1;11-2;0	3	5			
Total	34	65	11	13	3

### 6.2.1.3. Tercer año de vida.

Casi la totalidad de las vocales modales extraídas a lo largo del tercer año de vida fueron identificadas con el mismo timbre por cuatro de los cinco oyentes (cuarenta y cinco de cincuenta y dos). Como puede observarse en la tabla 6.5, la preferencia por los sonidos vocálicos anteriores volvió a manifestarse durante este tercer año de vida suponiendo la suma de estos el 53,2% del total. En cuanto a vocales aisladas, la vocal anterior semialta [e] fue la más identificada por los oyentes (28,8%) y a esta le siguió la vocal posterior semialta [o] que aventajó en una muestra a la vocal anterior alta [i] (26,6% frente al 24,4%). Por primera vez en los tres periodos analizados, la vocal posterior alta [u] superó con creces a la vocal central baja [a] siendo los porcentajes del 15,5% y del 4,4% respectivamente.

**Tabla 6.5. Sonidos vocálicos modales identificados en la niña en el tercer año de vida (2;0-3;0 años).**

Sonidos vocálicos modales niña. Tercer año de vida.					
Edad	a	e	i	o	u
2;0-2;1					
2;1-2;2		2	1		
2;2-2;3		2	2	3	
2;3-2;4	1		1	1	
2;4-2;5		2		2	
2;5-2;6		4		2	4
2;6-2;7	1		4	1	1
2;7-2;8		1	1		
2;8-2;9		1		1	1
2;9-2;10			1	1	1
2;10-2;11					
2;11-3;0		1	1	1	
Total	2	13	11	12	7

En cuanto a las vocales nasales extraídas en las emisiones de la niña en el mismo periodo (tabla 6.6), observamos que los transcritores coincidieron en ciento sesenta y dos vocales de las doscientas una seleccionadas en este tercer año de vida. La inclinación hacia las vocales anteriores persistió también en las señales nasales de este periodo representando estas más de la mitad de la muestra total (el 56,7%). Continuando con el patrón del año anterior, la

vocal anterior semialta [ẽ] fue la más identificada (41,9%), seguida de la vocal central baja [ã] (24,6%) y de la vocal posterior semialta [õ] que superó como en el segundo año de vida a la vocal anterior alta [ĩ] si bien tan solo en dos muestras (16% frente al 14,8%). Cuatro muestras de la vocal posterior alta fueron identificadas en este último año de análisis (2,4%).

**Tabla 6.6. Sonidos vocálicos nasales identificados en la niña en el tercer año de vida (2;0-3;0 años).**

Sonidos vocálicos nasales niña. Tercer año de vida.					
Edad	ã	ẽ	ĩ	õ	ũ
2;0-2;1	3	4	2		
2;1-2;2	2	5			
2;2-2;3	5	4	3	1	
2;3-2;4	4	7		3	
2;4-2;5	3	11	1	8	
2;5-2;6	10	17	2	4	
2;6-2;7		3	5	1	1
2;7-2;8	6	6	4	6	1
2;8-2;9	2	5		2	2
2;9-2;10	2		3		
2;10-2;11	2	1	1		
2;11-3;0	1	5	3	1	
Total	40	68	24	26	4

## 6.2.2. Timbres vocálicos identificados en las emisiones del niño.

### 6.2.2.1. Primer año de vida.

De cincuenta y cuatro vocales modales extraídas de las emisiones del niño durante el primer año de vida, los jueces coincidieron en la identificación de cuarenta y cinco (tabla 6.7). Como ocurrió con los sonidos de la niña emitidos en el primer periodo analizado, se trató en su mayoría de vocales anteriores semialtas (un 60%) seguidas de las anteriores altas (un 22,2%) y de las centrales bajas que representaron el 8,8% del total. Coincidiendo también con los datos de la niña, las vocales posteriores semialtas ocuparon la penúltima posición representando el 6,6% de la muestra y adelantando en número las posteriores altas que supusieron el 2,2%.

**Tabla 6.7. Sonidos vocálicos modales identificados en el niño en el primer año de vida (0;4-1;0 año).**

Sonidos vocálicos modales niño. Primer año de vida.					
Edad	a	e	i	o	u
0;4-0;5	1	3	1		
0;5-0;6		4		2	
0;6-0;7	1	3			
0;7-0;8		4	1	1	1
0;8-0;9	2	5			
0;9-0;10		4			
0;10-0;11		2	3		
0;11-1;0		2	5		
Total	4	27	10	3	1

En cuanto a los sonidos nasales, los transcripores atribuyeron el mismo timbre a cuarenta y nueve de los setenta y cuatro sonidos extraídos en este periodo (tabla 6.8). La predilección por la vocal anterior semialta se mantuvo también en los sonidos nasales puesto que la vocal [ẽ] representó el 65,3% de la totalidad. La vocal central baja fue la segunda vocal que más se repitió constituyendo un 20,4% del total y superando en el doble a la vocal anterior alta [ĩ] que estuvo presente en un 10,2%. En cuanto al ángulo posterior, la vocal posterior semialta [õ] apareció en dos ocasiones (4%) mientras que la posterior alta [ũ] no fue identificada en ninguno de los meses de este primer año.

**Tabla 6.8. Sonidos vocálicos nasales identificados en el niño en el primer año de vida (0;4-1;0 año).**

Sonidos vocálicos nasales niño. Primer año de vida.				
Edad	ã	ẽ	ĩ	õ
0;4-0;5	1	1	1	
0;5-0;6	2	4		1
0;6-0;7	1	3		
0;7-0;8		8	2	1
0;8-0;9		5	1	
0;9-0;10	2	4		

0;10-0;11	3	2	1	
0;11-1;0	1	5		
Total	10	32	5	2

### 6.2.2.2. Segundo año de vida.

En el segundo año de vida treinta y dos vocales modales de las treinta y cuatro extraídas fueron identificadas como la misma vocal por los transcriptores (tabla 6.9). Aunque la inclinación por las vocales anteriores se mantuvo en este periodo, la vocal anterior alta [i] apareció casi el doble de veces que la anterior semialta [e] constituyendo aquella el 40,6% de la muestra frente al 21,8% de esta última. Otra diferencia hallada con respecto al primer periodo analizado es que por primera vez las vocales posteriores, tanto la semialta [o] como la alta [u], sobrepasaron a la vocal central baja [a] aun siendo mínima la diferencia entre estas (la vocal posterior semialta representó un 15,6%, la posterior alta un 12,5% y la central baja un 9,3%).

**Tabla 6.9. Sonidos vocálicos modales identificados en el niño en el segundo año de vida (1;0-2;0 años).**

Sonidos vocálicos modales niño. Segundo año de vida.					
Edad	a	e	i	o	u
1;0-1;1		1	2		
1;1-1;2			1		
1;2-1;3					
1;3-1;4					
1;4-1;5			1		
1;5-1;6		1			
1;6-1;7		1	3		
1;7-1;8	2	3	2		1
1;8-1;9	1				1
1;9-1;10		1	2	3	1
1;10-1;11				1	
1;11-2;0			2	1	1
Total	3	7	13	5	4

Respecto a los sonidos vocálicos nasales (tabla 6.10), de ciento setenta y cinco vocales extraídas en este segundo año de vida los transcripores coincidieron en la identificación de ciento nueve. Resulta llamativo el hecho de que por primera vez en los dos periodos analizados (primer y segundo de vida) la vocal central baja [ã] aventaje a la anterior semialta [ẽ] representando la primera el 33,9% de la muestra y el 24,7% la segunda. Incluso si optamos por la suma del número de apariciones de todas las vocales anteriores identificadas, vemos que aunque el número de muestras total supera al de la vocal central baja, este lo hace en tan solo cinco vocales (42 frente a 37). Por otro lado, es también novedosa la circunstancia de que solo entre los 1;7 y los 1;8 años la cantidad de vocales centrales bajas fuera muy superior a la encontrada en el resto de meses de este periodo y del periodo anterior, tanto en los sonidos modales como en los nasales. Asimismo, es la primera vez que la [ã] aparece en todos los meses de este periodo.

En cuanto a las vocales anteriores y a diferencia de lo que ocurría con las vocales modales, fue la vocal anterior semialta [ẽ] la que mayor número de veces se repitió constituyendo un 24,7% de la muestra frente al 13,7% de la anterior alta [ĩ]. Respecto a las vocales posteriores, su presencia aumenta significativamente en los sonidos nasales de este periodo, especialmente la presencia de la vocal posterior semialta [õ] cuyo número de muestras llega a superar por primera vez en los dos periodos analizados y tanto en los sonidos vocálicos modales como en los nasales, al número de muestras de la vocal anterior alta [ĩ] (veintidós apariciones de la vocal posterior semialta frente a las quince de la vocal anterior alta). Finalmente, la vocal posterior alta [ũ] también registró un notable ascenso apareciendo hasta en ocho ocasiones y representando un 7,3% de la muestra total.

**Tabla 6.10. Sonidos vocálicos nasales identificados en el niño en el segundo año de vida (1;0-2;0 años).**

Sonidos vocálicos nasales niño. Segundo año de vida.					
Edad	ã	ẽ	ĩ	õ	ũ
1;0-1;1	1	3			
1;1-1;2	4	3	1	1	1
1;2-1;3	2	6	3		
1;3-1;4	2	2	3		
1;4-1;5	2		1	1	
1;5-1;6	3	6	2	3	3
1;6-1;7	4	2	2	3	2

1;7-1;8	11	3	2	4	
1;8-1;9	4			1	1
1;9-1;10	2	1		5	1
1;10-1;11	1			2	
1;11-2;0	1	1	1	2	
Total	37	27	15	22	8

### 6.2.2.3. Tercer año de vida.

Prácticamente la totalidad de las vocales modales extraídas en el tercer año de vida de las emisiones del niño fueron identificadas con el mismo timbre por cuatro de los cinco oyentes. (veintidós de veinticuatro). Sin embargo, la elevada coincidencia en la transcripción no mitigó el tan reducido número de señales válidas extraídas en este periodo resultando estas quizás algo escasas para determinar el timbre o timbres vocálicos más frecuentes del mismo. Entre las razones de esta insuficiencia, expuestas ya en el apartado 5.1.3, pueden subrayarse dos: la primera se debe a que la cantidad meses en los que pudieron realizarse grabaciones durante este tercer año de vida se redujo a diez; la segunda residió en el hecho de que en solo en ocho de los diez meses analizados se encontraron sonidos válidos para el análisis formántico, esto es sonidos modales, extrayéndose además en muchos de ellos una sola señal.

Si nos detenemos en la tabla posterior (tabla 6.11) dos circunstancias saltan a la vista: por un lado la ausencia de la vocal [a] que no fue identificada por ninguno de los oyentes en el tercer año de vida; y por otro el hecho de que por primera vez en el tiempo global que comprendió el estudio (los tres primeros años de vida) el número de vocales posteriores fue superior al de las vocales anteriores representando las primeras el 59% y las segundas el 40,8% de la totalidad. En el ángulo anterior encontramos que la vocal anterior semialta [e] aventajó en una muestra a la vocal anterior alta [i] (22,7% frente al 18,1%, respectivamente). En cuanto a la zona posterior y si bien una vez más la vocal posterior semialta [o] superó a la posterior alta [u], en esta ocasión tan solo lo hizo en una muestra representando la posterior semialta el 31,8% y la posterior alta el 27,2% del total.



**Tabla 6.11. Sonidos vocálicos modales identificados en el niño en el tercer año de vida (2;0-3;0 años).**

Sonidos vocálicos modales niño. Tercer año de vida.				
Edad	e	i	o	u
2;0-2;1				1
2;1-2;2	1			
2;2-2;3				
2;3-2;4				
2;4-2;5	3		1	
2;5-2;6				
2;6-2;7	1	2	1	
2;7-2;8			2	1
2;8-2;9		1	2	2
2;9-2;10			1	2
2;10-2;11		1		
2;11-3;0				
Total	5	4	7	6

Para finalizar este apartado mostraremos las vocales nasales a las que los transcripores atribuyeron la misma cualidad vocálica a lo largo del tercer año de vida y cuya cifra es de setenta y ocho de las ochenta y seis señales extraídas (tabla 6.12). Tal y como se desprende de la tabla, parece mantenerse la tendencia general encontrada hasta ahora en las emisiones del niño en cuanto a frecuencia de aparición vocálica. De este modo, las vocales anteriores fueron las más representadas seguidas de la central baja a la que le sucedieron las posteriores. Una vez más fue la vocal anterior semialta [ɛ̃] la que más veces se repitió alcanzando el 34,6% del conjunto de las señales de este mes. Le siguió la vocal central baja [ã] con un 26,9% y a esta le sucedió la posterior semialta [õ] que aventajó en una muestra a la anterior alta [ĩ] (17,9% y 16,6%, respectivamente). La posterior alta [ũ] apareció en tres ocasiones lo que supuso un 3,8% del total.

**Tabla 6.12. Sonidos vocálicos nasales identificados en el niño en el tercer año de vida (2;0-3;0 años).**

Sonidos vocálicos nasales niño. Tercer año de vida.					
Edad	ã	ẽ	ĩ	õ	ũ
2;0-2;1	1	3	1	1	
2;1-2;2	2	3			
2;2-2;3					
2;3-2;4					
2;4-2;5		2			
2;5-2;6					
2;6-2;7	4	1	4	2	
2;7-2;8	3	5	1	6	
2;8-2;9	5	3	2	2	2
2;9-2;10	5	7	3		
2;10-2;11		1	1		
2;11-3;0	1	2	1	3	1
Total	21	27	13	14	3

### **6.3. Comparación de los resultados obtenidos en el análisis acústico y perceptivo en la niña y el niño.**

#### **6.3.1. Comparación de los resultados obtenidos tras el análisis acústico.**

En el primer periodo de tiempo analizado, el que transcurrió entre los 0;4 y el 1;0 año, ambos niños mostraron una clara preferencia por las vocales situadas en los ángulos anterior y central del triángulo vocálico tal y como muestran los gráficos expuestos en el apartado 6.1 que representan las cartas de formantes. Dichos valores reflejaron que el número de vocales ubicadas en la parte derecha del gráfico, parte que se corresponde con las vocales articuladas en la zona posterior de la cavidad bucal, fue mucho más reducido en la mayoría de los periodos analizados que el de las señales distribuidas en la parte central e izquierda del mismo, sonidos que se identifican con aquellas vocales articuladas en la parte central y anterior de la boca. En cuanto a la altura de estos sonidos vocálicos, las vocales medias y bajas predominaron sobre las altas en ambos lados del triángulo (es decir, tanto en la zona anterior como en la posterior del

gráfico /cavidad bucal). No obstante, y a pesar de que ambos niños se inclinaron por este tipo de vocales, las preferencias individuales afloraron en determinados meses de forma que las vocales tendieron a agruparse o esparcirse en determinadas zonas de las cartas de formantes según el sujeto y el mes concreto analizado como veremos a continuación.

En el primer mes analizado, el análisis acústico de las señales de la niña situó a la mayoría de las vocales en la zona central y anterior de la carta de formantes, superponiéndose muchas de ellas, quedando la zona posterior de esta carta de formantes menos representada. En el mismo periodo de tiempo la mayor parte de las señales analizadas en el niño presentaron un segundo formante muy elevado que se tradujo en un predominio de la zona anterior de la cavidad bucal con una tímida representación de la central y con una clara ausencia de la posterior.

A la edad de entre 0;5 y 0;6 años la preferencia de la niña se invirtió y frente a la abundancia de señales con un elevado F2 del periodo anterior, en este mes predominaron los sonidos centrales y posteriores con un F2 más bajo. Las señales del niño se distribuyeron, sin embargo, de manera opuesta en la carta de formantes. De esta forma, la mayoría de las vocales se agrupó en el ángulo izquierdo del triángulo aunque un sonido, con un alto valor en el F1 y un valor medio en el F2, apareció en el centro del mismo. Debemos por un lado subrayar la aparición por primera vez en este primer periodo de las vocales posteriores que se reunieron en la parte derecha del gráfico, y debemos recordar por otro que aunque aparecen por primera vez en el niño en este mes, ya estaban presentes en la niña en el mes anterior. La incorporación de estas últimas vocales en el repertorio del niño permite dibujar el triángulo vocálico del español si bien con forma de triángulo escaleno invertido.

Por primera vez entre los 0;6 y los 0;7 años, la distribución de los valores de los formantes de las vocales de la niña delimitó, aunque de manera tímida, un esbozo del triángulo vocálico del español. Las cartas de formantes de ambos niños mostraron una clara preferencia por las vocales anteriores medias y altas. Asimismo, en ambos gráficos estuvieron presentes los sonidos centrales bajos aunque la cantidad fue superior en el caso de la niña. La única diferencia en las cartas de formantes de ambos niños parece ser la ausencia de vocales posteriores en la carta de formantes del niño.

Entre los 0;7 y los 0;8 años, las cartas de formantes de ambos niños volvieron a mostrar una marcada predilección por los sonidos anteriores medios con una menor representación de los altos. También en ambos casos encontramos señales en la parte derecha del gráfico que apuntaron a vocales posteriores de distintas alturas. Cabe subrayar la falta de vocales situadas en

el centro de la carta formántica del niño mientras que estas sí estuvieron presentes en la de la niña.

La ausencia de señales de bajo F1 y bajo F2 situadas en el ángulo derecho del gráfico del niño es la única circunstancia que diferencia las cartas de formantes de ambos niños entre los 0;8 y 0;9 años. Los sonidos anteriores y centrales estuvieron presentes en los dos sujetos siendo los primeros los más numerosos. En cuanto a estos últimos, la carta de formantes de la niña muestra una mayor cantidad de vocales anteriores altas que la carta de formantes del niño.

La propensión hacia las vocales anteriores medias y hacia las centrales bajas volvió a evidenciarse entre los 0;9 y los 0;10 años en las cartas de formantes de ambos niños si bien la proporción de estas varió siendo las primeras más frecuentes en el niño y las últimas más repetidas en la niña. Asimismo, la falta de vocales posteriores en las cartas de formantes de ambos sujetos en este periodo continuó asemejando las emisiones de los dos niños.

Una vez más las vocales anteriores (medias en su mayoría) y las vocales centrales bajas fueron las que estuvieron más presentes en los gráficos de ambos niños entre los 0;10 y los 0;11 años. Además, parece que dichas vocales fueron las únicas que aparecieron en la carta de formantes del niño mientras que la de la niña, y tal y como ocurría en otros periodos, sí que pareció reflejar sonidos posteriores.

En el último mes de este primer periodo analizado (0;11-1;0 año) las señales situadas en el ángulo izquierdo del gráfico volvieron a estar presentes en ambos sujetos siendo estas las únicas encontradas en el caso del niño. En la carta de formantes de la niña se adivinaron también vocales centrales bajas. Por otro lado, parece que ninguno de los gráficos, o al menos no el del niño, recogió vocales posteriores.

La preferencia por las vocales anteriores se mantuvo en el segundo año de vida en ambos sujetos aunque a diferencia del periodo anterior el número de sonidos posteriores aumentó en ambos niños.

En el periodo comprendido entre el año y el año y medio (1;0-1;6 años), podemos observar como todas las vocales del niño se situaron a la izquierda de la carta de formantes mientras que las de la niña se distribuyeron por todos los extremos del gráfico pudiéndose identificar por tanto vocales anteriores, centrales y posteriores de distintas alturas.

Entre los 1;6 y los 2;0 años de vida las cartas de formantes de ambos niños presentaron muestras en todos sus ángulos por lo que podemos deducir que todos los timbres vocálicos

estuvieron presentes. Por otro lado, el predominio de las vocales anteriores sobre el resto de timbres vocálicos es también común en ambos niños en este periodo.

En la primera mitad del tercer año de vida, esto es entre los 2;0 y los 2;6 años, llama la atención la escasez de vocales analizadas en las emisiones del niño en comparación con la cantidad examinada en la niña. En cuanto a la ubicación de las mismas, la carta de formantes del niño solo presentó vocales en el extremo izquierdo y en el derecho evidenciándose así una total ausencia de vocales centrales. En el caso de la niña, las señales se distribuyeron por todos los ángulos de la carta y fue llamativo el notable aumento de las vocales posteriores en este periodo.

En el último periodo temporal en el que se agruparon las señales de cada niño y que fue de los 2;6 a los 3;0 años, encontramos en ambos sujetos una escasez de vocales centrales y un notorio aumento de señales posteriores, señales que en el caso del niño superaron a las vocales anteriores tal y como muestra la carta de formantes.

### **6.3.2. Comparación de los resultados obtenidos tras el test de percepción.**

A lo largo de los tres años de vida estudiados en ambos sujetos y tanto en los sonidos modales como en los nasales extraídos en uno y otro sujeto, las vocales que fueron identificadas un mayor número de veces por los oyentes fueron las vocales anteriores, en particular, la anterior semialta [e]. Asimismo, el test de percepción nos ha permitido observar que la vocal que registró el menor número de muestras en ambos niños y en ambos tipos de señales (modales y nasales) fue la vocal posterior alta [u]. No obstante, pueden destacarse algunas particularidades según el tipo de señal, el sujeto y el periodo de tiempo concreto analizado. Comenzaremos exponiendo las similitudes y diferencias que arrojó el análisis perceptivo realizado sobre las emisiones modales de ambos niños y finalizaremos comparando los resultados hallados en las señales nasales.

En el primer periodo observado y dentro de la categoría de “sonido modal”, las vocales anteriores y la vocal central baja fueron las más frecuentes en ambos sujetos. La primera posición la ocupó la vocal anterior semialta [e] que fue la más común tanto en el niño como en la niña. La distinción apareció en torno a la segunda vocal más frecuente. En el caso de la niña la vocal central baja [a] ocupó el segundo puesto en el orden frecuencial manifestando aún una clara ventaja sobre el resto de timbres vocálicos. En el niño sin embargo, dicha vocal fue superada en seis muestras por la vocal anterior alta [i], vocal que fue identificada casi el mismo número de veces en ambos niños (diez veces en el niño y trece en la niña). La cantidad de

sonidos posteriores fue la menor en ambos sujetos siendo entre estos los posteriores altos los menos comunes.

La inclinación por los sonidos vocálicos anteriores se mantuvo en ambos sujetos en el segundo año de vida aunque preferencias individuales asomaron en cuanto a la altura vocálica de los mismos. De esta forma, si en la niña la vocal anterior semialta [e] fue la más identificada por los transcriptores seguida de la vocal anterior alta [i], en el niño el orden se invirtió y fue la vocal anterior alta la más presente seguida de la anterior semialta. El tercer puesto fue ocupado en ambos sujetos por la vocal posterior semialta [o] que presentó en ambos niños la misma frecuencia de aparición. A esta vocal le siguió en las emisiones de la niña la vocal central baja [a] quedando la posterior alta [u] en último lugar. En el caso del niño el orden volvió a invertirse si bien por muy poco ya que aunque la vocal posterior alta [u] superó a la vocal central baja [a], esta tan solo lo hizo en una muestra.

En cuanto al tercer año de vida, subrayamos el hecho de que por primera vez los sonidos posteriores aventajaron a los anteriores en las emisiones del niño y prácticamente presentaron el mismo número de muestras en el caso de la niña (diecinueve frente a veinticuatro). Llama la atención que la vocal posterior semialta [o] sea la más frecuente en las emisiones del niño, y que a esta le siga la vocal posterior alta [u] quedando en tercer lugar la vocal anterior semialta [e] y en último la anterior alta [i]. En el caso de la niña, aunque la vocal anterior semialta [e] ocupa el primer puesto, esta tan solo supera en una muestra a la vocal posterior semialta [o] que aventaja también en una muestra a la vocal anterior alta [i]. El penúltimo lugar lo ocupa en ambos sujetos la vocal posterior alta [u] que al igual que la posterior semialta [o] manifiesta un notable aumento en este periodo con respecto a los años anteriores y que supera en ambos niños a la central baja [a] (vocal que en el caso del niño no fue identificada).

En lo que respecta a las señales nasales, los sonidos vocálicos anteriores volvieron a ser los preferidos por ambos niños en el tiempo que comprendió el estudio. No obstante y a diferencia de lo que ocurría con los sonidos modales, estos compartieron en determinados periodos el primer puesto en cuanto a frecuencia de aparición con la vocal central baja [ã], cuyas muestras fueron superiores en todos los meses a las de la vocal posterior semialta [õ], que a veces aventajó a la primera en determinadas edades en los sonidos modales. Sin embargo, y al igual que sucedía en los sonidos vocálicos modales, las vocales posteriores manifestaron un notable aumento en el segundo y en el tercer año de vida, sobre todo la vocal posterior semialta [õ]. Una vez más, la vocal posterior alta [ũ] ocupó el último lugar en frecuencia de aparición en

todo el periodo de estudio y en ambos sujetos. Comencemos pues por el primer periodo analizado.

El primer año de vida es un buen ejemplo de la alternancia de dos timbres vocálicos en cuanto a frecuencia de aparición ya que en el caso de la niña fue la vocal central baja [ã] la vocal identificada un mayor número de ocasiones ocupando la vocal anterior semialta [ẽ] el segundo lugar, mientras que en las emisiones del niño la [ẽ] ocupó este primer puesto superando a la [ã], segunda vocal más frecuente, en un gran número de muestras. El tercer puesto fue ocupado en las emisiones del niño por la vocal anterior alta [ĩ] quedando la vocal posterior semialta [õ] en última posición y no identificándose la posterior alta [ũ] en ninguna ocasión. En cuanto a las emisiones de la niña, y debido a la ausencia total de la vocal anterior alta [ĩ] y de la vocal posterior alta [ũ], la vocal posterior semialta [õ] ocupó el tercer lugar.

El orden se invirtió en el segundo año de vida y la vocal anterior semialta [ẽ] pasó a ocupar la primera posición en las emisiones de la niña mientras que fue la vocal central baja [ã] la más frecuente en las vocalizaciones del niño (no obstante, la suma de todos los sonidos anteriores hallados en las emisiones del niño en este periodo superó al número de muestras de la vocal central baja [ã]). Estas mismas vocales se alternaron en la segunda posición que cambió según el sujeto y que se invirtió respecto al primer año: la vocal central baja [ã] ocupó el segundo puesto en la niña y la anterior semialta [ẽ] en el niño. El tercer lugar lo ocupó en ambos sujetos la vocal posterior semialta [õ] que manifestó un notable aumento en este segundo año de vida llegando incluso a superar a la vocal anterior alta [ĩ] que se situó en la cuarta posición en ambos niños. La vocal posterior alta [ũ] quedó una vez más en último lugar en ambos sujetos aunque resulta llamativo el considerable incremento de muestras de esta vocal en el caso del niño donde fue identificada hasta en ocho ocasiones. Este aumento de las vocales posteriores en los sonidos vocálicos nasales coincide con el incremento de estas vocales registrado en el segundo año de vida en los sonidos vocálicos modales de ambos niños.

La vocal más frecuente en las emisiones nasales de ambos sujetos en el tercer año de vida analizado fue la anterior semialta [ẽ], seguida de la vocal central baja [ã] y esta de la vocal posterior semialta [õ]. Se repite por tanto el patrón del segundo año de vida según el cual el aumento que experimenta la vocal posterior semialta [õ] hace que esta supere incluso a la anterior alta [ĩ], vocal que ocupa en ambos sujetos la cuarta posición quedando en último lugar la vocal posterior alta [ũ]. Asimismo, vuelve a encontrarse una similitud entre los sonidos vocálicos modales y nasales extraídos en ambos niños en el tercer año de vida puesto que en ambos las vocales posteriores aumentan de manera considerable.

Tras esta comparativa, podemos determinar que son más las semejanzas que las diferencias en cuanto a timbre vocálico halladas en ambos sujetos y en ambos tipos de señales (modales y nasales) a lo largo de todo el periodo de estudio analizado a pesar de que particularidades inherentes al sujeto fueron encontradas en determinados periodos temporales (preferencias individuales).



## 7. ANÁLISIS ESTADÍSTICOS.

En este apartado mostraremos el promedio de los dos primeros formantes de las señales que comparten el timbre vocálico en cada uno de los meses de los tres periodos estudiados, así como el promedio anual de estos sonidos ya que el análisis de la media de los dos primeros formantes de todas las vocales halladas en cada uno de los meses de los tres periodos analizados (esto es, sin diferenciación de timbre) no tiene mucho sentido al no revelar este datos sobre las vocales concretas sino sobre el conjunto de señales del periodo. De la misma forma, expondremos la desviación estándar de los dos primeros formantes en cada uno de los meses de los tres años de vida revelando la media anual del conjunto de valores.

Los estudios se harán con ayuda de las herramientas (=PROMEDIO( )) y (=DESVEST( )) de *Excel*. El objetivo del primero de los análisis no es otro que el de observar si se producen aumentos o descensos en los promedios de los dos primeros formantes a lo largo del periodo de tiempo que comprende el estudio, o si por el contrario estos se mantienen estables (finalidad ya manifestada entre los objetivos del capítulo tercero de esta tesis). En cuanto al segundo, se trata de calcular la dispersión de los valores del F1 y F2 respecto a su media en cada uno de los meses analizados. Se optó por el cálculo de la desviación estándar frente al de la varianza por razones tales como: por un lado porque el resultado se encuentra expresado en las mismas unidades de la variable, esto es en hertzios y no en valores elevados al cuadrado como en la varianza; por otro, por ser este tipo de análisis el más común en los estudios de análisis acústico revisados. Finalmente, y para averiguar si los valores del F1 y F2 descienden de manera significativa con el paso del tiempo, realizaremos un análisis de varianza o ANOVA (*Analysis of variance*) que nos ayudará a saber si los valores de los dos formantes se asemejan a lo largo de todo el estudio longitudinal, o si por el contrario hay diferencias entre los distintos periodos en que este se divide. Para calcularlo, nos valdremos de la fórmula estadística correspondiente.

### **7.1. Análisis estadísticos realizados sobre las señales que poseen el mismo timbre vocálico en cada uno de los meses analizados en los tres primeros años de vida.**

En este apartado procederemos a realizar los análisis estadísticos de promedio y desviación estándar de los dos primeros formantes de los sonidos que dentro de cada mes fueron identificados por los transcritores con el mismo timbre vocálico. Esta manera de proceder nos permitirá comprobar la evolución de la media del F1 y F2 en cada una de las vocales del español durante los tres años que comprendió el estudio sin que la mayor o menor presencia de un timbre vocálico dentro de un mes concreto pueda determinar los aumentos o descensos del

promedio del primer y segundo formante en ese mes y en ese año. Del mismo modo se analizará en qué medida los valores de ambos formantes se apartan de la media mensual y anual dentro de un mismo timbre vocálico.

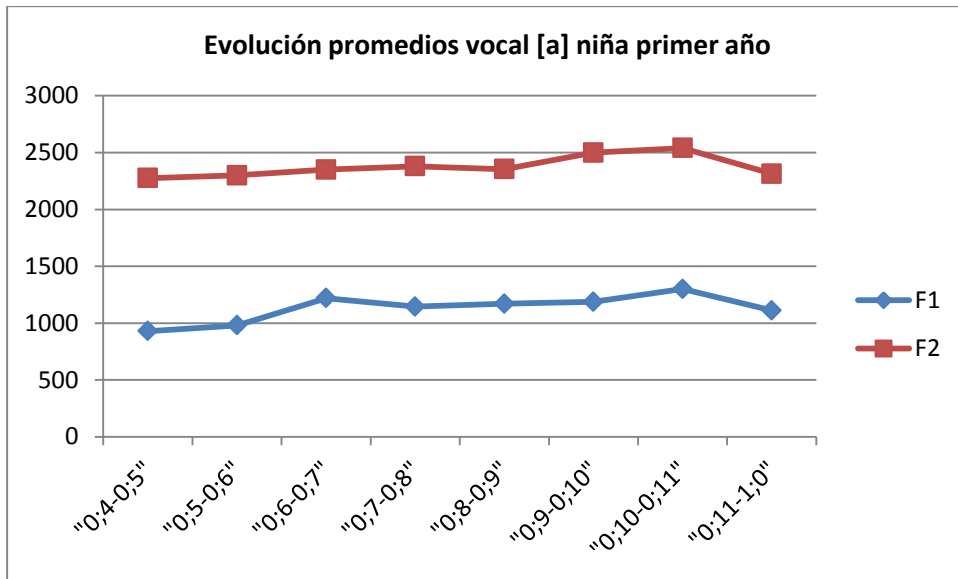
### **7.1.1. Evolución de la media y de la desviación estándar de los dos primeros formantes en las vocales de la niña.**

#### ***7.1.1.1. Vocal central baja, [a].***

Lo primero que observamos al contemplar el gráfico de la evolución del promedio del F1 y F2 de la vocal central baja [a] extraída a lo largo del primer año de vida de la niña (gráfico 7.1), es que dicha vocal aparece en todos los meses del periodo examinado. También salta a la vista la evolución constante de ambos formantes que se mantienen alrededor de unos determinados valores a lo largo de todo el periodo sin cambios violentos. En relación al F1, advertimos un aumento suave pero progresivo desde el primer mes analizado en el periodo, mes que registró el valor más bajo, hasta el penúltimo mes del primer año de vida estudiado, mes en el que se alcanzó el valor máximo. Entre los 0;8 y los 0;9 años los valores prácticamente no variaron (1145 Hz vs 1170 Hz). A partir de los 0;11 años y hacia el último mes de vida analizado, se produce un descenso del F1 algo menos sutil pero de ninguna manera pronunciado, sin que en este decrecimiento se volviera a alcanzar el valor mínimo (registrado en el primer mes de este año). Los valores del F1 se mantuvieron así alrededor de los 1000 Hz durante todo el periodo.

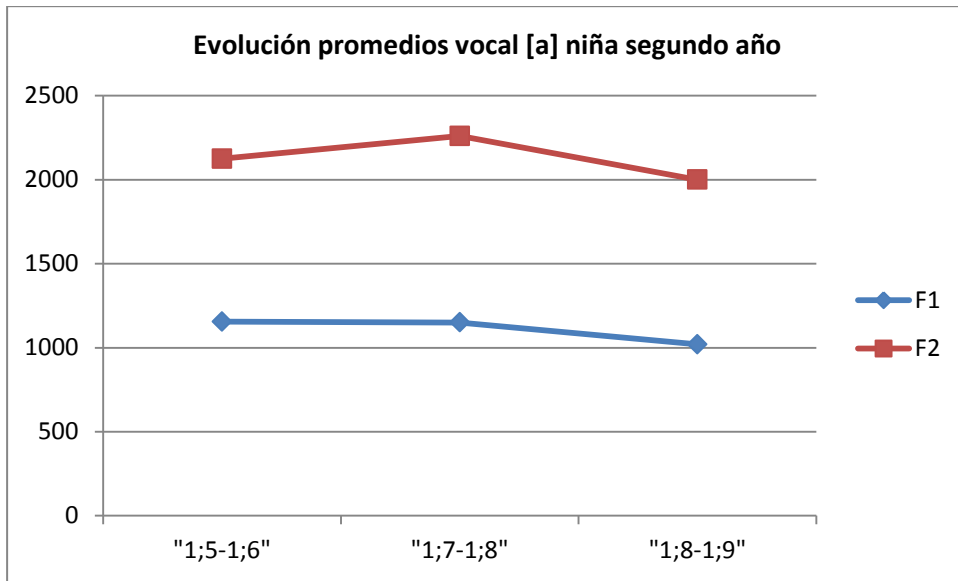
En cuanto al F2, los ascensos y descensos fueron también muy suaves situándose los valores alrededor de los 2300 Hz y 2500 Hz a lo largo de todo el periodo. Durante los tres primeros meses encontramos que los valores se mantuvieron prácticamente invariables con un pequeño aumento y un ligero descenso entre los 0;7 y los 0;9 años de vida, mes tras el cual los valores ascendieron de manera progresiva hasta alcanzar el valor máximo entre los 0;10 y los 0;11 años. Tal y como ocurría con el F1, el segundo formante también descendió en el último mes del primer año de vida sin llegar a registrarse en este mes el valor mínimo del periodo.

**Gráfico 7.1. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año).**



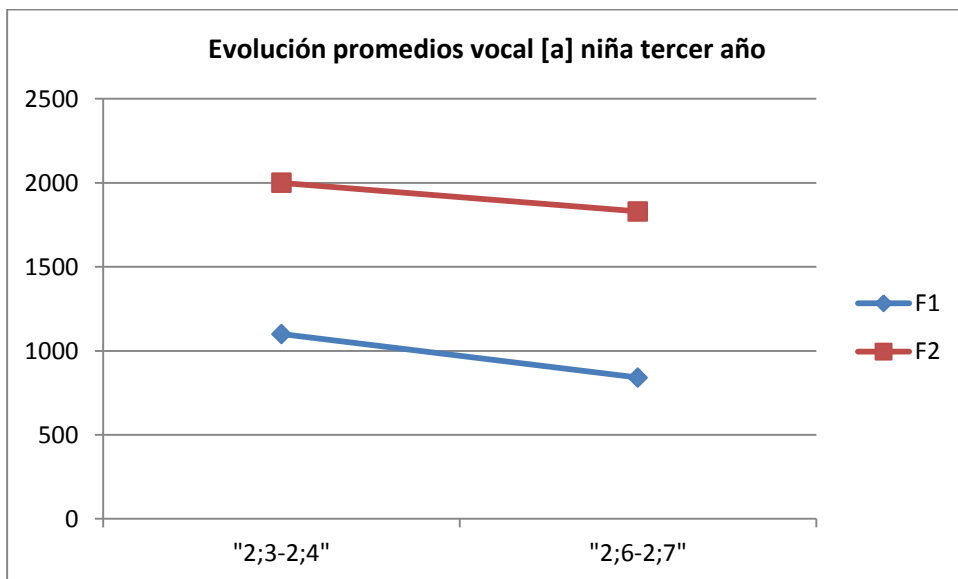
Tan solo en tres meses del segundo año de vida examinado pudieron extraerse muestras identificadas como aes por los oyentes (gráfico 7.2). Respecto al F2, este descendió unos 188 Hz con respecto al último mes del periodo anterior, aumentó hasta los 1;8 años y descendió entre los 1;8 y 1;9 años alcanzando en este último mes el valor más bajo (2000 Hz). Sin embargo en el F1, cuyos valores aumentaron 43 Hz entre el último mes del primer periodo analizado y el primero del segundo año de vida, la tendencia fue hacia el descenso en los tres meses analizados de este segundo año registrándose también el valor más bajo entre los 1;8-1;9 años (1020 Hz).

**Gráfico 7.2. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años).**



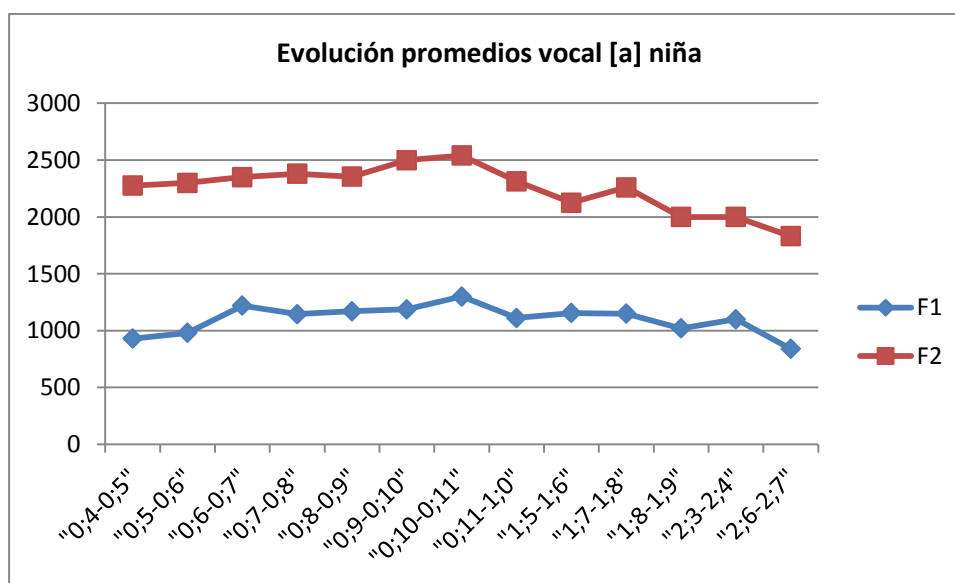
En el periodo comprendido entre los 2;0 y los 3;0 años de vida tan solo fueron dos los meses en los que se identificó esta vocal de manera conjunta (gráfico 7.3). Ambos formantes comenzaron con un valor semejante respecto al último mes del segundo año de vida analizado siendo el F1 tan solo 80 Hz mayor y manteniendo el F2 el mismo valor (2000 Hz). Los valores tendieron a bajar en el último mes analizado del periodo registrando ambos formantes en dicho mes los valores mínimos del tercer año.

**Gráfico 7.3. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en el tercer año de vida (2;0-3;0 años).**



En el siguiente gráfico (gráfico 7.4), que resume la evolución de los dos primeros formantes de esta vocal a lo largo de los tres años de vida estudiados, podemos observar que ambos formantes comienzan un descenso progresivo a partir de los 0;11 años (mes en el que dichos formantes alcanzaron el valor máximo) apartándose más el F2 de su valor más alto de lo que lo hizo el F1.

**Gráfico 7.4. Evolución de los promedios de los dos primeros formantes de la vocal [a] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Ofrecemos ahora las tablas que recogen el promedio de los valores del primer y segundo formante en cada uno de los meses que conforman los tres años de vida observados, así como la media anual de los mismos (los valores recogidos aquí han sido redondeados).

**Tabla 7.1. Valores de los promedios de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;4-0;5	930	2275
0;5-0;6	980	2300
0;6-0;7	1220	2350
0;7-0;8	1145	2380
0;8-0;9	1170	2355
0;9-0;10	1187	2500
0;10-0;11	1300	2540
0;11-1;0	1112	2313

Media anual	1103	2359
-------------	------	------

**Tabla 7.2. Valores de los promedios de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;5-1;6	1155	2125
1;7-1;8	1150	2260
1;8-1;9	1020	2000
Media anual	1120	2128

**Tabla 7.3. Valores de los promedios de los dos primeros formantes de la vocal [a] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;3-2;4	1100	2000
2;6-2;7	840	1830
Media anual	970	1915

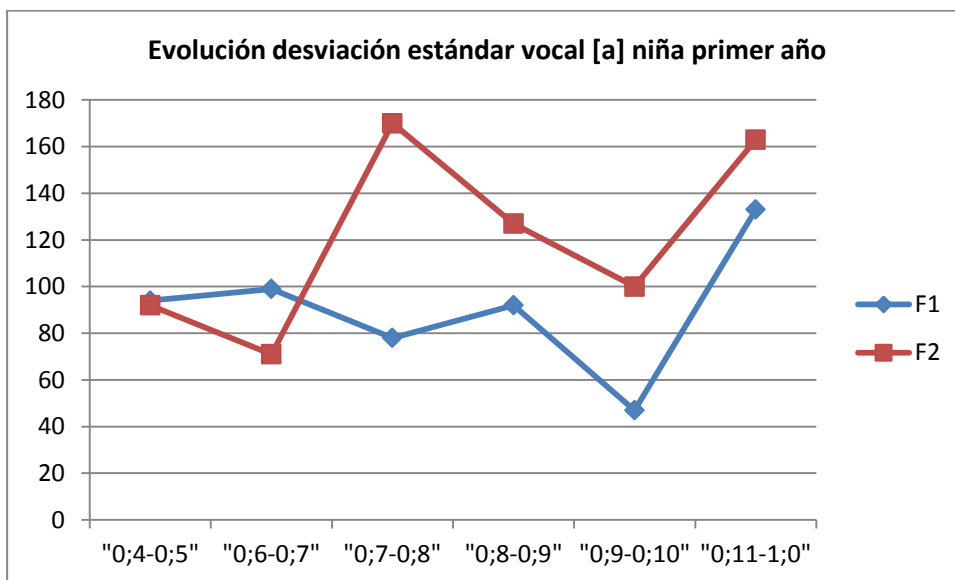
Una vez examinados los valores anuales, podemos observar que en el F1 se produjo un aumento de 17 Hz entre el primer y el segundo año de vida analizado pero un descenso de 150 Hz entre el segundo y el tercero. La diferencia de valores entre el primer y el tercer año de vida fue de 133 Hz. A pesar de que a simple vista ni el incremento ni la reducción en el F1 parecían muy notables, decidimos realizar el ANOVA para averiguar si estos cambios eran o no significativos. El análisis estadístico demostró que efectivamente el promedio anual de los tres años de vida era muy similar y que no se produjeron diferencias significativas entre el primer y el segundo año, entre el segundo y el tercero ni entre el primero y el tercero aceptándose por tanto la hipótesis nula al comparar todos los periodos (ninguna diferencia significativa se observó tampoco entre los cuatro primeros meses analizados en el primer año de vida y los cuatro últimos, esto es, entre 0;4-0;8 años y entre 0;8-1;0 año).

En cuanto al F2, la disminución en los valores fue más acusada. De este modo, una caída de 231 Hz se produjo entre el primer y el segundo año de vida, una de 213 Hz entre el segundo y el tercero, y una de 444 Hz entre el primero y el tercero. El ANOVA reveló que existía una diferencia significativa entre el primer y el segundo año de vida y para averiguar en qué mes del segundo año de vida los valores descendían significativamente con respecto al

primero, repetimos los análisis estadísticos en cada una de las edades de este segundo año. El análisis de varianza informó de que los valores se mantenían estables hasta los 1;5 años descendiendo significativamente entre los 1;5 y los 1;6 años de edad. Un tercer análisis nos descubrió que a partir de los dieciocho meses de edad los valores del F2 no descendieron de manera significativa siendo similares por tanto los datos obtenidos entre los 1;0-2;0 y los 2;0-3;0 años.

Respecto a la desviación estándar de la [a] en este primer año de vida y a pesar de que encontramos alguna subida y alguna bajada algo bruscas tal y como presenta el gráfico 7.5, si nos detenemos un poco más en el mismo observamos que estas se produjeron en un rango de valores estrecho: de los 47 Hz a los 133 Hz para el F1 (valor mínimo y máximo) y de los 71 Hz a los 170 Hz para el F2. En el caso del F1, la desviación de este formante comenzó con un pequeño ascenso y un ligero descenso prosiguiendo este zigzag en los meses siguientes si bien los aumentos y la disminución fueron mayores (alcanzándose el valor más alto en el último mes del primer año de vida y el más bajo en el antepenúltimo). En cuanto al F2, la desviación de este experimentó el mayor descenso entre los 0;6 y los 0;7 años, mes en el que registró el valor más bajo, y presentó el mayor aumento entre los 0;7 y los 0;8 años, mes en el que alcanzó el valor máximo. A partir de aquí el F2 descendió y volvió a ascender acercándose en el último mes analizado al valor más alto del periodo.

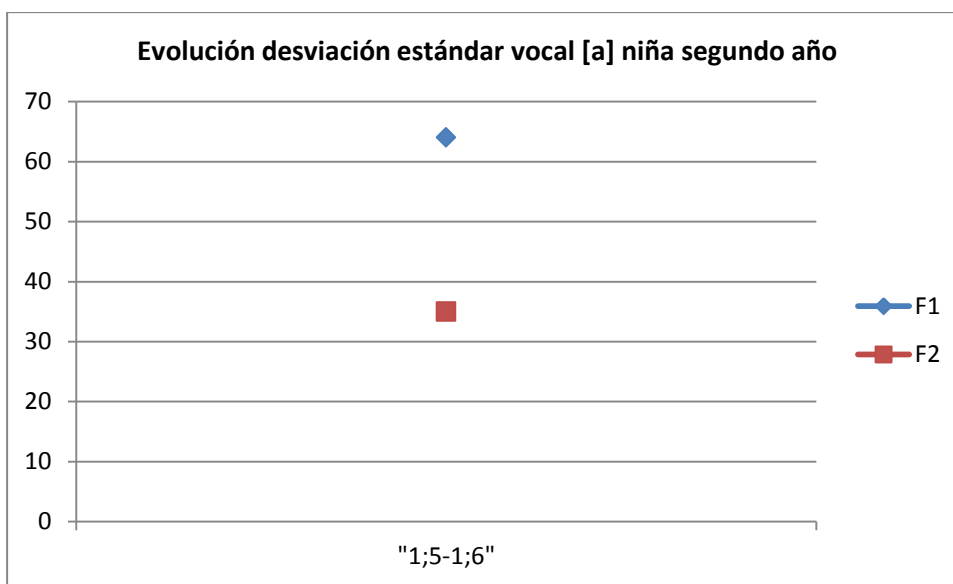
**Gráfico 7.5. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año).**



En el gráfico que mostramos a continuación (gráfico 7.6) tan solo hallamos un punto en la evolución de la desviación estándar debido a que de los tres meses de este segundo año en los

que los oyentes coincidieron en la identificación de la vocal central baja [a], tan solo en uno se encontró más de una señal, hecho imprescindible para calcular la desviación estándar. Dicha desviación se dio entre las señales de los 1;5 y 1;6 años de vida y tal y como percibimos en el gráfico, el grado en el que se alejaron de la media los valores del F1 y F2 de los sonidos agrupados en ese mes fue muy bajo (64 Hz el primero y 35 Hz el segundo). Con respecto al último mes del año anterior, la desviación estándar de ambos formantes en este segundo año fue muy similar siendo la del F1 69 Hz menor y la del F2 128 Hz más baja.

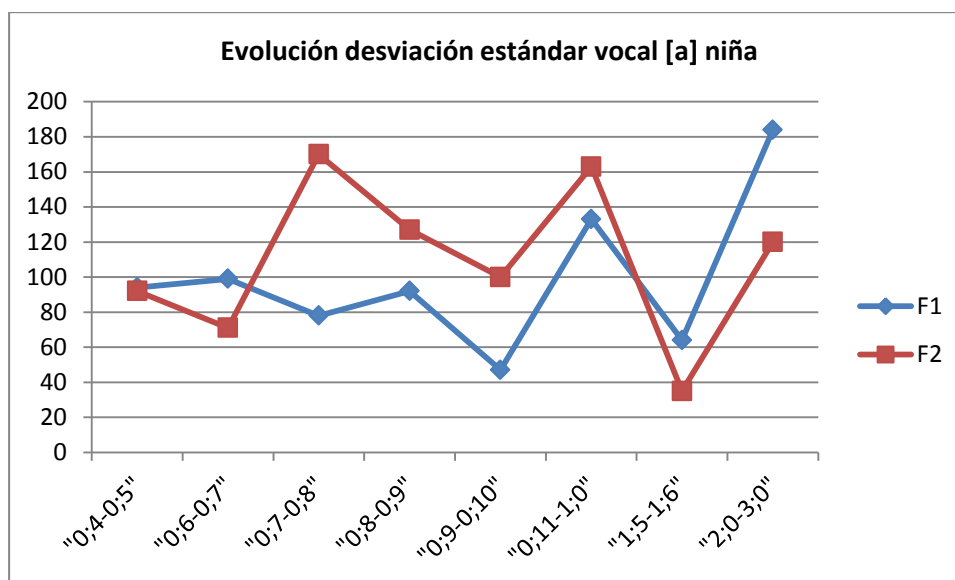
**Gráfico 7.6. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años).**



La ausencia de más de una señal en cada uno de los tres meses del tercer año de vida en los que los oyentes identificaron la [a], nos ha llevado a mostrar directamente el gráfico que recoge la desviación estándar de los tres periodos analizados (gráfico 7.7). El patrón de montaña rusa predomina en ambos formantes sin que los aumentos y los descensos de la desviación fueran en general bruscos durante el primer año de vida aunque sí lo fueron algo más en el segundo y en el tercero. Como diferencia podemos apuntar que en el caso del F2 la desviación máxima se alcanza en el primer año y la mínima en el segundo mientras que en el F1 la desviación mínima se registra en el primero y la máxima se alcanza en el tercero.



**Gráfico 7.7. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Ofrecemos seguidamente las tablas en las que se encuentran los valores concretos de la desviación estándar del F1 y F2 en cada uno de los meses de los tres años de vida y el promedio anual de la misma:

**Tabla 7.4. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;4-0;5	94	92
0;6-0;7	99	71
0;7-0;8	78	170
0;8-0;9	92	127
0;9-0;10	47	100
0;11-1;0	133	163
Desviación estándar anual	141	128

**Tabla 7.5. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;5-1;6	64	35
Desviación estándar anual	76	108

**Tabla 7.6. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
Desviación estándar anual	184	120

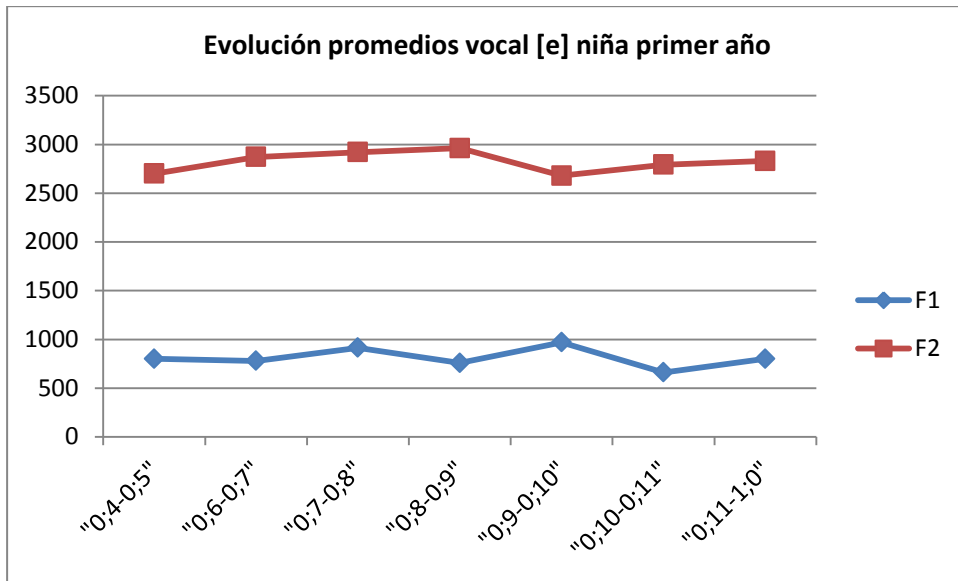
Según se desprende de las tablas, la desviación típica de los dos primeros formantes se mantuvo alrededor de los mismos valores a lo largo de los tres periodos estudiados. En el caso del F1, esta descendió entre el primer y el segundo año de vida y aumentó entre el segundo y el tercero si bien la diferencia entre la desviación del primer y del tercer año fue de tan solo 43 Hz. Respecto al F2, la desviación estándar de este formante también descendió entre el primer y el segundo año de vida y aunque aumentó entre el segundo y el tercero, la desviación del tercer año de vida fue muy similar a la del primero (128 Hz en el primer año y 120 Hz en el tercero).

#### ***7.1.1.2. Vocal anterior semialta, [e].***

Proseguimos nuestro estudio con el análisis de las vocales anteriores y comenzamos con el de la vocal anterior semialta [e], vocal que sigue a la [a] en el orden alfabético español. Llama la atención al contemplar el primer gráfico, el que recoge el promedio del F1 y F2 de dicha vocal durante el primer año de vida examinado (gráfico 7.8), la elevada presencia de dicha vocal a lo largo del periodo ya que esta apareció en todos los meses del mismo excepto en el que transcurrió entre los 0;5 y 0;6 años (algo similar a lo que ocurría con la vocal central baja, [a]).

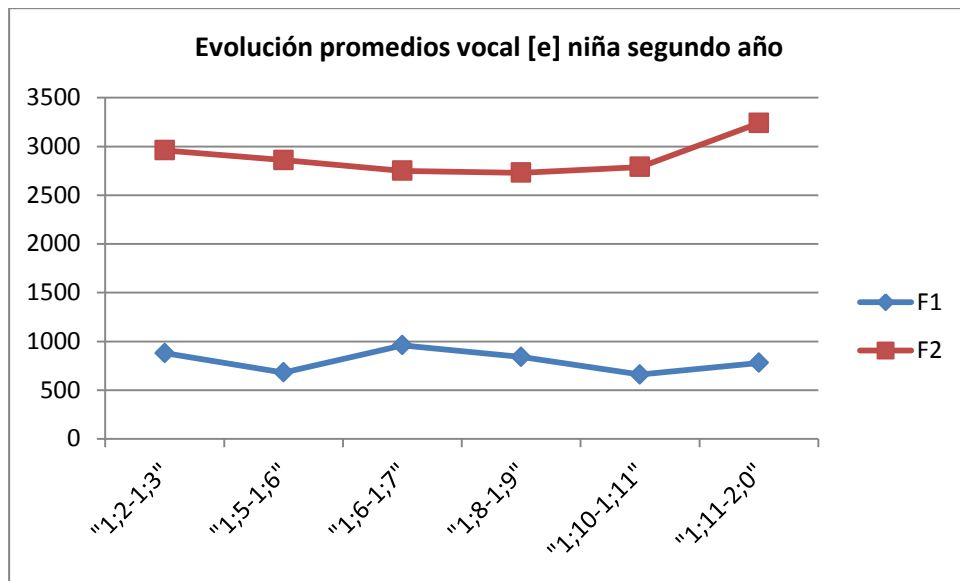
En cuanto al promedio de los formantes, podemos observar la evolución prácticamente lineal de la media del primero de ellos con ligeros ascensos y descensos de los valores. Dichos valores se situaron entre los 700 Hz y 900 Hz a lo largo de todo el periodo temporal excepto entre los 0;9 y 0;10 años, mes en el que se alcanzó el valor máximo (de 970 Hz) y la pendiente del incremento fue algo más notoria. Respecto al F2, los valores se situaron por encima de los 2600 Hz y por debajo de los 3000 Hz durante todo el periodo analizado alcanzándose entre los 0;8 y 0;9 años el promedio máximo. Como en el F1, la evolución de los mismos fue prácticamente lineal y como en aquel encontramos un periodo en el que las pendientes del incremento y del descenso fueron más notables situándose en este caso entre los 0;9 y los 0;10 años. El resto de desniveles hallados en el periodo fueron más graduales.

**Gráfico 7.8. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año).**



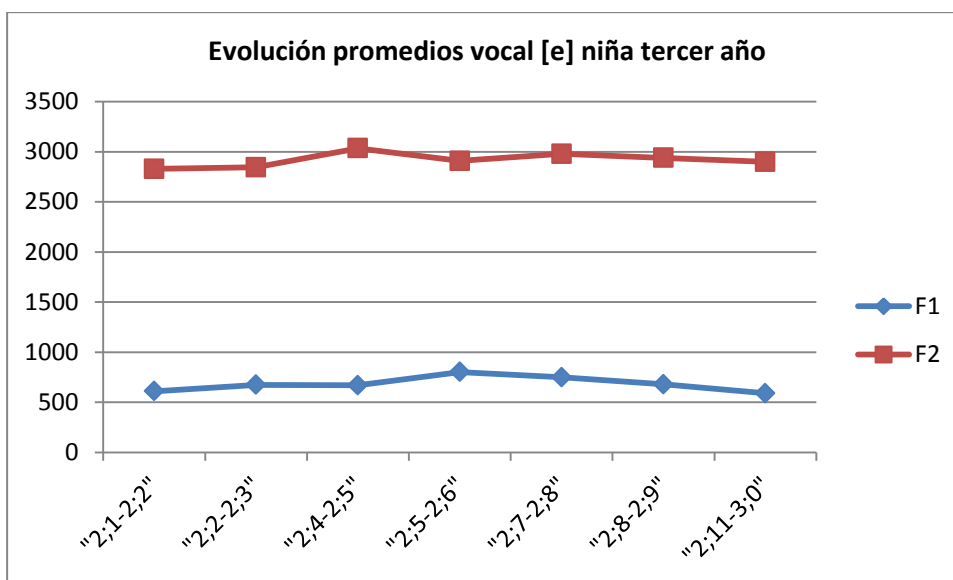
En el segundo año de vida el número de meses en los que la vocal anterior semialta estuvo presente fue de seis (gráfico 7.9). A primera vista e incluyendo ambos formantes, dos observaciones pueden ser hechas: por un lado, que ni los promedios del F1 ni los del F2 parecieron descender o aumentar con respecto al primer año de vida sino que más bien resultaron muy similares; por otro, la ausencia de cambios repentinos asemejó la evolución de los promedios de los dos primeros formantes que se vieron no obstante alejados en el tipo de evolución. En relación al F1, un pequeño aumento (de 80 Hz) se produjo entre el último mes del primer año de vida y el primer mes analizado del segundo (1;2-1;3 años). A continuación, los valores descendieron y aumentaron de manera progresiva hasta los 1;10-1;11 años, mes en el que se registró el promedio mínimo del periodo. El promedio del F2 por otro lado, también experimentó un aumento (de 130 Hz) entre el último mes del primer año de vida y el primer mes examinado del segundo. En cuanto a su evolución, la media del segundo formante descendió sutilmente entre los 1;2 y los 1;9 años, mes en el que registró el valor mínimo (2730 Hz), y aumentó también de manera suave entre los 1;9 y los 2;0 años alcanzando en el último mes del segundo año de vida el valor más alto (3240 Hz).

**Gráfico 7.9. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años).**



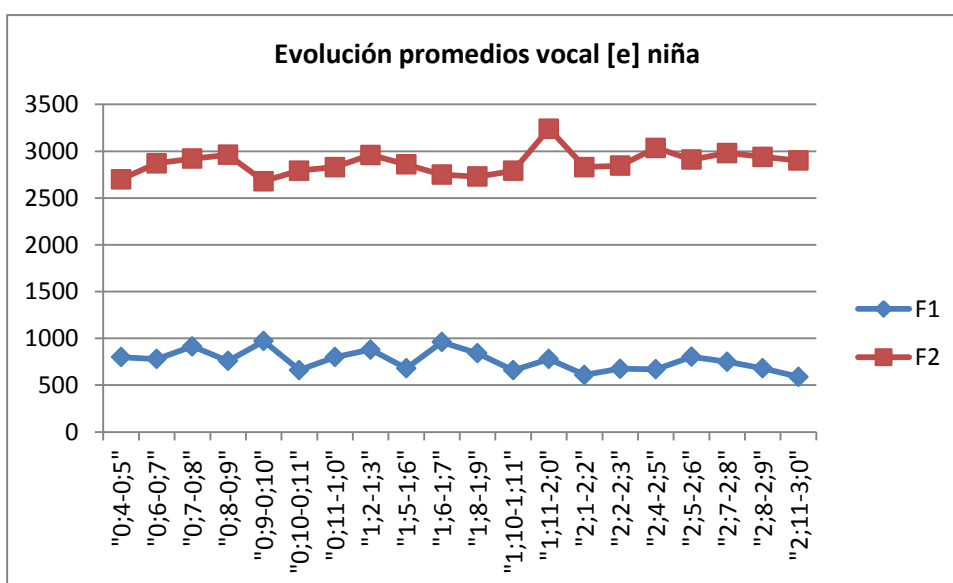
En siete de los doce meses del tercer año de vida estuvo presente la vocal anterior semialta (gráfico 7.10) y al igual que ocurría entre el primer y el segundo periodo analizados, ninguno de los dos formantes pareció presentar cambios importantes en los valores de sus promedios. Un descenso de 170 Hz tuvo lugar en el primer formante entre el último mes del segundo año de vida (1;11-2;0 años) y el primer mes analizado del tercero (2;1.2;2 años). El promedio de este formante aumentó sutilmente hasta la primera mitad del tercer año de vida (con un casi imperceptible descenso de 5 Hz entre los 2;4 y 2;5 años), momento en el que alcanzó el valor máximo, y descendió también de manera suave en la segunda mitad de este tercer año registrando en el último mes del periodo el valor mínimo. En cuanto al F2, este experimentó pequeñas subidas y bajadas a lo largo de todo el periodo registrando el valor mínimo en el primer mes analizado (2;1-2;2 años) y alcanzando a los 2;5 años el valor máximo.

**Gráfico 7.10. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años).**



En el siguiente gráfico (gráfico 7.11) queda mejor reflejada la ausencia de descensos y aumentos dramáticos tanto en los promedios del F1 como en los del F2 a lo largo de los tres periodos anuales estudiados. Se observa por el contrario, que la línea evolutiva progresa alrededor de unos mismos valores con pequeños incrementos y descensos. El perfil es por tanto diferente al hallado para la vocal central baja [a] en cuya evolución los valores, al menos en el F2, caían notoriamente a partir de un mes determinado sin que volvieran a alcanzarse los máximos anteriores.

**Gráfico 7.11. Evolución de los promedios de los dos primeros formantes de la vocal [e] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Presentamos a continuación las tablas con los valores concretos de los promedios que registraron el F1 y F2 a lo largo de los tres años de vida observados:

**Tabla 7.7. Valores de los promedios de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;4-0;5	800	2700
0;5-0;6		
0;6-0;7	780	2871
0;7-0;8	914	2921
0;8-0;9	758	2962
0;9-0;10	970	2680
0;10-0;11	661	2793
0;11-1;0	800	2830
Media anual	795	2876

**Tabla 7.8. Valores de los promedios de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;2-1;3	880	2960
1;5-1;6	680	2860
1;6-1;7	960	2750
1;8-1;9	840	2730
1;10-1;11	660	2790
1;11-2;0	780	3240
Media anual	768	2873

**Tabla 7.9. Valores de los promedios de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años).**

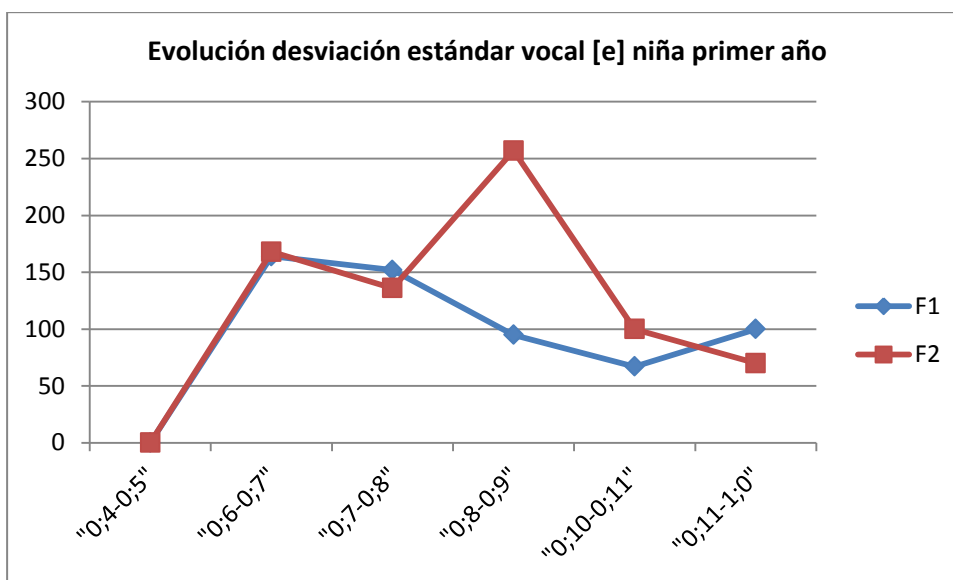
Edad	Media F1	Media F2
2;1-2;2	610	2830
2;2-2;3	675	2845
2;4-2;5	670	3035

2;5-2;6	803	2910
2;7-2;8	750	2980
2;8-2;9	680	2940
2;11-3;0	590	2900
Media anual	703	2914

Tal y como se desprende de las tablas el promedio anual del F1 fue de 795 Hz en el primer año de vida y de 768 Hz en el segundo produciéndose un descenso mínimo de 27 Hz. Entre el segundo y el tercer año el promedio del F1 se redujo 65 Hz habiendo por tanto una diferencia de 92 Hz entre el primer y el tercer año de vida. En el F2, los valores se mantuvieron entre el primer y el segundo año de vida (2876 Hz y 2873 Hz respectivamente) y aumentaron en el tercero (2914 Hz). A pesar de que las diferencias entre los promedios de los distintos años nos parecieron mínimas tanto en el F1 como en el F2, realizamos un análisis de varianza sobre el conjunto de datos para comprobarlo. Los resultados indicaron que no había diferencias significativas en los promedios del F1 y F2 de los tres años de vida no advirtiéndose por tanto ningún descenso ni aumento pertinente de los valores entre los cuatro primeros y los cuatro últimos meses del primer año de vida, entre el primer y el segundo año ni entre el segundo y el tercero. Sin embargo, el ANOVA arrojó un descenso significativo del F1 entre el primer y el tercer año de vida a pesar de que en el primer año la media del F1 fue de 795 Hz y en el tercero de 703 Hz, es decir, tan solo hubo 92 Hz de diferencia, por lo que quizás este resultado se deba a la considerable diferencia en el número de muestras halladas en el primer y el tercer año de vida (37 datos en el primer año y tan solo 13 en el segundo). El segundo formante de la vocal anterior semialta [e] presentó por tanto un patrón muy distinto al encontrado en la evolución del promedio de este formante para la vocal central baja [a], que descendió significativamente en el segundo año de vida.

Respecto a la desviación estándar que presentan los valores de los dos primeros formantes en el primer año de vida (gráfico 7.12), encontramos que tanto el F1 como el F2 parten de una desviación estándar mínima (0 Hz) que aumenta en el mes siguiente alcanzando en ambos formantes prácticamente el mismo valor (164 Hz en el F1 y 168 Hz en el F2). En el caso del F1, esta es la desviación estándar máxima de todo el periodo y a partir de este mes la desviación típica tiende a descender hasta el último mes de este primer año donde aumenta ligeramente. En cuanto al F2, la desviación presenta bajadas y subidas registrándose la desviación estándar máxima entre los 0;8 y los 0;9 años.

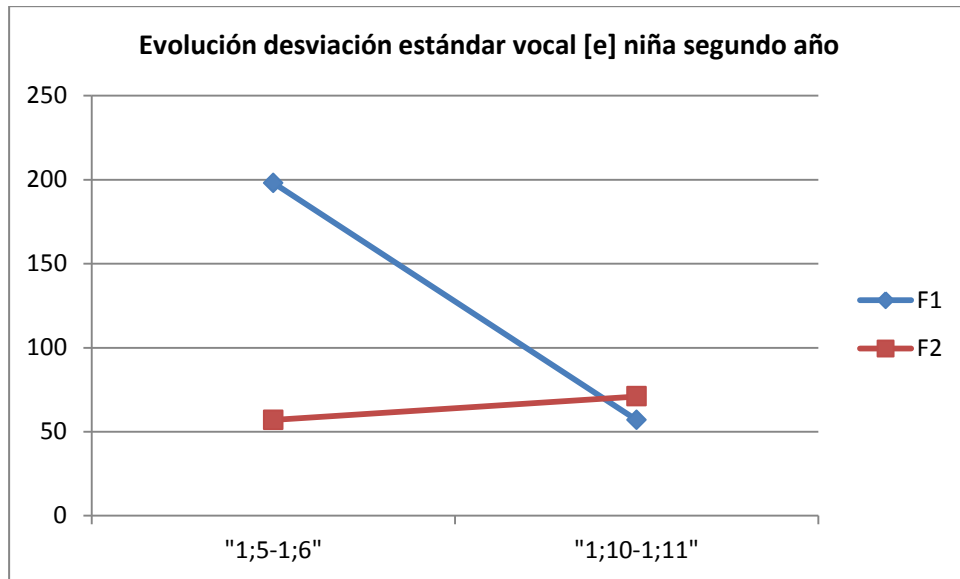
**Gráfico 7.12. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año).**



Tan solo en dos meses del segundo año de vida pudo calcularse la desviación estándar de los dos primeros formantes: a mitad del periodo y en el penúltimo mes del mismo (gráfico 7.13). La desviación típica del F1 y del F2 presentó, tal y como muestra el gráfico inferior, un patrón evolutivo contrario yendo el F1 de una desviación estándar máxima a una mínima (de 198 Hz a 57 Hz) y ascendiendo el F2 de una desviación típica mínima a una máxima (de 57 Hz a 71 Hz). En cuanto a la relación entre la desviación estándar del último mes del primer año de vida y el primer mes analizado del segundo, la desviación estándar del F1 aumentó 98 Hz y la del F2 disminuyó 13 Hz.

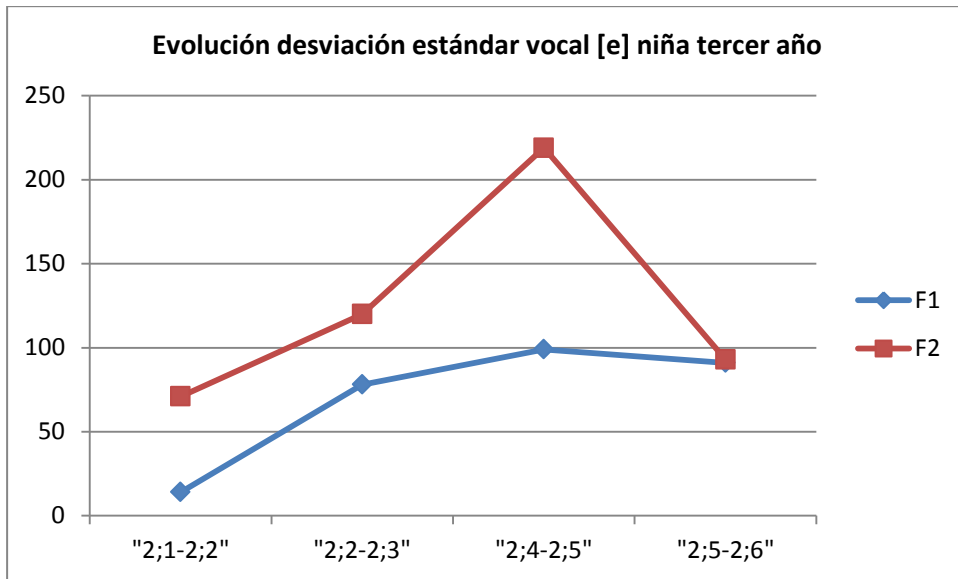


**Gráfico 7.13. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años).**



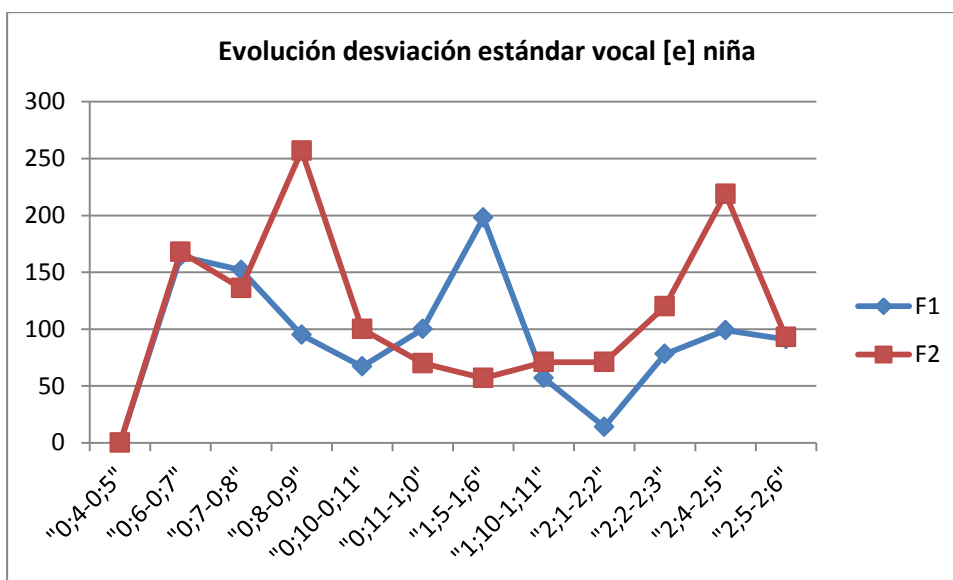
Tal y como puede observarse en el gráfico 7.14, la desviación estándar de los dos primeros formantes en el tercer año de vida presentó una evolución muy similar. De esta forma, ambos formantes registraron la desviación estándar mínima en el primer mes analizado y presentaron un aumento de la misma hasta los 2;4-2;5 años, mes en el que ambos formantes alcanzaron la desviación estándar máxima (con una pendiente más acusada en el F2), y a partir del cual la desviación descendió (de manera más brusca en el segundo formante). Además, la desviación estándar del F1 descendió 43 Hz entre el último mes analizado del segundo año de vida y el primero del tercero, manteniéndose igual en el F2 (71 Hz).

**Gráfico 7.14. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años).**



El gráfico inferior (gráfico 7.15) resume la evolución de la desviación estándar de la vocal anterior semialta a lo largo de los tres años de vida estudiados. Tal y como se desprende del mismo, la desviación con respecto a la media tanto del F1 como del F2 comenzó registrando el valor mínimo en el primer mes del primer periodo y alcanzó el valor máximo antes del tercer año (en el primero para el F2 y en el segundo para el F1). En cuanto a la evolución, podemos observar un patrón contrario pues cuando la desviación típica del F1 desciende la del F2 aumenta, y viceversa, hasta el tercer año de vida donde la desviación aumenta y desciende en ambos formantes.

**Gráfico 7.15. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Las tablas siguientes muestran la desviación típica de los valores del F1 y F2 en los tres periodos analizados:

**Tabla 7.10. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;4-0;5	0	0
0;6-0;7	164	168
0;7-0;8	152	136
0;8-0;9	95	257
0;10-0;11	67	100
0;11-1;0	100	70
Desviación estándar anual	143	181

**Tabla 7.11. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;5-1;6	198	57
1;10-1;11	57	71
Desviación estándar anual	139	169

**Tabla 7.12. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
2;1-2;2	14	71
2;2-2;3	78	120
2;4-2;5	99	219
2;5-2;6	91	93
Desviación estándar anual	99	112

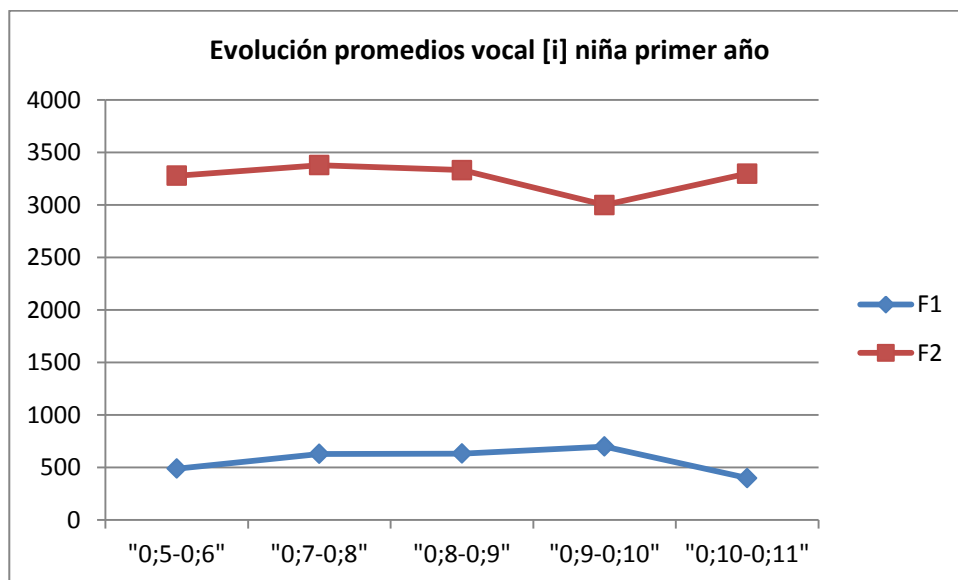
Como puede desprenderse de las tablas, la media anual de la desviación típica de ambos formantes descendió con el paso del tiempo. Esto fue lo que sucedió con el F1 cuya desviación estándar pasó de los 143 Hz a los 139 Hz entre el primer y el segundo año de vida y a los 99 Hz en el tercero, y también con el F2 cuya desviación estándar pasó de los 181 Hz a los 169 Hz entre el primer y el segundo año de vida y a los 112 Hz en el tercero. Por último y tras comparar la desviación estándar hallada en la vocal anterior semialta con la de la vocal central baja, dos apreciaciones pueden ser hechas: por un lado, en la vocal central baja la desviación estándar de ambos formantes disminuye entre el primer y el segundo año de vida pero aumenta ligeramente entre el segundo y el tercero mientras que la desviación estándar de los dos primeros formantes de la vocal anterior semialta tiende a descender con el paso del tiempo; por otro, los valores anuales de dichas desviaciones tanto en el F1 como en el F2 fueron muy similares en ambas vocales.

### ***7.1.1.3. Vocal anterior alta, [i].***

Proseguimos exponiendo la evolución del promedio del F1 y del F2 de la vocal anterior alta [i] durante el primer año de vida (gráfico 7.16). Lo primero que puede señalarse es que a diferencia de lo que ocurría con las vocales [e], que fue identificada por los sujetos que realizaron el test de percepción en siete de los ocho meses analizados en el primer año de vida, y [a], que fue reconocida en la totalidad del mismo, la [i] solo fue observada en cinco de los ocho meses del primer periodo. En cuanto al movimiento de las medias, llama la atención la evolución casi lineal de estas hasta los 0;8-0;9 años en el caso del segundo formante, y los 0;9-0;10 años en el caso del primero. En relación al F2, a partir de este mes se produce un descenso algo más notorio registrándose en el mes siguiente (0;9-0;10 años) el promedio más bajo si bien

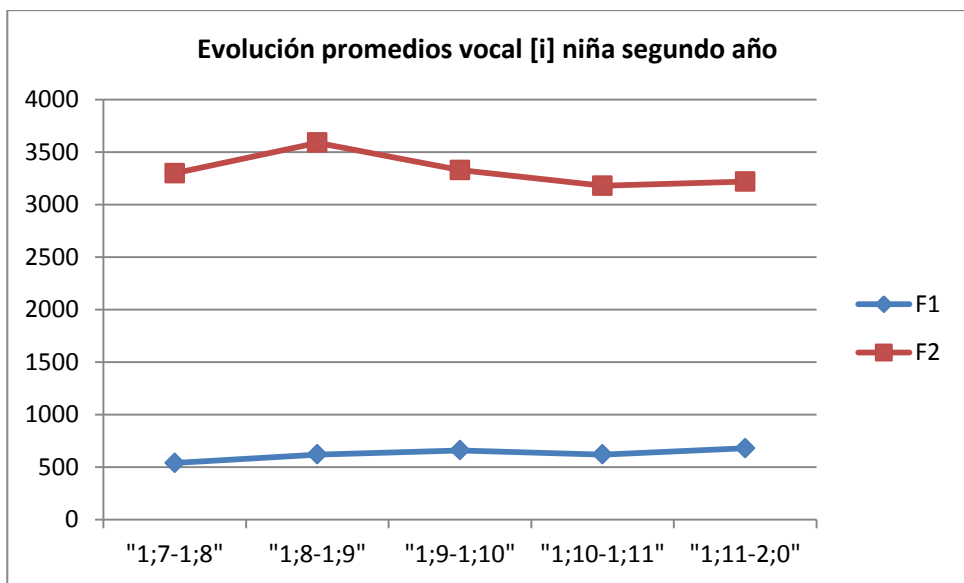
un mes más tarde se vuelven a alcanzar los valores anteriores. En el F1, la evolución lineal se prolonga hasta los 0;9-0;10 años produciéndose el descenso en el mes siguiente que además registra el promedio menor de todo el año.

**Gráfico 7.16. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año).**



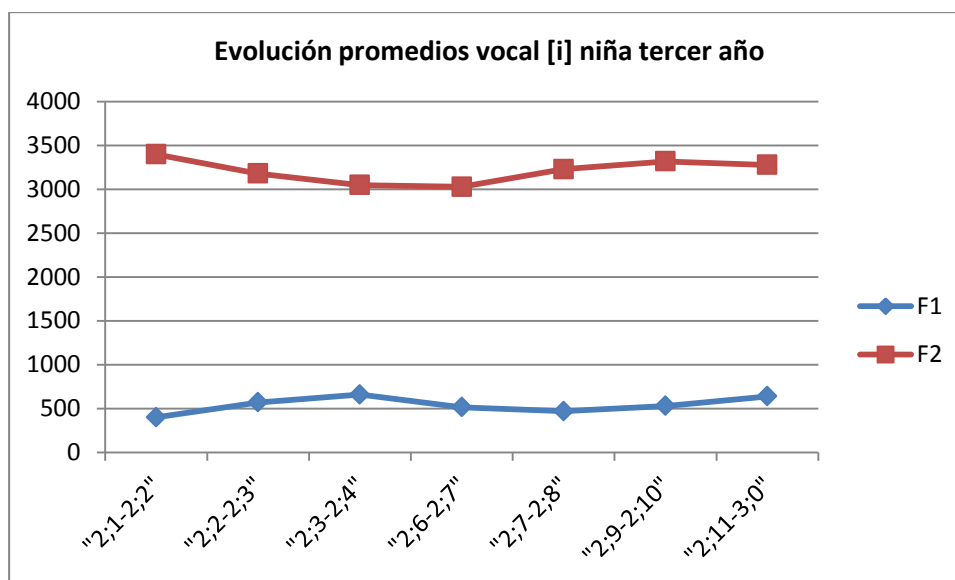
La progresión lineal del promedio del F1 hallada durante el primer año de vida pareció repetirse en el segundo (gráfico 7.17). Dicha media experimentó un pequeño aumento con respecto al último mes analizado del primer periodo estudiado (de 140 Hz), siendo este primer mes del segundo año de vida el que registró el valor mínimo. A continuación se produjo un pequeñísimo ascenso que se prolongó hasta el último mes examinado en el segundo año. Respecto al F2, no hubo diferencia entre el último mes analizado del primer año de vida y el primero del segundo puesto que ambos presentaron un promedio de 3300 Hz. En el mes siguiente se produjo un aumento y el F2 alcanzó la media máxima del periodo (entre los 1;8-1;9 años). A partir de este mes los valores descendieron hasta registrar el valor mínimo en el penúltimo mes observado de este segundo año.

**Gráfico 7.17. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años).**



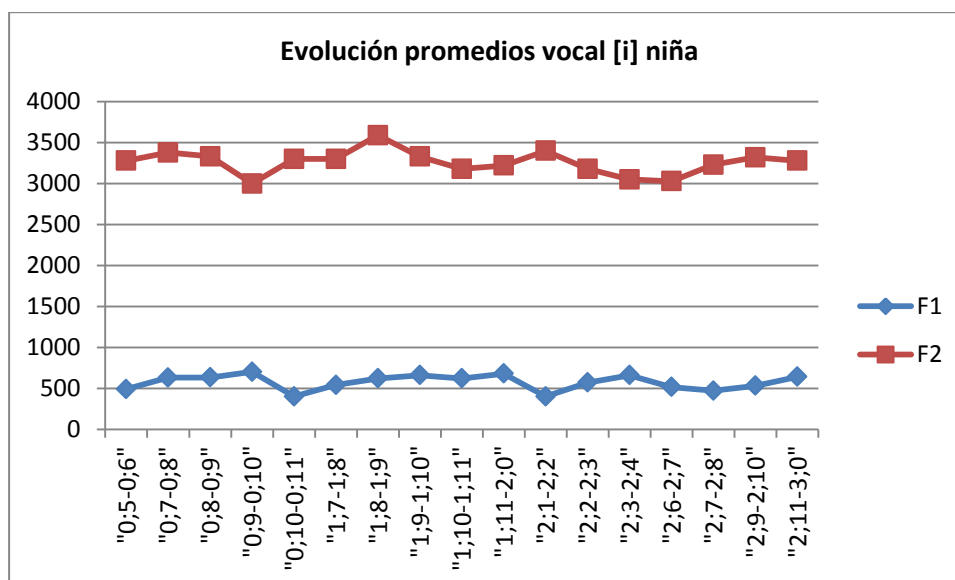
De la misma forma que sucedía en los años anteriores, los valores del F1 en el tercer año de vida no variaron demasiado pudiéndose evidenciar una vez más una progresión bastante rectilínea de este formante durante el tercer periodo temporal (gráfico 7.18). Entre el último mes examinado del segundo año de vida y el primero del tercero, el primer formante descendió 280 Hz registrándose en este mes el valor mínimo del periodo. Dos meses más tarde sin embargo, el F1 alcanzó su valor máximo (660 Hz) sin que una inclinación demasiado pronunciada pudiera observarse entre ambos. En los meses sucesivos el F1 descendió situándose alrededor de los 500 Hz y volvió a incrementarse ligeramente en el último mes de la etapa. La media del F2, a diferencia de lo que ocurría con la del F1, aumentó 180 Hz entre el último mes del segundo año de vida y el primero del tercero alcanzando en este primer mes del tercer año de vida el promedio máximo de todo el periodo. A partir de aquí el F2 experimentó pequeños descensos hasta los 2;6-2;7 años (mes en el que se registró el valor más bajo), volvió a incrementarse también de manera suave hasta los 2;9-2;10 años y descendió ligeramente en el último mes del periodo.

**Gráfico 7.18. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años).**



La evolución del promedio de los valores del F1 y del F2 a lo largo de los tres años de vida puede contemplarse de manera conjunta en el siguiente gráfico (gráfico 7.19). Como puede desprenderse de dicho gráfico, el F1 siguió una progresión bastante lineal a lo largo de todo el periodo de estudio y aunque el F2 experimentó cambios algo más notorios, sobre todo en el primer año de vida, el promedio no descendió a partir de un punto determinado (como entre los 0;9 y 0;10 años) sin que volvieran a registrarse los valores anteriores sino que la media volvió a recuperarse registrando incluso sus valores máximos en los periodos de tiempo posteriores (como entre los 1;8 y 1;9 años). La ausencia de una caída considerable de los valores formánticos y de un punto de inflexión a partir del cual dichos valores no consiguieran volver a registrar valores anteriores, asemeja la evolución de la media del F1 y F2 de la [i] a la de la otra vocal anterior ya estudiada, la anterior semialta [e].

**Gráfico 7.19. Evolución de los promedios de los dos primeros formantes de la vocal [i] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Las tablas siguientes recogen los promedios exactos obtenidos en los meses analizados en cada uno de los años así como la media final anual de los tres periodos examinados.

**Tabla 7.13. Valores de los promedios de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;5-0;6	490	3280
0;7-0;8	630	3380
0;8-0;9	633	3333
0;9-0;10	700	3000
0;10-0;11	400	3300
Media anual	609	3304

**Tabla 7.14. Valores de los promedios de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;7-1;8	540	3300
1;8-1;9	620	3590
1;9-1;10	660	3330
1;10-1;11	620	3180



1;11-2;0	680	3220
Media anual	629	3304

**Tabla 7.15. Valores de los promedios de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;1-2;2	400	3400
2;2-2;3	570	3180
2;3-2;4	660	3050
2;6-2;7	515	3030
2;7-2;8	470	3230
2;9-2;10	530	3320
2;11-3;0	640	3280
Media anual	536	3160

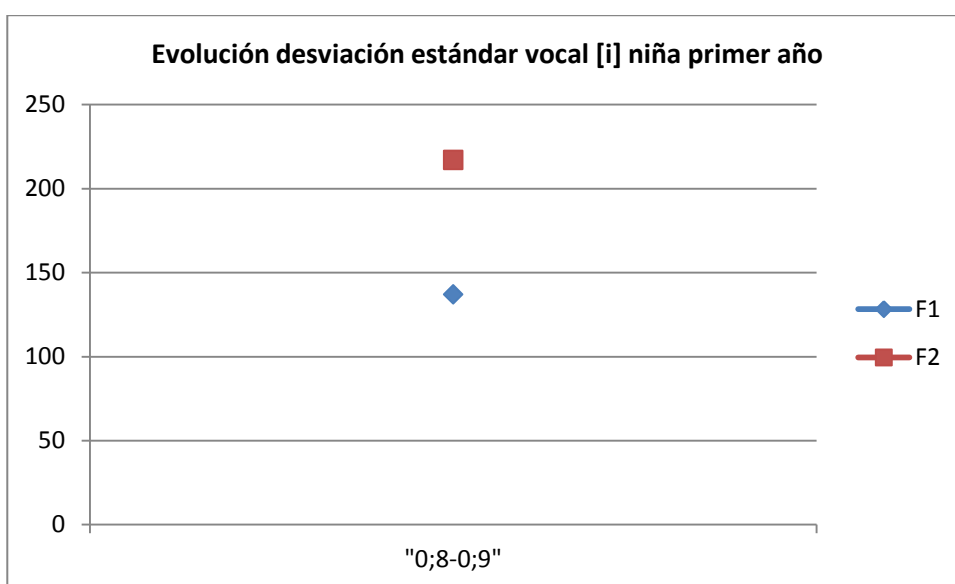
Como podemos observar, el promedio del F1 aumentó 20 Hz entre el primer y segundo año de vida y descendió 93 Hz entre el segundo y el tercero. Respecto al F2, la media del primer año de vida se mantuvo en el segundo (3304 Hz en ambos periodos) y descendió 144 Hz entre el segundo y el tercero. A simple vista ninguno de estos cambios parece muy acentuado pero una vez más decidimos realizar un ANOVA para corroborarlo. El análisis de varianza refrendó nuestras suposiciones aceptándose la hipótesis nula según la cual el promedio no sufrió alteraciones significativas con el paso del tiempo ni en el F1 ni en el F2. De este modo, el descenso experimentado por ambos formantes entre el primer y el tercer año (de 73 Hz en el F1 y de 144 Hz en el F2) no fue relevante. Estos resultados volvieron a asemejar la evolución de los promedios de los dos primeros formantes de la vocal anterior alta [i] con los de la vocal anterior semialta [e], cuyos promedios tampoco variaron significativamente a lo largo del estudio longitudinal.

Nos ocupamos ahora de la desviación estándar que presentaron los valores del F1 y del F2 en la vocal [i] durante el periodo de tiempo que comprendió el estudio. Subrayamos el hecho de que este análisis tan solo pudiera realizarse en cinco de los diecisiete meses en los que esta vocal fue identificada. La razón de tal escasez residió en la circunstancia de que a pesar de que dicha vocal fue reconocida por cuatro de los oyentes en el resto de los meses seleccionados, el número de señales presentes en los distintos periodos fue en su mayoría de uno, resultando

imposible de calcular la desviación que los valores de las distintas íes presentaban con respecto a la media dentro de un mismo mes.

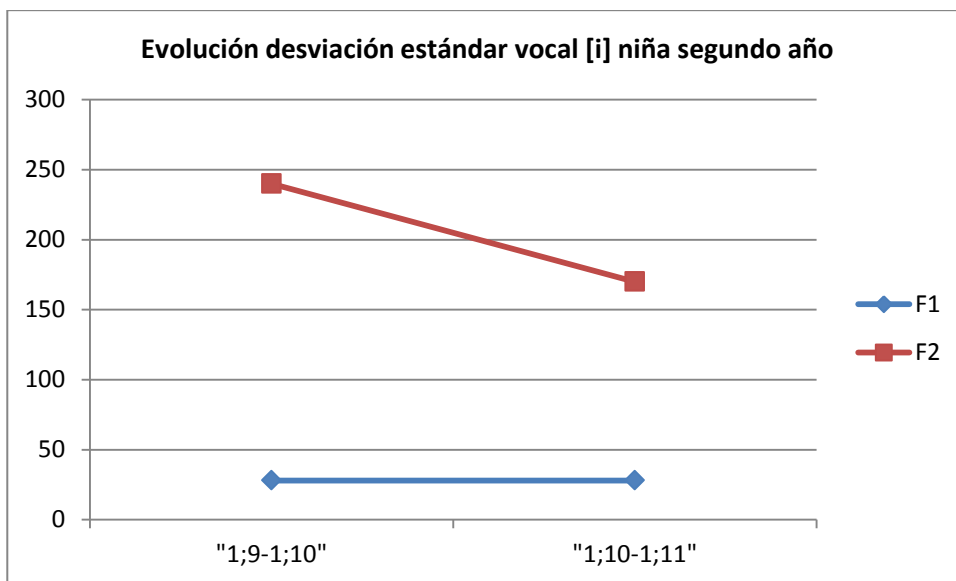
Poco podemos decir de la desviación estándar del primer año de vida puesto que en ella tan solo encontramos un valor, el de 137 Hz para el F1 y el de 217 Hz para el F2 (gráfico 7.20). Estos valores nos servirán de guía para ver en el segundo año si dicha desviación aumenta o disminuye con respecto al primero.

**Gráfico 7.20. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año).**



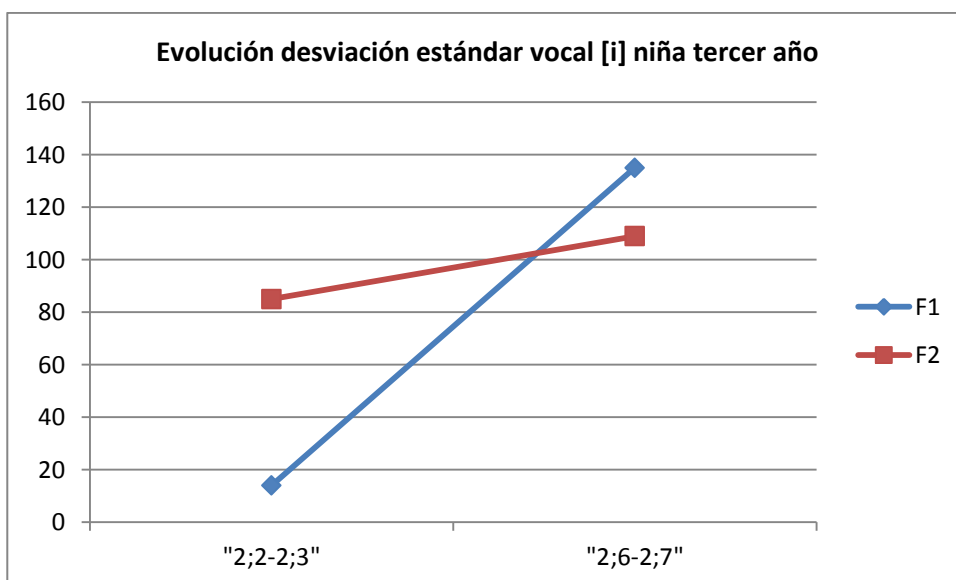
En dos de los cinco meses del segundo año de vida donde fue hallada esta vocal pudo calcularse la desviación estándar (gráfico 7.21). Como podemos observar, la evolución de esta desviación en el F1, que descendió 109 Hz con respecto al mes del periodo anterior, fue lineal a lo largo de este segundo año mientras que la del F2, que aumentó 23 Hz respecto al periodo anterior, descendió en este segundo año de vida (disminución de 70 Hz).

**Gráfico 7.21. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años).**



La evolución de la desviación estándar de los dos primeros formantes en el tercer año de vida presentó un patrón muy similar, tal y como apreciamos en el gráfico 7.22, ya que ambas desviaciones partieron de sus valores mínimos (en el caso del F1 la desviación fue 14 Hz menor que la del último mes del año anterior y en el del F2 la diferencia fue de 85 Hz), y aumentaron en el mes siguiente para registrar la desviación típica mayor.

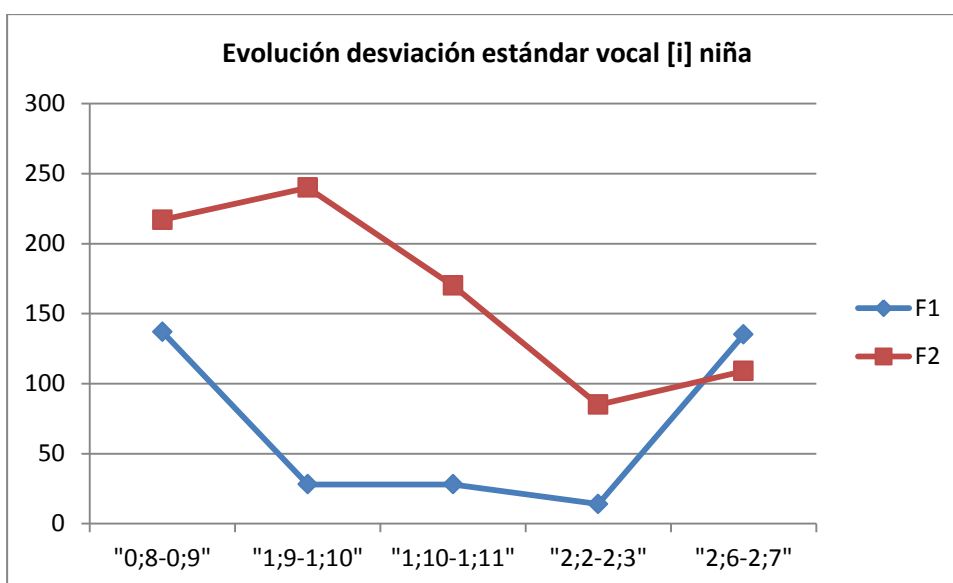
**Gráfico 7.22. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años).**



El gráfico inferior (gráfico 7.23) recoge la evolución de la desviación estándar de los dos primeros formantes a lo largo de los tres años de vida estudiados. Como podemos observar,

el patrón de ambos formantes se opone entre el primer y el segundo año de vida ya que entre los 0;8-0;9 años y los 1;9-1;10 años la desviación del F1 desciende, aunque no llega a registrar su valor mínimo (28 Hz), mientras que la del F2 aumenta para alcanzar su valor máximo (240 Hz). A partir de aquí la evolución de ambas desviaciones es similar puesto que la del F2 desciende progresivamente hasta los 2;2-2;3 años, mes a partir del cual aumenta, y la del F1 se mantiene en el transcurso al mes siguiente pero desciende entre este mes y los 2;2-2;3 años para aumentar como lo hizo la del F2 en el último mes analizado.

**Gráfico 7.23. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en los tres primeros años de vida (0;4-3;0 años).**



A continuación, se exponen los valores concretos de dicha desviación a lo largo de los tres años analizados:

**Tabla 7.16. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;8-0;9	137	217
Desviación estándar anual	136	201

**Tabla 7.17. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;9-1;10	28	240

1;10-1;11	28	170
Desviación estándar anual	49	186

**Tabla 7.18. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
2;2-2;3	14	85
2;6-2;7	135	109
Desviación estándar anual	104	150

A pesar de la escasez de datos, el haber podido calcular al menos una desviación típica por año nos ofreció la posibilidad de observar la evolución de la misma. La tendencia observada en el F1 es algo cambiante puesto que la desviación estándar disminuyó en 87 Hz entre el primer y el segundo año de vida, pero aumentó 55 Hz entre el segundo y el tercero. Los valores de la desviación con respecto a la media fueron no obstante muy similares a los hallados en el F1 de la vocal anterior semialta y a los encontrados en el F1 de la vocal central baja, ambas explicadas anteriormente. Respecto al F2, este tendió a disminuir a lo largo del periodo de estudio cayendo 15 Hz entre el primer y el segundo año de vida, y 36 Hz entre el segundo y el tercero. Siguiendo con la comparación, la evolución de la desviación típica de los dos primeros formantes de la vocal [i] fue muy similar a la hallada para las vocales [a] y [e] ya que la desviación estándar del F2 disminuyó con el paso del tiempo como en estas vocales, y la del F1 se redujo entre el primer y el segundo año de vida y aumentó entre el segundo y el tercero como en el caso de la [a] (en el de la [e] la disminución del F1 fue constante a lo largo del tiempo).

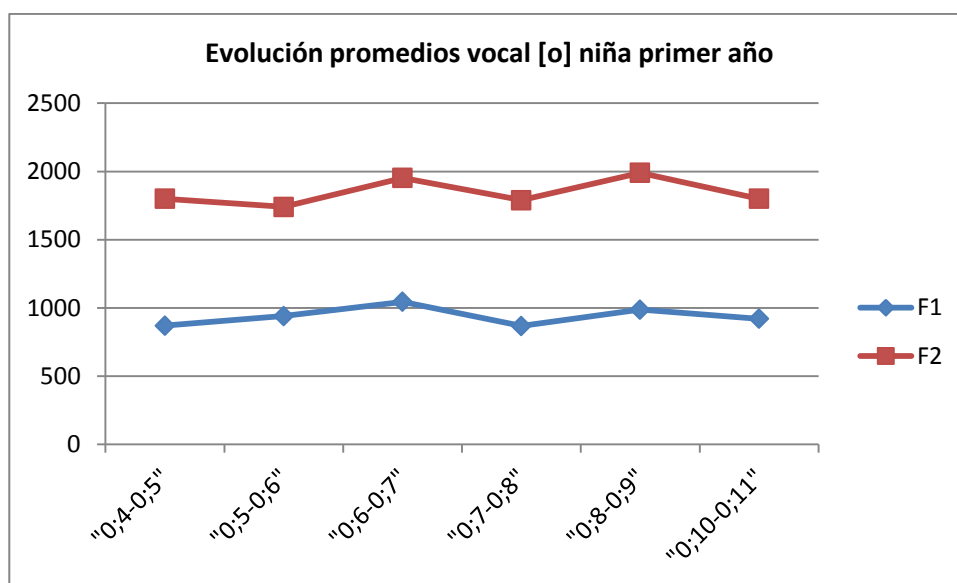
#### ***7.1.1.4. Vocal posterior semialta, [o].***

Continuamos con las vocales posteriores y comenzamos con el cálculo estadístico del promedio de los valores de los dos primeros formantes de la vocal posterior semialta [o] en el primer año de vida (análisis que pudo realizarse en más de la mitad de los meses examinados de este periodo).

Según el gráfico inferior (gráfico 7.24) la evolución de la media del F1 no mostró aumentos ni descensos pronunciados sino que más bien los cambios fueron escalonados y los valores permanecieron en torno a los 900-1000 Hz durante todo el periodo. El promedio más elevado fue alcanzando entre los 0;6 y 0;7 años (1045 Hz) y un mes más tarde se registró el

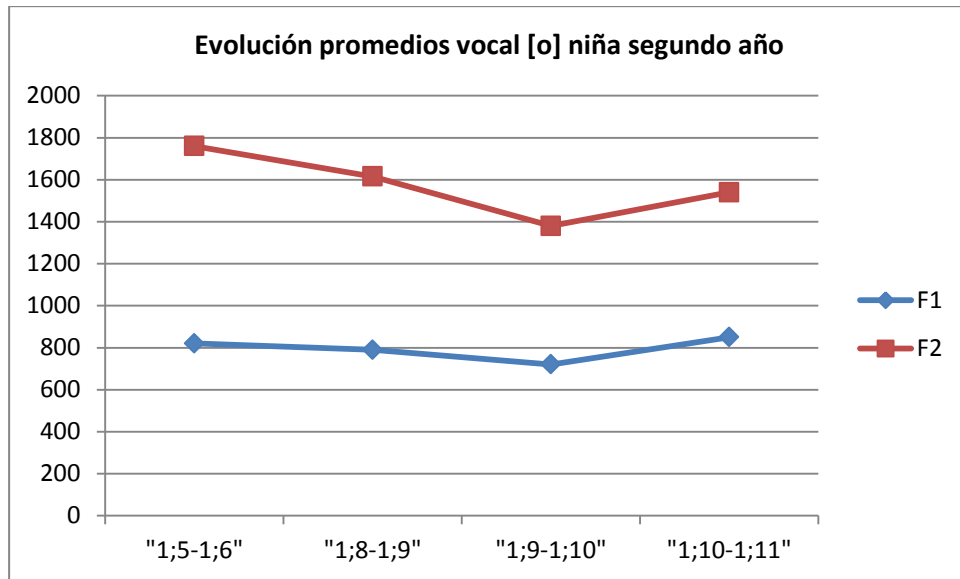
valor mínimo (867 Hz). En relación al F2, el promedio en el primer y en el último mes examinados del periodo fue el mismo (1800 Hz). Los meses intermedios mostraron oscilaciones produciéndose pequeños descensos, como el que tuvo lugar entre los 0;5 y los 0;6 años donde se registró el promedio mínimo (1740 Hz), y sutiles aumentos como el que se produjo entre los 0;8 y los 0;9 años y en el que se alcanzó el valor máximo (1990 Hz).

**Gráfico 7.24. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año).**



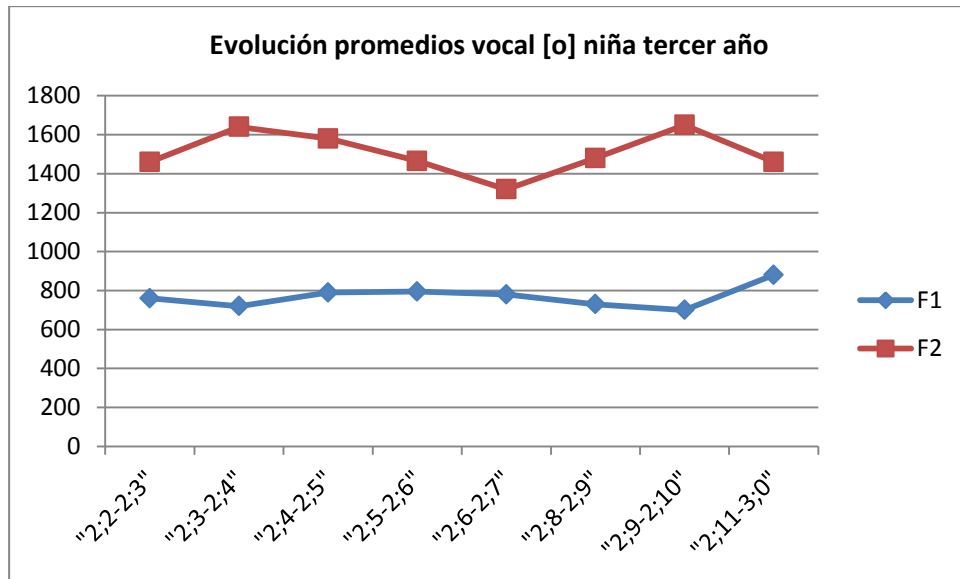
A pesar de presentar el segundo año de vida un mayor número de meses para analizar, tan solo en cuatro de ellos fue reconocida la vocal [o] por cuatro de los transcripores. Tras una rápida ojeada al gráfico 7.25 puede advertirse un patrón evolutivo muy similar en ambos formantes. Respecto al F1, y en comparación con el último mes considerado del primer año de vida, la media del primer formante descendió 100 Hz y siguió reduciéndose ligeramente hasta los 1;10 años, mes en el que registró el valor mínimo (720 Hz). Después de presentar el valor mínimo en este mes, el F1 ascendió alcanzando en el último mes analizado del periodo su valor máximo (850 Hz). De la misma manera, el F2 comenzó el periodo descendiendo en comparación con el último mes del año anterior (disminución de 40 Hz) y continuó haciéndolo hasta los 1;10 años, mes en el que como el F1 registró el promedio mínimo (1380 Hz). No obstante, el descenso del F2 presentó una pendiente mucho más pronunciada que la que manifestó el F1 en ese mismo periodo. En el último mes, el F2 aumentó aproximándose al promedio máximo pero sin alcanzarlo (dicho promedio apareció en el primer mes analizado de este segundo año de vida y fue de 1760 Hz).

**Gráfico 7.25. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años).**



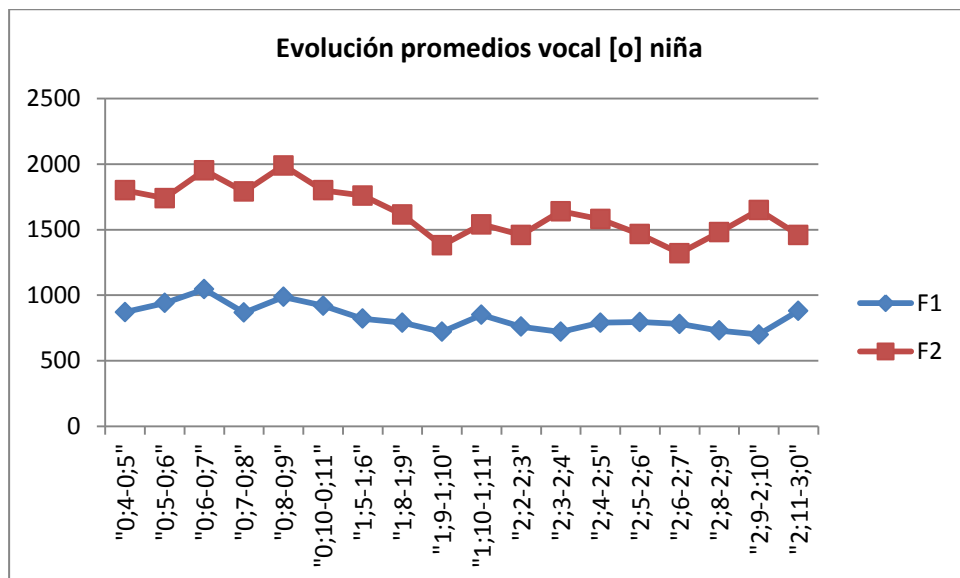
En el tercer año de vida la evolución del promedio de ambos formantes fue inversa (gráfico 7.26). De esta manera, si el F1 (que descendió 90 Hz con respecto al último mes analizado en el periodo anterior) presentó un pequeño descenso en el segundo mes analizado y después experimentó un aumento hasta mediados del periodo, el F2 (que también se redujo entre el último mes del segundo año de vida y el primero del tercero) aumentó entre el primer y el segundo mes analizado y después fue disminuyendo hasta los 2;7 años, mes en el que registró el valor mínimo (1320 Hz). En la segunda mitad del periodo el F1 descendió hasta los 2;10 años, registrándose en este mes el promedio mínimo (700 Hz), y experimentó un ligero ascenso en el último mes analizado (valor máximo de 880 Hz). Por el contrario, el F2 manifestó un ascenso hasta los 2;10 años (alcanzando en este mes el valor máximo) y descendió en el transcurso al último mes analizado.

**Gráfico 7.26. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años).**



En el gráfico siguiente (gráfico 7.27), que recoge la evolución de los promedios de los dos primeros formantes en los tres primeros años de vida, puede contemplarse una disminución tanto en los valores del F1 como en los del F2, si bien el descenso parece ser mucho más acusado en el caso del segundo formante.

**Gráfico 7.27. Evolución de los promedios de los dos primeros formantes de la vocal [o] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Los valores concretos del promedio estadístico para el F1 y F2 quedan recogidos en las tablas inferiores:



**Tabla 7.19. Valores de los promedios de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;4-0;5	870	1800
0;5-0;6	940	1740
0;6-0;7	1045	1953
0;7-0;8	867	1790
0;8-0;9	987	1990
0;10-0;11	920	1800
Media anual	952	1878

**Tabla 7.20. Valores de los promedios de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;5-1;6	820	1760
1;8-1;9	790	1615
1;9-1;10	720	1380
1;10-1;11	850	1540
Media anual	794	1582

**Tabla 7.21. Valores de los promedios de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;2-2;3	760	1460
2;3-2;4	720	1640
2;4-2;5	790	1580
2;5-2;6	795	1465
2;6-2;7	780	1320
2;8-2;9	730	1480
2;9-2;10	700	1650
2;11-3;0	880	1460
Media anual	772	1502

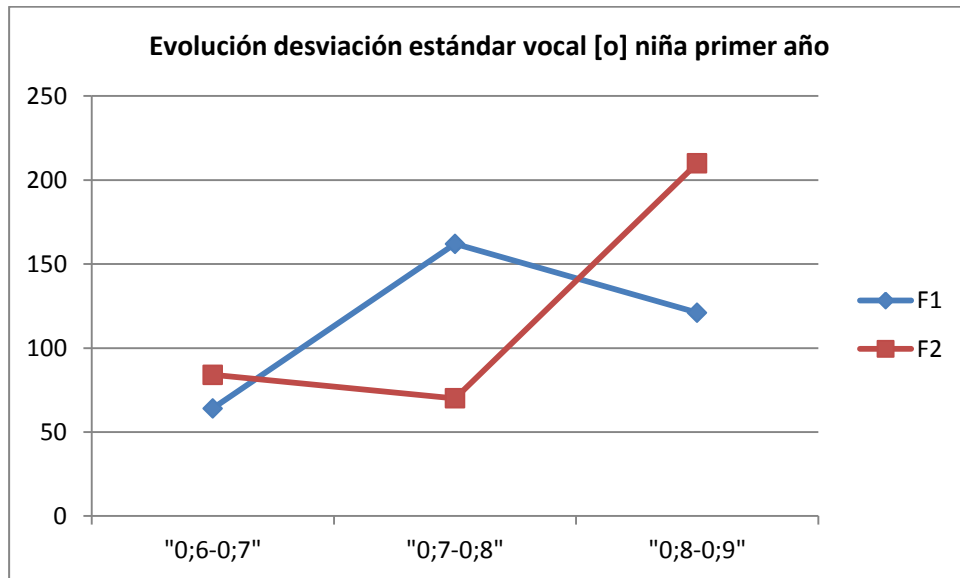
La media del primer formante descendió 158 Hz entre el primer y segundo año de vida y 22 Hz entre el segundo y el tercero (produciéndose una disminución de 180 Hz entre el primer y el tercer año). Por otro lado, el promedio del F2 disminuyó 296 Hz entre el primer y el segundo año y 80 Hz entre el segundo y el tercero, reduciéndose entre el primer y el tercer periodo 376 Hz. Los valores del F1 no descendieron significativamente entre los cuatro primeros meses analizados del primer año de vida y los cuatro últimos pero a diferencia de lo ocurrido en el resto de vocales analizadas, el análisis estadístico evidenció que un descenso significativo se producía en el F1 tras comparar los valores del primer año de vida con los del segundo. No obstante, y si bien a simple vista parece que el descenso significativo del F1 tiene lugar entre los 1;9 y 1;10 años ya que la media de este periodo es de 720 Hz frente a los 952 Hz del primer año de vida, es decir hay 232 Hz de diferencia, al hacer el cálculo estadístico entre la media del F1 del primer año de vida y la de los distintos meses del segundo año para ver en cuál de ellos se produce el descenso significativo, el análisis de varianza no señala ninguno de estos meses debido al distinto número de señales que nos encontramos al comparar el conjunto de meses del primer año de vida, que tiene en total doce muestras, con cada uno de los meses analizados en el segundo en los que solo encontramos una o dos muestras. Un descenso también significativo se observó tras comparar los valores del primer año con los del tercero mientras que tras equipararse los valores del segundo año con los del tercero se halló que estos eran similares.

En cuanto al F2, el ANOVA manifestó que los valores de los cuatro primeros y los cuatro últimos meses analizados en el primer año de vida eran muy similares pero que un descenso significativo de los mismos se producía tras comparar los valores del primer año con los del segundo y que dicha disminución tenía lugar entre los 1;8 y los 1;9 años. Dichos valores no sufrieron cambios pertinentes cuando se comparó el segundo con el tercer año de vida pero, y al igual que ocurrió con el F1, sí se produjo un descenso significativo entre los valores del primer y tercer año analizados. Los resultados del F2 son similares a los encontrados en la vocal central [a] en la que un descenso significativo del F2 tuvo lugar a partir del mes diecisiete (entre los 1;5 y los 1;6 años).

Exponemos a continuación los datos del segundo análisis estadístico que en el primer año de vida pudo realizarse en tres meses (gráfico 7.28). Tal y como muestra el gráfico, la evolución de la desviación estándar de los dos primeros formantes es opuesta. De esta manera, mientras que el F1 parte de la desviación más baja (64 Hz), aumenta en el mes siguiente hasta alcanzar la desviación estándar máxima (162 Hz) y vuelve a descender en el último mes en que

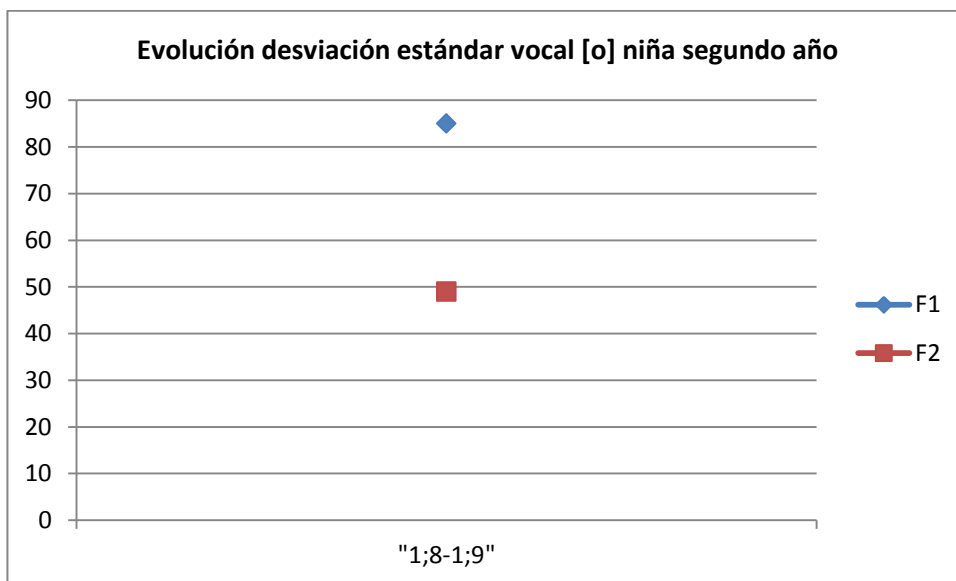
pudo calcularse, la desviación del F2 desciende entre el primer y el segundo mes analizados, registrando entre los 0;7 y los 0;8 años su valor mínimo (70 Hz), y aumenta en el último mes del periodo alcanzando en este su valor máximo (210 Hz).

**Gráfico 7.28. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año).**



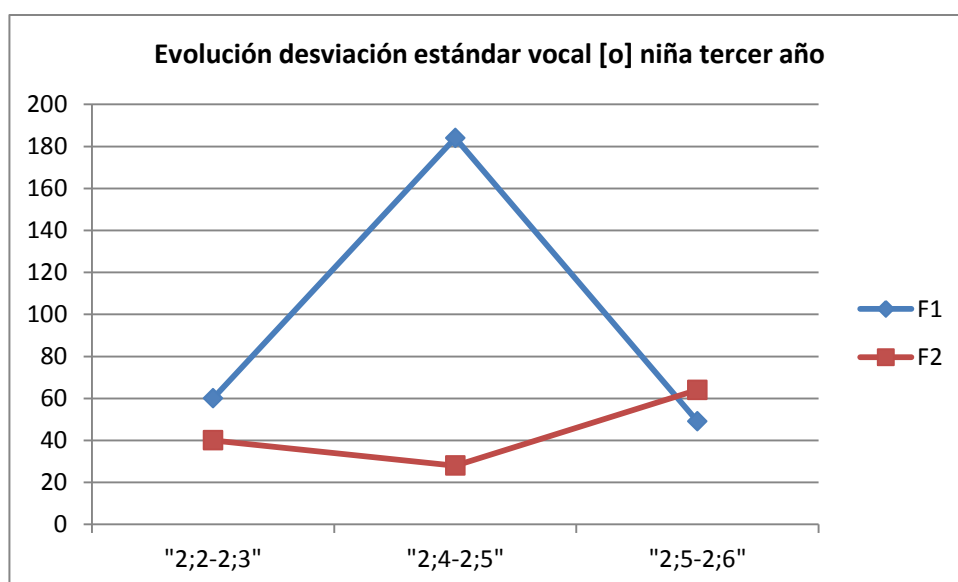
Tan solo en un mes del segundo año de vida pudo realizarse el análisis de la desviación estándar (gráfico 7.29). En cuanto al F1, este descendió 36 Hz con respecto al último mes del primer periodo analizado y respecto al F2 este disminuyó 161 Hz tras comparar estos dos últimos meses.

**Gráfico 7.29. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años).**



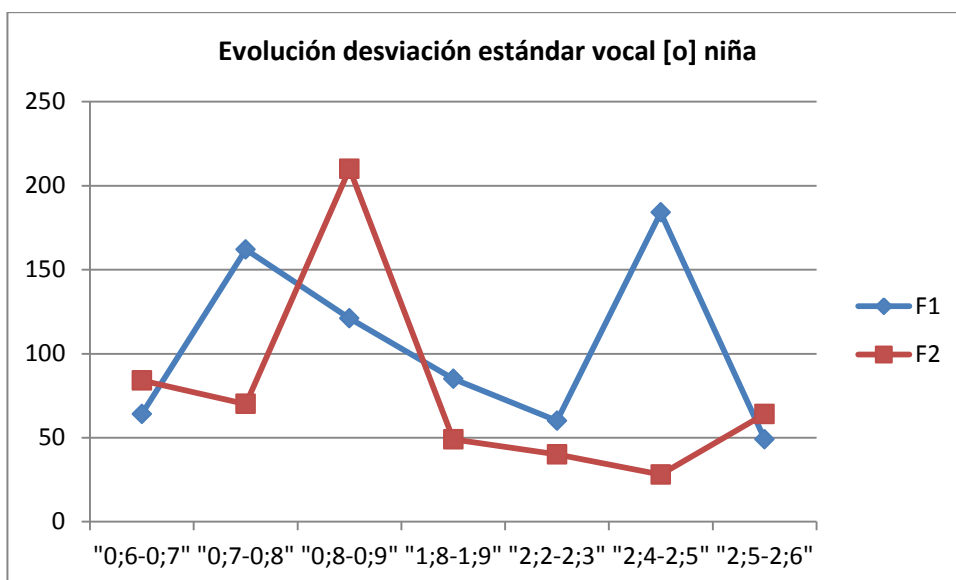
La evolución de la desviación estándar de ambos formantes en el tercer año de vida presentó un patrón muy similar al de la evolución del primer año siendo dichas evoluciones una vez más contrarias (gráfico 7.30). Así, la desviación del F1 aumentó entre el primer y el segundo mes en los que se calculó, alcanzándose la desviación estándar máxima entre los 2;4 y 2;5 años (184 Hz), y descendió en el último mes analizado registrando en este su valor mínimo (49 Hz). Por otro lado, la desviación del F2 disminuyó entre el primer y el segundo mes analizados, registrando en este su valor mínimo (28 Hz), y aumentó entre el segundo y el tercero, alcanzando en este último el valor máximo (64 Hz). La desviación del F1 descendió 25 Hz entre el último mes del segundo año de vida y el primero del tercero, y la del F2 disminuyó 9 Hz.

**Gráfico 7.30. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años).**



El gráfico siguiente recoge la evolución de la desviación estándar de los dos primeros formantes en los distintos meses de los tres periodos estudiados (gráfico 7.31). Como podemos observar y como hemos venido apuntando en los gráficos anteriores, la evolución de la desviación estándar de ambos formantes es opuesta durante el primer año de vida, coincide en el transcurso del primer año al segundo y del segundo al primer mes del tercero (ya que la desviación de ambos formantes tiende a disminuir) y vuelve a oponerse en el tercer año.

**Gráfico 7.31. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Las tablas siguientes muestran los meses de los tres primeros años de vida en los que la desviación típica pudo calcularse, así como la desviación estándar anual.

**Tabla 7.22. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;6-0;7	64	84
0;7-0;8	162	70
0;8-0;9	121	210
Desviación estándar anual	116	142

**Tabla 7.23. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;8-1;9	85	49
Desviación estándar anual	64	140

**Tabla 7.24. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña en el tercer año de vida (2;0-3;0 años).**

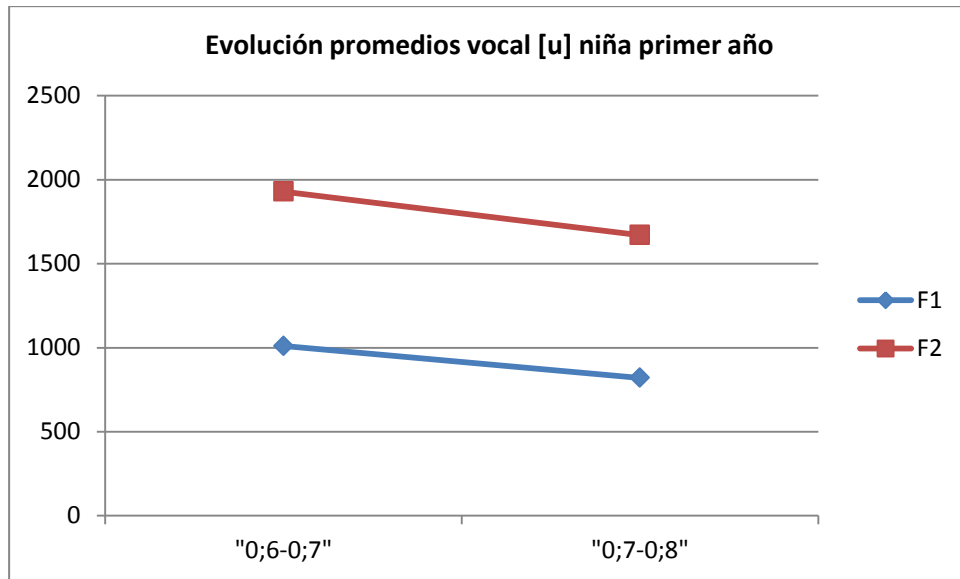
Edad	Desviación estándar F1	Desviación estándar F2
2;2-2;3	60	40
2;4-2;5	184	28
2;5-2;6	49	64
Desviación estándar anual	78	97

Un descenso de 52 Hz se produjo en la desviación típica del F1 entre el primer y el segundo año de vida y un ligero aumento tuvo lugar entre el segundo y el tercero (14 Hz). Respecto al F2, la desviación estándar se mantuvo prácticamente igual entre el primer y el segundo año puesto que la media fue de 142 Hz en el primero y de 140 Hz en el segundo, y disminuyó 43 Hz entre el segundo y el tercero produciéndose por tanto una disminución de 45 Hz entre el primer y el tercer año de vida. Si comparamos estos datos con los obtenidos en las vocales anteriormente analizadas, encontramos que el grado de apartamiento de la media de los valores del F1 en la vocal [o] fue levemente inferior al hallado en el resto de vocales ya que los valores arrojados tras el cálculo de la desviación estándar son menores, si bien la diferencia es mínima, mientras que los del F2 son muy similares.

#### ***7.1.1.5. Vocal posterior alta, [u].***

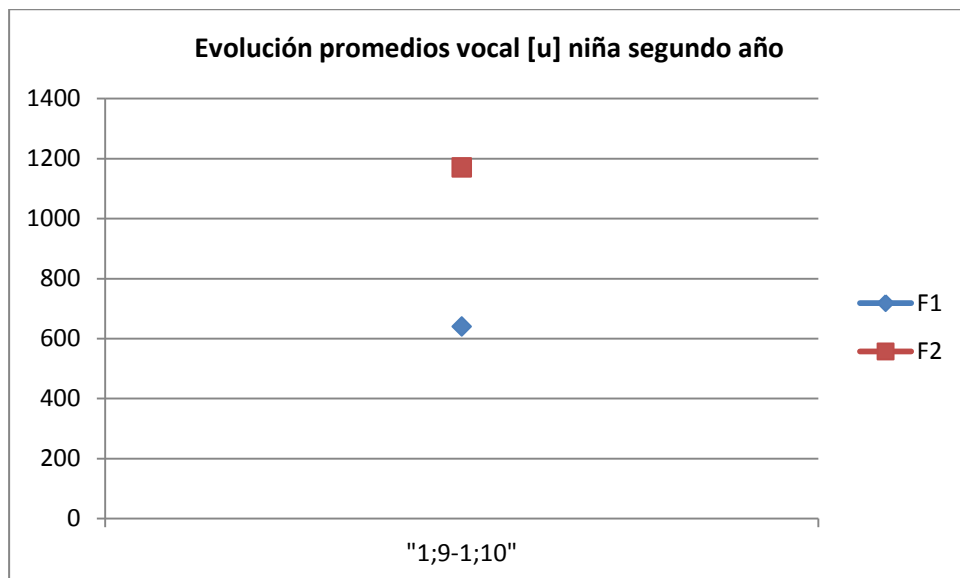
Tal y como expusimos en las tablas del apartado 6.2.1, la vocal posterior alta [u] fue identificada por cuatro de los transcritores en solo siete de los meses examinados de forma que el número de datos sobre el promedio y la desviación estándar de los valores es muy reducido en comparación con el del resto de vocales. En el gráfico que recoge la evolución de la media durante el primer año de vida (gráfico 7.32) observamos que tanto el F1 como el F2 descienden entre los 0;6 y los 0;8 años.

**Gráfico 7.32. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en el primer año de vida (0;4-1;0 año).**



En el transcurso hacia el segundo año de vida tan solo contamos con el promedio obtenido en uno de los meses (gráfico 7.33). Respecto al año anterior, encontramos que el F1 descendió 180 Hz en comparación con el último mes analizado del primer año de vida y que el F2 lo hizo en 500 Hz.

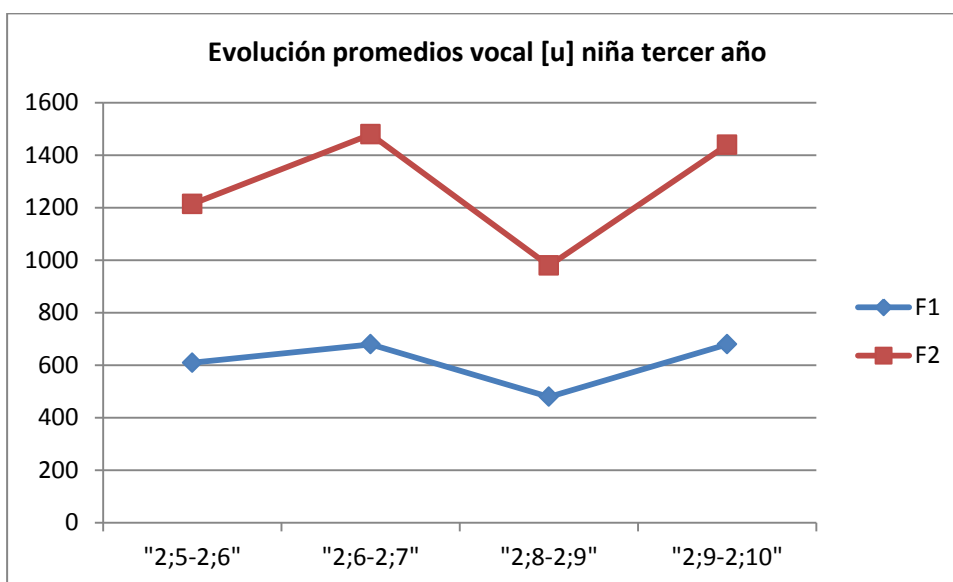
**Gráfico 7.33. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en el segundo año de vida (1;0-2;0 años).**



En el intervalo de tiempo que hay entre el segundo y el tercer año de vida, la evolución del promedio de ambos formantes muestra un patrón prácticamente idéntico si bien los aumentos y descensos del F2 son mucho más pronunciados que los del F1 (gráfico 7.34). En

cuanto a este primer formante, encontramos que un descenso de 30 Hz se produjo entre el último mes del segundo año de vida y el primero del tercero. A un suave ascenso en el que la vocal alcanzó el promedio máximo (680 Hz), le siguió un leve descenso en el que se registró el valor mínimo (480 Hz), y un nuevo aumento estuvo presente en último mes analizado en el que el promedio fue el mismo que el valor máximo alcanzado tres meses antes. La evolución del segundo formante fue exactamente la misma si bien las subidas y bajadas del promedio fueron mucho más notorias. Este formante comenzó con un aumento de 45 Hz con respecto al promedio del último mes analizado del periodo anterior, alcanzó su máximo en el mes siguiente (1480 Hz) y el mínimo después de este (980 Hz).

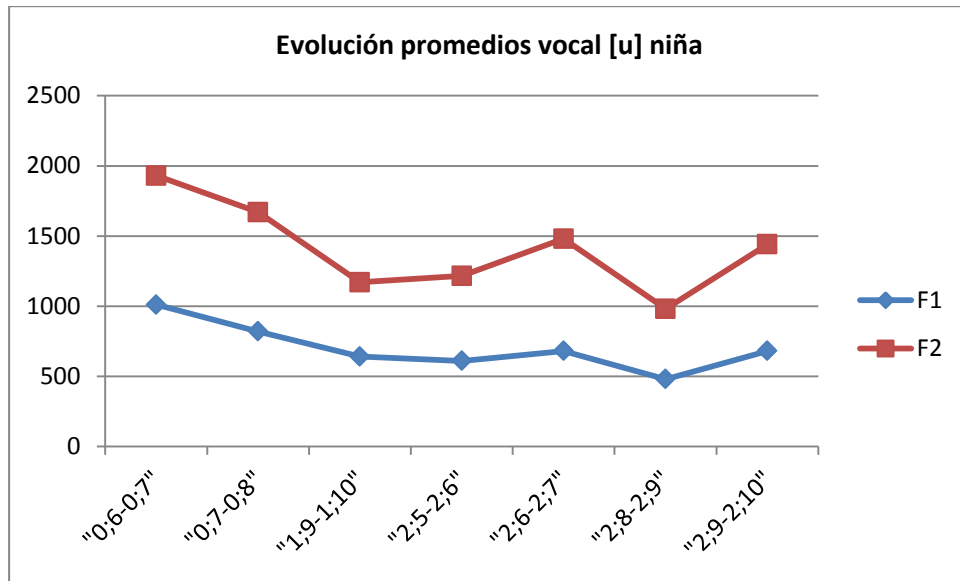
**Gráfico 7.34. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en el tercer año de vida (2;0-3;0 años).**



La evolución del promedio del F1 y del F2 en los tres primeros años de vida se presenta en el gráfico posterior (gráfico 7.35) y en este queda muy bien reflejado que a partir del primer mes analizado los promedios de los dos primeros formantes descienden y no vuelven a recuperarse en los meses que restan hasta el final del periodo. Por otro lado, las subidas y bajadas en la media del F1 se muestran mucho más escalonadas aunque la tendencia de este primer formante también es hacia el descenso.



**Gráfico 7.35. Evolución de los promedios de los dos primeros formantes de la vocal [u] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Los promedios obtenidos en cada uno de los intervalos temporales analizados se muestran en las siguientes tablas:

**Tabla 7.25. Valores de los promedios de los dos primeros formantes de la vocal [u] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;6-0;7	1010	1930
0;7-0;8	820	1670
Media anual	915	1800

**Tabla 7.26. Valores de los promedios de los dos primeros formantes de la vocal [u] de la niña en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;9-1;10	640	1170
Media anual	640	1170

**Tabla 7.27. Valores de los promedios de los dos primeros formantes de la vocal [u] de la niña en el tercer año de vida (2;0-3;0 años).**

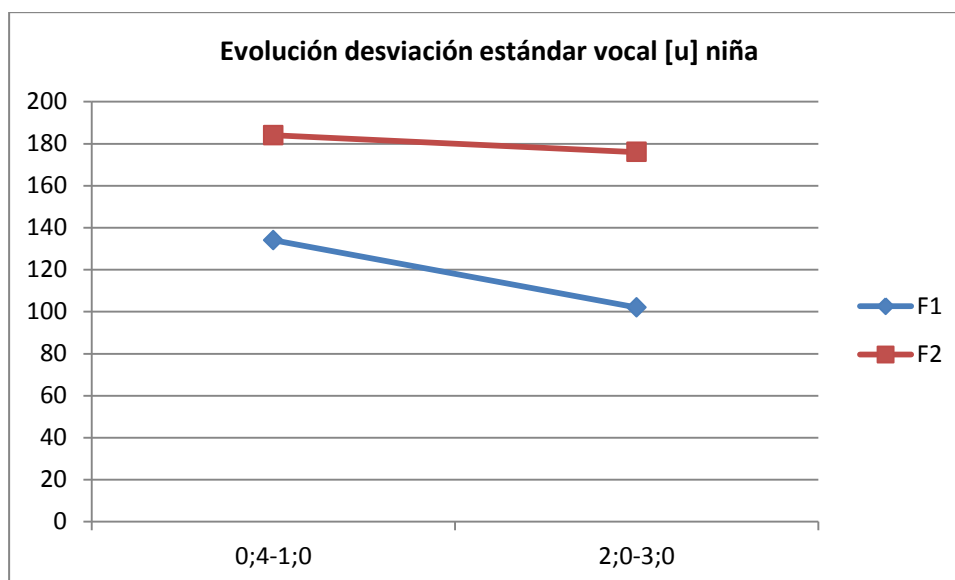
Edad	Media F1	Media F2
2;5-2;6	610	1215

2;6-2;7	680	1480
2;8-2;9	480	980
2;9-2;10	680	1440
Media anual	611	1251

La media del F1 se redujo 315 Hz con el paso del tiempo disminuyendo 275 Hz entre el primer y el segundo periodo analizados y 29 Hz entre el segundo y el tercero. El promedio del F2 también se redujo entre el primer y el segundo año de vida (disminución de 630 Hz) mientras que entre el segundo y el tercero este aumentó 81 Hz por lo que la distancia entre el primer y el tercer periodo examinados fue de 549 Hz. En cuanto al análisis de varianza, en esta ocasión fue imposible realizar el ANOVA para comprobar si un descenso significativo tenía lugar en el F1 y en el F2 entre el primer y el segundo año de vida debido a la presencia de solo dos señales en el primer año y de tan solo una en el segundo. No obstante, la diferencia en el F1 de 275 Hz entre el primer y el segundo año de vida, y la diferencia en el F2 de 630 Hz en el mismo periodo parece indicar que un descenso significativo tuvo lugar entre estos periodos (concretamente entre los 1;9 y 1;10 años). Sí pudo realizarse el ANOVA entre el segundo y el tercer año y este mostró que ni el F1 ni el F2 descendían de manera significativa (el promedio del F2 aumentó incluso), mientras que tras realizarlo entre el primer y el tercer año un descenso significativo asomó en ambos formantes (el ya señalado a los 1;10 años).

En relación a la desviación estándar y debido a que tan solo el mes que transcurrió entre los 2;5 y los 2;6 años presentó más de una vocal posterior alta redondeada que permitió que este análisis estadístico pudiera ser calculado, creemos conveniente presentar un solo gráfico en el que pueda observarse la desviación típica anual y no la mensual (desviación que además para esta vocal tan solo pudo calcularse en el primer y el tercer periodo analizados). Como podemos observar en el gráfico inferior (gráfico 7.36), la tendencia de la desviación típica de ambos formantes es hacia el descenso si bien la disminución es algo mayor en el caso del F1 (32 Hz) que en el del F2 (8 Hz).

**Gráfico 7.36. Evolución de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña en los tres primeros años de vida (0;4-3;0 años).**



Los valores exactos de los datos señalados en el gráfico lineal aparecen en las siguientes tablas:

**Tabla 7.28. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
Desviación estándar anual	134	184

**Tabla 7.29. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña en el tercer año de vida (2;0-3;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
"2;5-2;6"	109	82
Desviación estándar anual	102	176

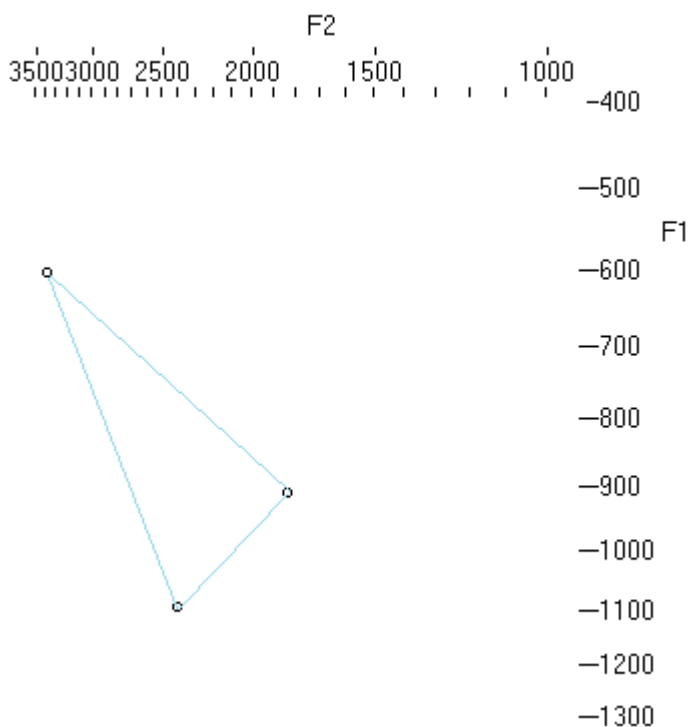
Comprobamos que entre el primer y el tercer periodo estudiados la desviación estándar del F1 descendió 32 Hz y la del F2 8 Hz. A pesar de que hay que tener en cuenta la ausencia del grado de apartamiento de los valores del F1 y F2 con respecto a la media en el periodo intermedio de este estudio (el segundo año de vida), podemos observar que los valores de la desviación típica tanto del F1 como del F2 se asemejaron a los encontrados en las vocales anteriormente analizadas.

### 7.1.1.6. Evolución del área vocálica a partir de los valores obtenidos en las señales de la niña.

Dedicaremos este último apartado a la representación del área vocálica que quedará dibujada a partir de los promedios de los formantes de las tres vocales de los extremos del español, a saber [i], [a] y [u], en cada uno de los periodos estudiados. Nuestro objetivo es el de observar si se producen cambios llamativos en la misma a lo largo del tiempo o si por el contrario esta no presenta variaciones importantes. Para ello crearemos gráficos en los que tras marcar los promedios formánticos de las vocales mencionadas, los uniremos mediante líneas para delimitar el área vocálica que ha de presentar un aspecto triangular al ser tres las vocales consideradas.

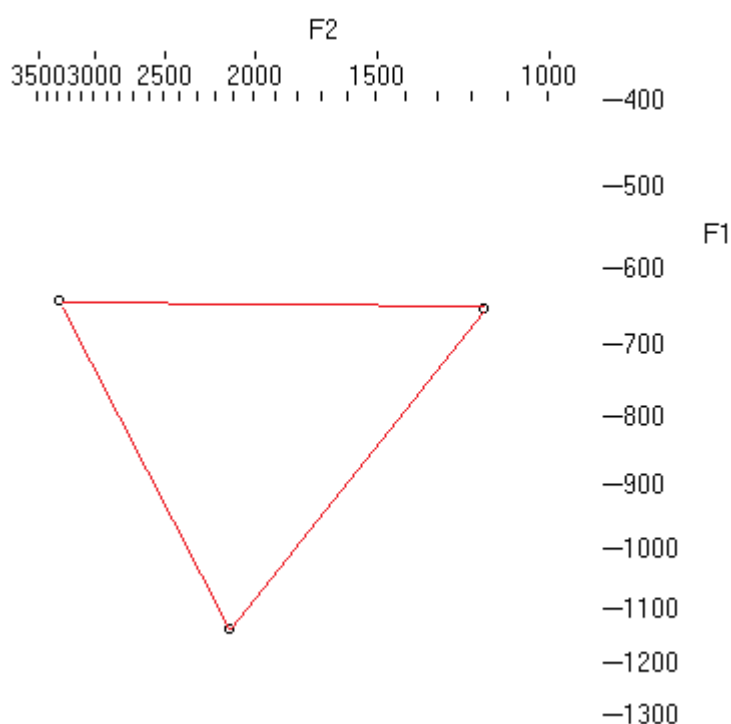
En el primer año de vida analizado (gráfico 7.37), el área vocálica recreó la silueta de un triángulo vocálico que se situó a la izquierda del gráfico debido a los elevados segundos formantes de las tres vocales de los extremos (frecuencias que sobrepasaron los 3000 Hz como en el caso de la [i], se acercaron a los 2500 Hz como en el de la [a] y fueron de 1800 Hz como en el de la [u]). Por otro lado, su ubicación en la parte baja del gráfico se debe a unos primeros formantes que también presentaron valores altos, valores que sobrepasaron los 500 Hz como en la [i], superaron los 1000 Hz como en la [a], y se situaron en los 900 Hz como en la [u].

Gráfico 7.37. Área vocálica de la niña en el primer año de vida (0;4-1;0 año).



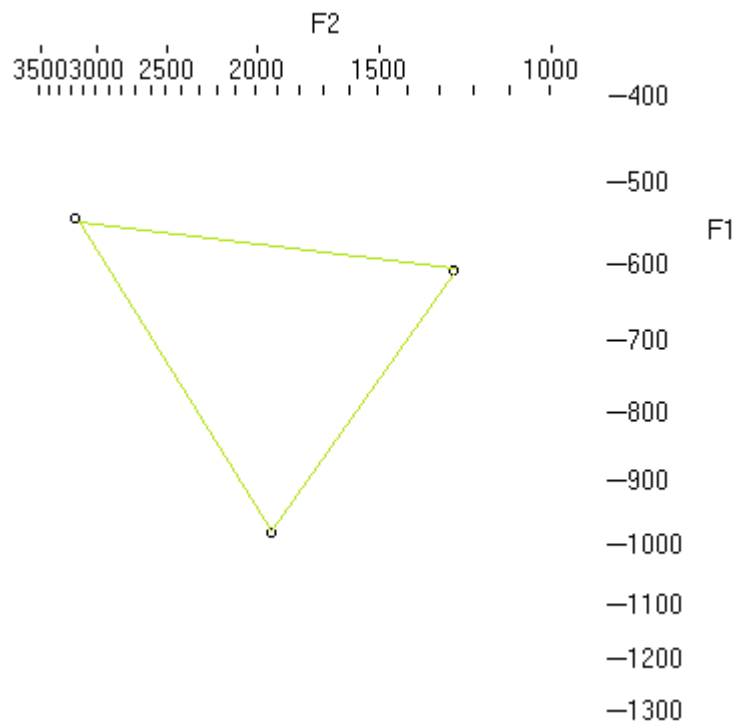
De nuevo el perfil de un triángulo delimitó el área vocálica de la niña en el segundo periodo examinado (gráfico 7.38). La reducción significativa de los valores de los segundos formantes de las vocales [a] y [u], sobre todo de esta última, fue la responsable de que el triángulo vocálico se encontrara más centrado en el gráfico en el transcurso del primer al segundo año de vida. Sin embargo, este desplazamiento de las vocales [u] y [a] hacia la parte derecha del gráfico en conjunción con la ausencia de un descenso significativo del F2 en la [i] (tal y como se comprobó en el apartado 7.1.1.3) provocaron que el área triangular aumentara en este segundo año de vida pasando de los 226903 del primer periodo a los 517429 del segundo. Respecto a la altura del triángulo, el triángulo siguió manteniéndose en la parte baja del gráfico.

**Gráfico 7.38. Área vocálica de la niña en el segundo año de vida (1;0-2;0 años).**



La circunstancia anterior se repitió en este tercer periodo (gráfico 7.39) de forma que la similitud en los promedios del segundo formante de la [i] en el primer, segundo y tercer año de vida, y la notoria disminución de dicho formante que tuvo lugar en las vocales [a] y [u] en el segundo periodo analizado y que se ha mantenido en el tercero, delinearon un triángulo vocálico que se alargó y se estrechó en la parte izquierda del gráfico, como si los promedios de la vocal anterior alta estuvieran estirando de él, y cuya área vocálica fue mayor que la hallada en el primer año estudiado (367566). La sutil reducción de los primeros formantes de todas las vocales en este tercer año de vida hizo que dicho triángulo no se localizara tan abajo en la carta formántica.

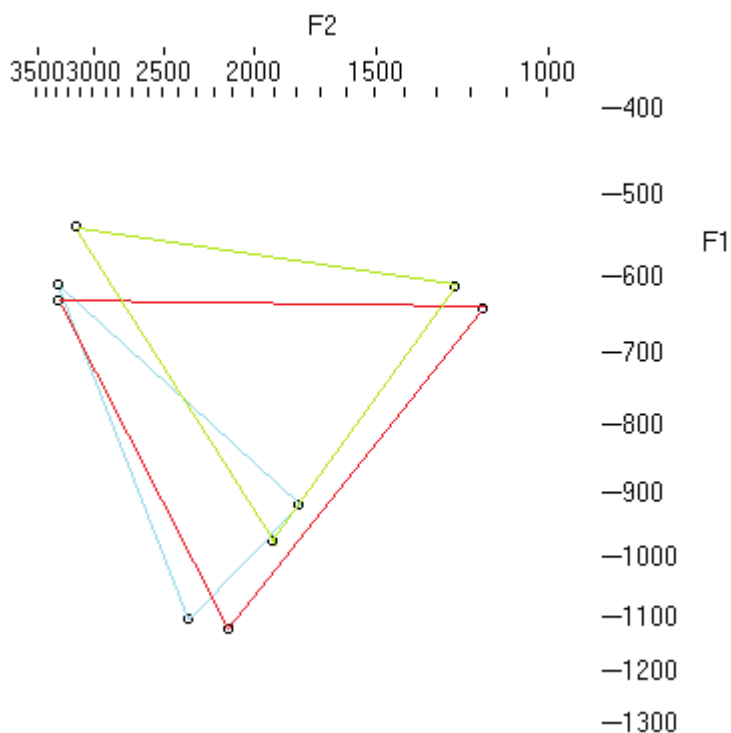
**Gráfico 7.39. Área vocálica de la niña en el tercer año de vida (2;0-3;0 años).**



Si centramos nuestra atención en el último gráfico que recoge la evolución del área vocálica en los tres periodos estudiados, podremos extraer diferentes conclusiones (gráfico 7.40). Por un lado y en cuanto a la ubicación de las áreas vocálicas, observamos que en el segundo y tercer año de vida dicha área vocálica siguió situándose en la parte izquierda del gráfico aunque la reducción significativa del segundo formante de la [a] y de la [u] en el segundo año de vida, y que se mantuvo en el tercero, provocó que dicho triángulo se centrara algo más en la carta de formantes. Además y debido a la reducción posiblemente significativa del F1 de la [u] en el segundo año de vida (recordemos que no pudo realizarse el ANOVA), el alargamiento que presenta el triángulo del primer año de vida, producido por un alto primer formante de la vocal [u], desaparece en el segundo y tercer año al acercarse los valores de este primer formante de la [u] a los de la [i]. Cabe destacar también que al no disminuir el promedio del F2 de la vocal anterior alta de manera significativa en ninguno de los periodos estudiados y al sí hacerlo el de la vocal central baja y el de la vocal posterior alta, el área vocálica no solo no disminuyó con el paso del tiempo sino que aumentó entre el primer año de vida y el segundo y entre el primero y el tercero (porque los valores de la [i] se mantuvieron en el mismo punto mientras que los de la [a] y la [u] se movieron hacia la derecha, ampliándose por tanto el área del triángulo). Por otro lado, el aumento de los promedios del primer formante en las vocales [i] y [a] en el segundo periodo analizado provocó que el área del triángulo del segundo año de vida resultara incluso mayor que la del primero. Finalmente, la reducción de los valores del F1 de

todas las vocales entre el segundo y el tercer año de vida (reducción que como vimos no fue significativa), conllevó que el área vocálica se estrechara en el tercer periodo analizado aunque esta estrechez no impidió que el área del mismo siguiera siendo mayor que la del primer periodo examinado.

**Gráfico 7.40. Evolución del área vocálica de la niña en los tres primeros años de vida (0;4-3;0 años).**



### 7.1.2. Evolución de la media y de la desviación estándar de los dos primeros formantes en las vocales del niño.

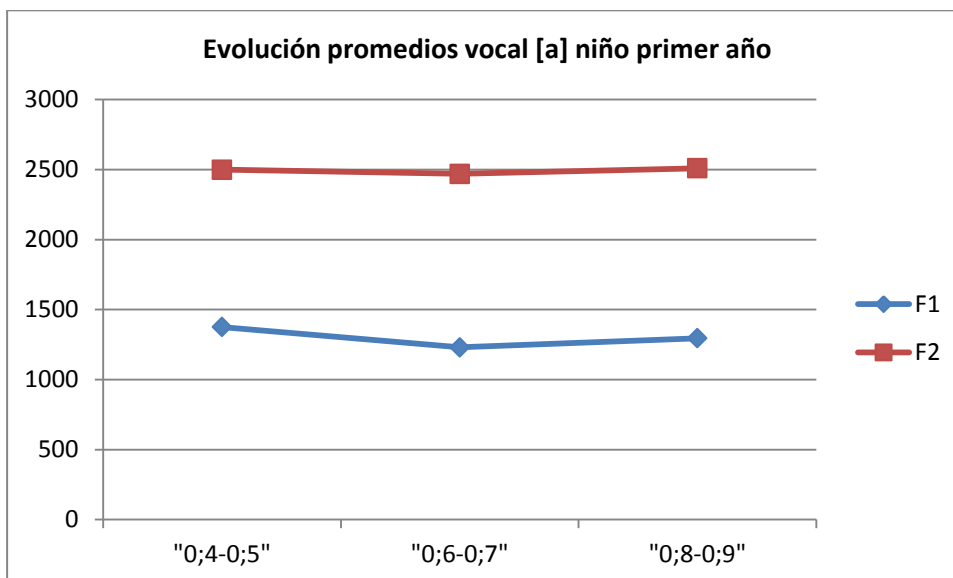
#### 7.1.2.1. Vocal central baja, [a].

Antes de comenzar con la descripción de la evolución formántica que la vocal central baja presentó durante el periodo de estudio, debemos indicar que los transcriptores solo coincidieron en su identificación en los dos primeros años de vida, quedando el tercero por tanto sin representación alguna de esta vocal.

En el gráfico que exhibe la evolución de la media del F1 y del F2 de la vocal central baja a lo largo del primer año de vida (gráfico 7.41), llaman la atención dos hechos: por un lado que esta tan solo pudo extraerse en tres de los ocho meses analizados en dicho periodo, y por

otro, que en tan reducido número de meses el promedio presentó una evolución muy similar ya que descendió en un primer momento y volvió a incrementarse al final, si bien ambos movimientos fueron muy sutiles. La única diferencia es que el primer formante presentó el promedio más alto en el primer mes mientras que el segundo lo hizo en el postrero.

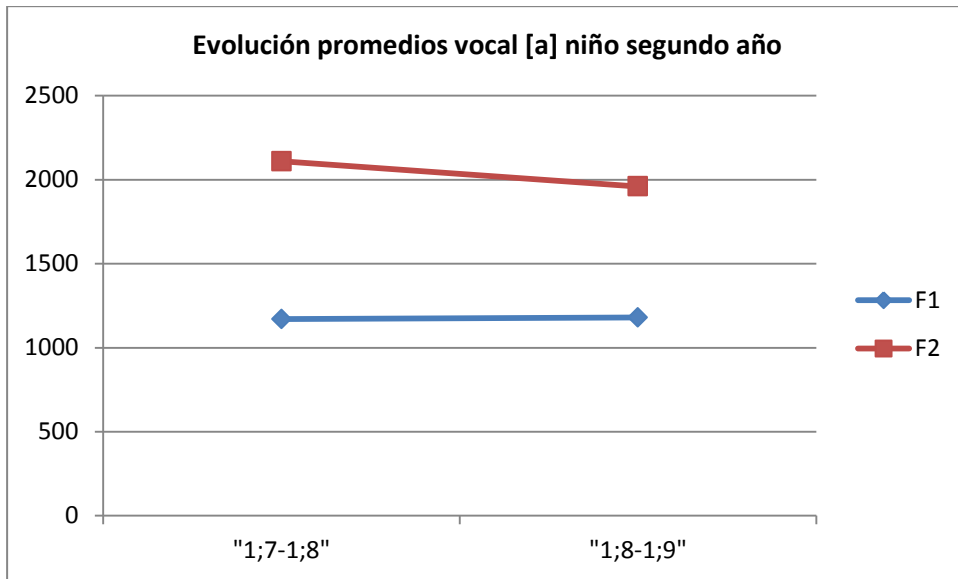
**Gráfico 7.41. Evolución de los promedios de los primeros formantes de la vocal [a] del niño en el primer año de vida (0;4-1;0 año).**



En el segundo intervalo de tiempo analizado (gráfico 7.42), fueron dos los meses en los que cuatro de los cinco transcriptores identificaron la vocal central baja y es por ello que tan solo contamos con dos datos del promedio formántico. En relación al F1, la media de este descendió 125 Hz respecto al último mes del primer periodo analizado y se incrementó ligeramente entre los 1;8 y los 1;9 años de vida aunque debido a la levedad de dicho aumento (10 Hz), este apenas puede visualizarse en el gráfico. En cuanto al F2, una brusca disminución de 400 Hz tuvo lugar entre el último mes examinado del primer año de vida y el primero del segundo. El descenso continuó en el transcurso hacia los 1;9 años aunque en esta ocasión este fue inferior (de 150 Hz).

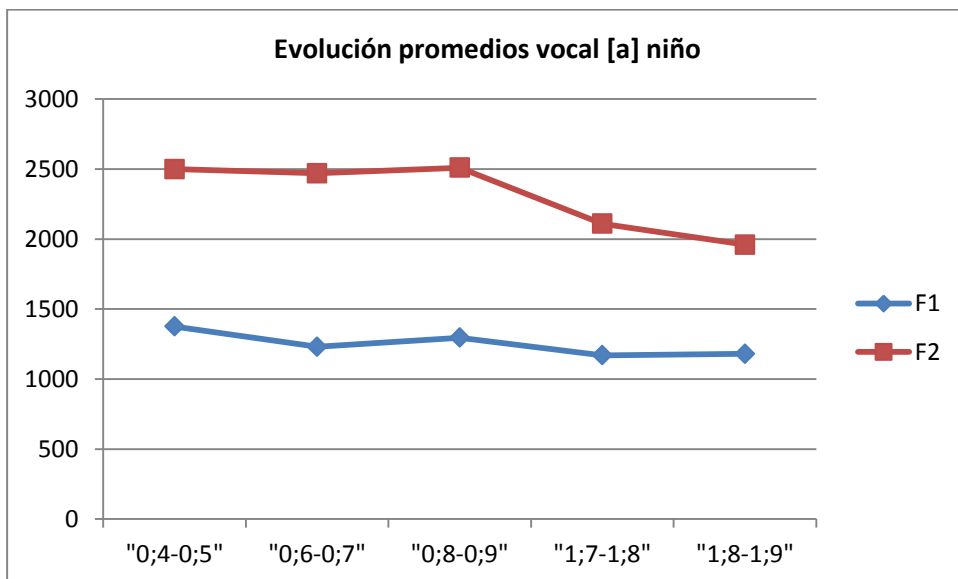


**Gráfico 7.42. Evolución de los promedios de los primeros formantes de la vocal [a] del niño en el segundo año de vida (1;0-2;0 años).**



La constancia en los valores del F1 y la disminución de los valores del F2 entre el primer y segundo periodo pueden observarse conjuntamente en el gráfico inferior que recoge la evolución de ambos formantes en los dos primeros años de vida (gráfico 7.43):

**Gráfico 7.43. Evolución de los promedios de los primeros formantes de la vocal [a] del niño en los dos primeros años de vida (0;4-2;0 años).**



El hecho de que los transcritores no coincidieran en el reconocimiento de la vocal [a] en ninguno de los meses que componen el tercer año de vida, nos lleva directamente a mostrar las tablas que recogen los valores exactos de los promedios formánticos en los dos primeros años del estudio.

**Tabla 7.30. Valores de los promedios de los dos primeros formantes de la vocal [a] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;4-0;5	1375	2500
0;6-0;7	1230	2470
0;8-0;9	1295	2510
Media anual	1299	2498

**Tabla 7.31. Valores de los promedios de los dos primeros formantes de la vocal [a] del niño en el segundo año de vida (1;0-2;0 años).**

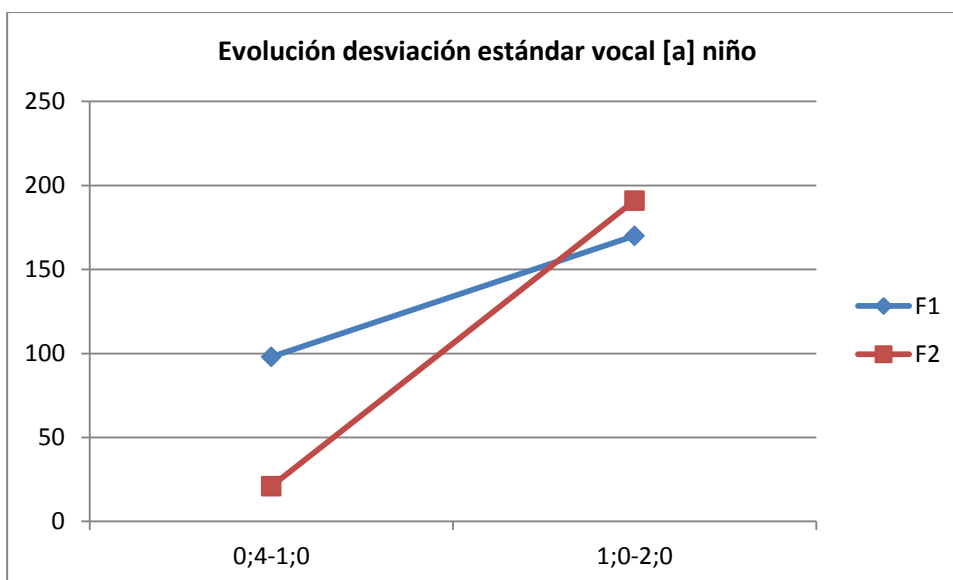
Edad	Media F1	Media F2
1;7-1;8	1170	2110
1;8-1;9	1180	1960
Media anual	1173	2060

Como recogen las tablas, el promedio del F1 disminuyó de manera escalonada cayendo 126 Hz entre el primer y el segundo año de vida. Por el contrario, la media del F2 sufrió un descenso mayor, de 438 Hz, en el mismo periodo de tiempo. El ANOVA realizado mostró que ningún descenso significativo de los valores se producía en el primer formante entre el primer y el segundo año de vida ni entre el segundo y el tercero, pero sí reveló un descenso significativo del segundo formante en el segundo año de vida (la reducción se produjo en concreto entre los 1;7 y 1;8 años, primer mes del que pudieron extraerse muestras en el segundo año de vida). Por otro lado, el promedio no descendió de manera significativa en ningún formante durante el primer año de vida.

El número total de meses sobre los que pudo calcularse la desviación estándar para la vocal central baja en el niño fue de dos, uno en el primer año de vida y otro en el segundo. En el transcurso del primer año al segundo, ambos formantes mostraron el mismo patrón que fue el de un incremento de la desviación típica aunque en el F1 el aumento fue de 106 Hz y en el F2 de 226 Hz. Debido a esta escasez de meses en los que pudo calcularse la desviación estándar en los dos primeros años de vida, presentamos a continuación el gráfico que recoge la desviación estándar anual de la vocal [a] en cada uno de los dos primeros periodos (gráfico 7.44). Este gráfico vuelve a reflejar, aunque ya de manera anual, el incremento de la desviación estándar de

ambos formantes entre el primer y el segundo año de vida (aumento de 72 Hz en el F1 y de 170 Hz en el F2):

**Gráfico 7.44. Evolución de la desviación estándar de los dos primeros formantes de la vocal [a] del niño en los dos primeros años de vida (0;4-2;0 años).**



En las siguientes tablas encontramos los valores exactos de la desviación estándar de los dos meses en los que esta pudo calcularse así como la desviación típica anual:

**Tabla 7.32. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;8-0;9	134	14
Desviación estándar anual	98	21

**Tabla 7.33. Valores de la desviación estándar de los dos primeros formantes de la vocal [a] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;7-1;8	240	240
Desviación estándar anual	170	191

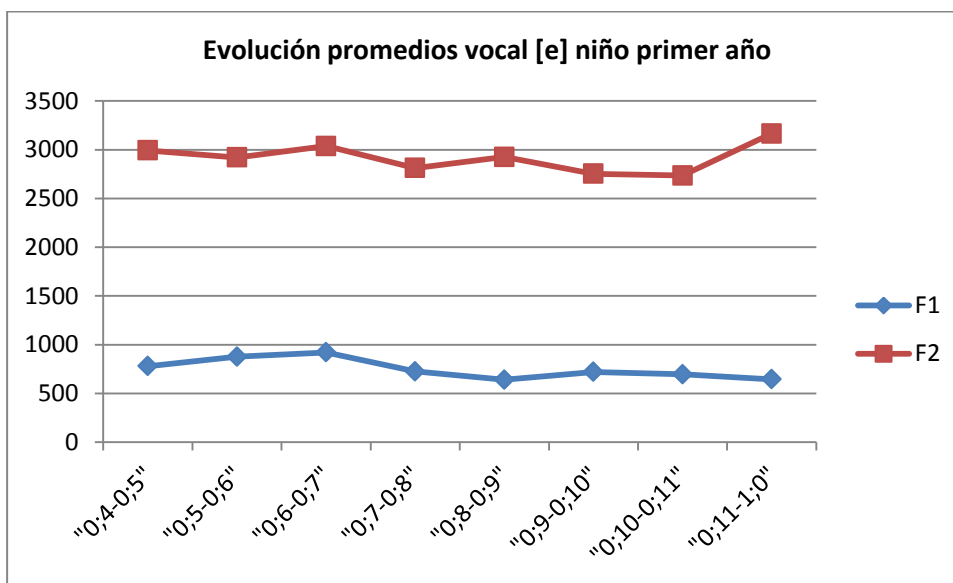
El incremento de la desviación típica se repitió lógicamente cuando calculamos la desviación anual de los valores. De esta forma, entre el primer y el segundo año de vida el grado

de apartamiento de la media de los valores del F1 aumentó 72 Hz, haciéndolo en 170 Hz el del F2.

### 7.1.2.2. *Vocal anterior semialta, [e].*

La evolución del promedio de los dos primeros formantes durante el primer año de vida pudo extraerse en todos los meses que conformaron el periodo, y no pareció manifestar cambios bruscos tal y como se refleja en el gráfico inferior (gráfico 7.45). Por un lado, el F1 exhibió los valores más altos en los tres primeros meses analizados de este primer año alcanzando el promedio máximo entre los 0;6 y los 0;7 años. Un descenso gradual se produjo hasta el final del periodo (exceptuando el ligero ascenso entre los 0;9 y los 0;10 años) registrándose el promedio más bajo entre los 0;8 y 0;9 años. Por otro lado, el F2 fue disminuyendo con el paso del tiempo y de manera leve hasta los 0;10-0;11 años, presentando a lo largo de esta disminución pequeños aumentos y descensos, y alcanzó en el último mes del periodo el promedio máximo (el promedio más bajo se registró en el penúltimo mes de este primer año).

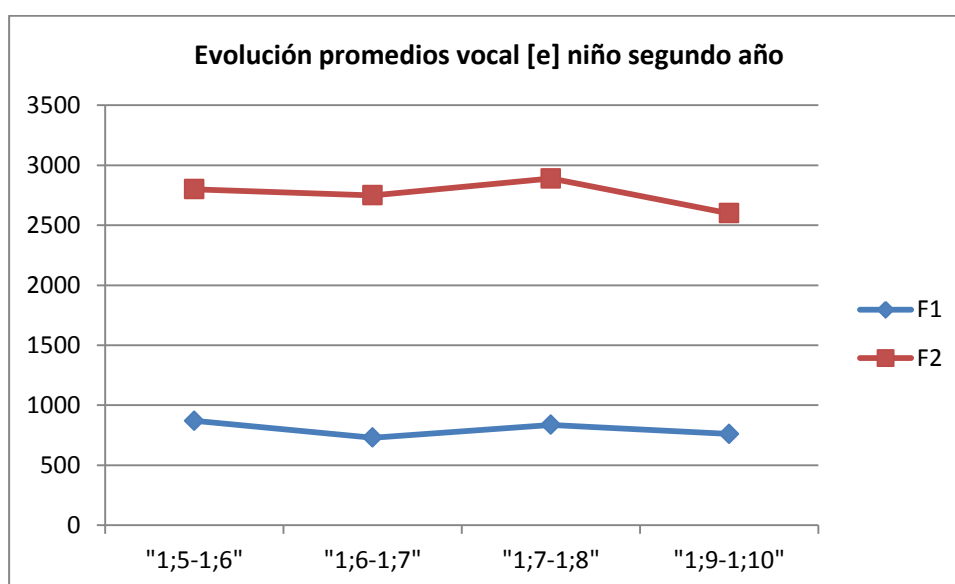
**Gráfico 7.45. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año).**



Al igual que en el periodo anterior, los ascensos y descensos de los promedios formánticos fueron muy escalonados en el segundo año de vida pero a diferencia de aquel, en este segundo año el F1 y el F2 presentaron un patrón evolutivo casi idéntico: disminución de los valores, incremento y una última reducción (gráfico 7.46). Si comenzamos por el F1, debemos señalar que este experimentó un aumento (de 225 Hz) entre el último mes del primer periodo y el primer mes del segundo alcanzando en este primer mes del segundo periodo el valor más alto

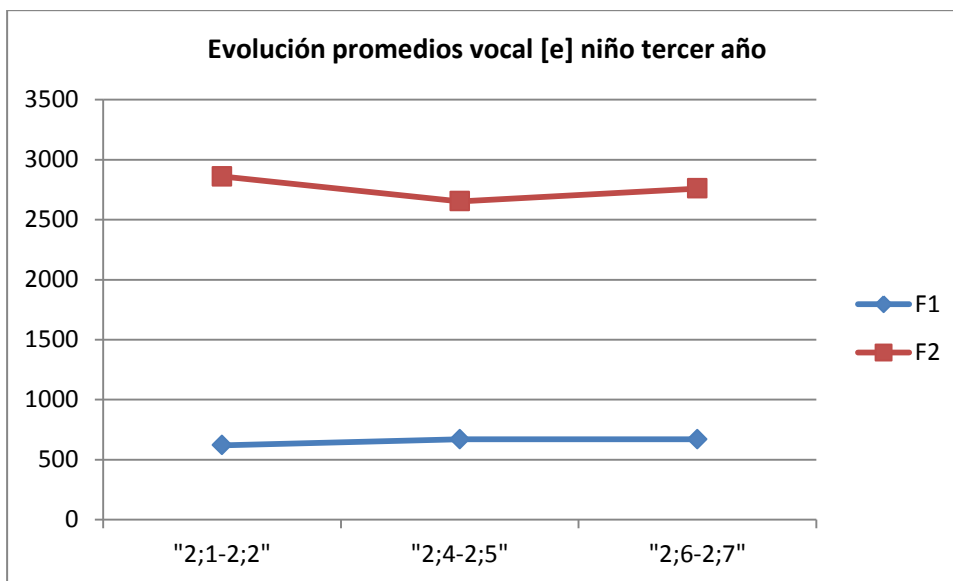
(870 Hz). A continuación, la media del primer formante descendió para registrar el promedio más bajo (de 730 Hz), aumentó en el mes siguiente y volvió a disminuir sin llegar en el último mes analizado a los valores mínimos. Contrariamente, el F2 sufrió un descenso entre el último mes del primer año de vida y el primer mes analizado del segundo (de 365 Hz) aunque presentó después a lo largo de este segundo periodo una evolución muy similar a la del F1. De esta forma, un leve descenso tuvo lugar en el mes sucesivo alcanzándose el promedio más elevado en el mes siguiente (2890 Hz) y registrándose la media más baja en el último mes analizado (2600 Hz).

**Gráfico 7.46. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años).**



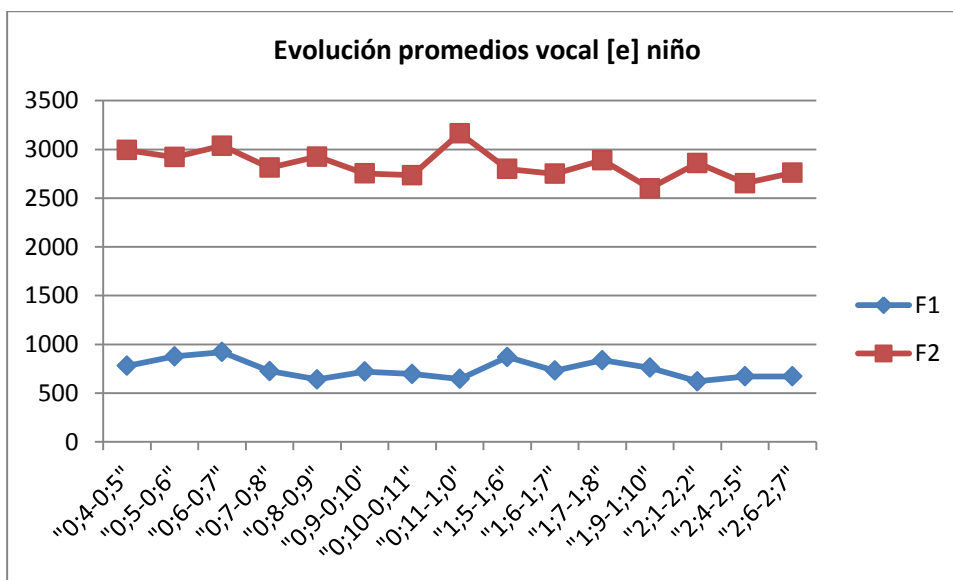
En el tercer año de vida sin embargo, las evoluciones de los promedios de ambos formantes fueron opuestas (gráfico 7.47). El F1 comenzó el periodo con un descenso de 140 Hz en comparación con el último mes analizado en el periodo anterior, y registró en este primer mes del tercer año de vida el valor mínimo. El promedio máximo se alcanzó en el mes siguiente y se mantuvo igual en el último mes del periodo. El F2 por otro lado, comenzó el periodo presentando el promedio máximo y habiendo experimentado un aumento de 260 Hz con respecto al último mes examinado del periodo anterior. En el mes siguiente el F2 disminuyó hasta registrar el valor mínimo y participó de un leve aumento en el último mes de este tercer año.

**Gráfico 7.47. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años).**



La evolución de los promedios del F1 y F2 queda recogida de manera conjunta en el gráfico inferior (gráfico 7.48). En él se refleja lo que ya hemos observado al comentar los gráficos anteriores respecto a los aumentos y descensos graduales y en nada bruscos que presentaron ambos formantes en todos los periodos estudiados:

**Gráfico 7.48. Evolución de los promedios de los dos primeros formantes de la vocal [e] del niño en los tres primeros años de vida (0;4-3;0 años).**



Las evoluciones comentadas en los gráficos anteriores responden a los valores de los promedios mensuales concretos recogidos en las siguientes tablas:

**Tabla 7.34. Valores de los promedios de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;4-0;5	780	2993
0;5-0;6	876	2921
0;6-0;7	920	3037
0;7-0;8	725	2813
0;8-0;9	640	2926
0;9-0;10	720	2753
0;10-0;11	695	2735
0;11-1;0	645	3165
Media anual	751	2906

**Tabla 7.35. Valores de los promedios de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;5-1;6	870	2800
1;6-1;7	730	2750
1;7-1;8	837	2890
1;9-1;10	760	2600
Media anual	787	2860

**Tabla 7.36. Valores de los promedios de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;1-2;2	620	2860
2;4-2;5	670	2653
2;6-2;7	670	2760
Media anual	660	2716

El promedio anual del primer formante aumentó entre el primer y el segundo año de vida en 36 Hz y descendió entre el segundo y el tercero en 127 Hz. Sin embargo, la diferencia

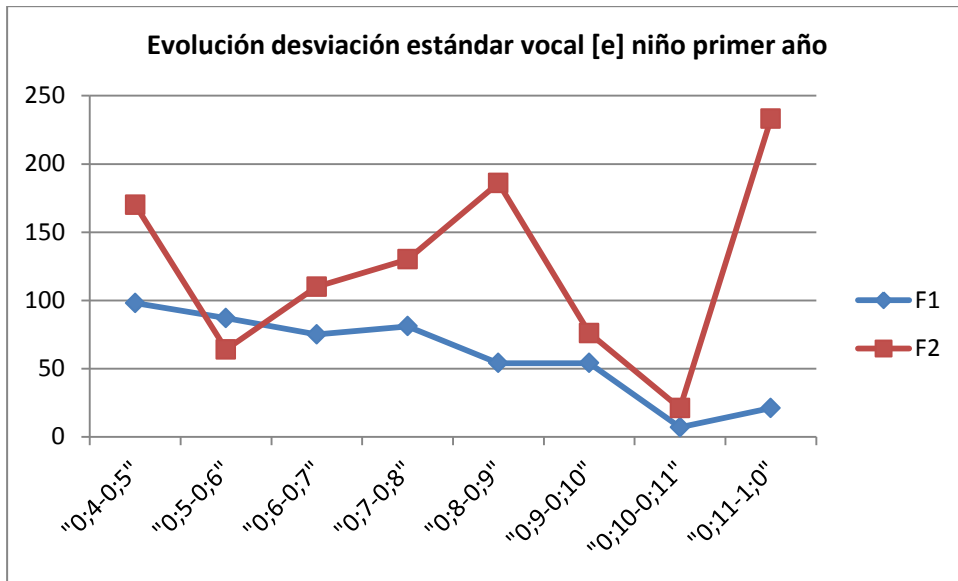
entre la media del primer y del tercer año fue solo de 91 Hz. El promedio del F2 por otro lado, disminuyó a lo largo de los tres periodos estudiados: 46 Hz entre el primer y el segundo año, 144 Hz entre el segundo y el tercero, y 190 Hz entre el primero y el tercero. Los análisis de varianza realizados sobre los valores del F2 demostraron que a pesar del descenso progresivo que dicho formante experimentó a lo largo de los tres periodos analizados, este no fue significativo cuando se compararon los valores del primer año con los del segundo y los del segundo con los del tercero pero sí lo fue tras equiparar los del primero con los del tercero. Sin embargo, este descenso fue mínimo (diferencia mínima entre el estadístico y el valor crítico) por lo que quizás pudo deberse a la diferencia en el número de señales entre el primer año de vida y el tercero (veintisiete y cinco respectivamente).

Por otro lado en el F1, el ANOVA mostró que ningún descenso significativo de los valores tenía lugar entre el primer y el segundo año de vida, entre el segundo y el tercero y entre el primero y el tercero pero sí que se producía un descenso al comparar los valores de los cuatro primeros meses del primer año de vida con los valores de los cuatro últimos de este mismo periodo. No obstante, el hecho de que en el segundo año de vida los valores del F1 no solo no fueran inferiores a los del primero sino que fueran de hecho más altos (media de 787 Hz en el segundo año de vida frente a los 674 Hz de los cuatro últimos meses del primero) demuestra que este descenso que se produce entre los dos periodos del primer año de vida no puede deberse a cambios anatómicos ya que si esta fuera la razón los valores del F1 no se recuperarían, llegando incluso a aumentar en el segundo año de vida, sino que serían iguales o inferiores. Puede este descenso deberse por tanto a la coincidencia de una agrupación de señales con un F1 más alto en los cuatro primeros meses del primer año de vida en comparación con las señales de los cuatro últimos meses de este primer periodo.

En cuanto a la desviación estándar, y al igual que ocurría con los promedios, esta pudo calcularse en los ocho meses que conformaron el primer año de vida. El siguiente gráfico (gráfico 7.49) presenta la evolución de la misma y en él puede observarse que aunque la tendencia de ambas desviaciones fue hacia el descenso registrándose el valor más bajo tanto para el F1 como para el F2 entre los 0;10 y 0;11 años (7 Hz para el F1 y 21 Hz para el F2), los puntos de aumento y descenso así como las pendientes de estos no siempre coincidieron, como tampoco hubo coincidencia en el mes en el que se alcanzó la desviación típica máxima (el primer mes en el caso del F1 y el último en el del F2).

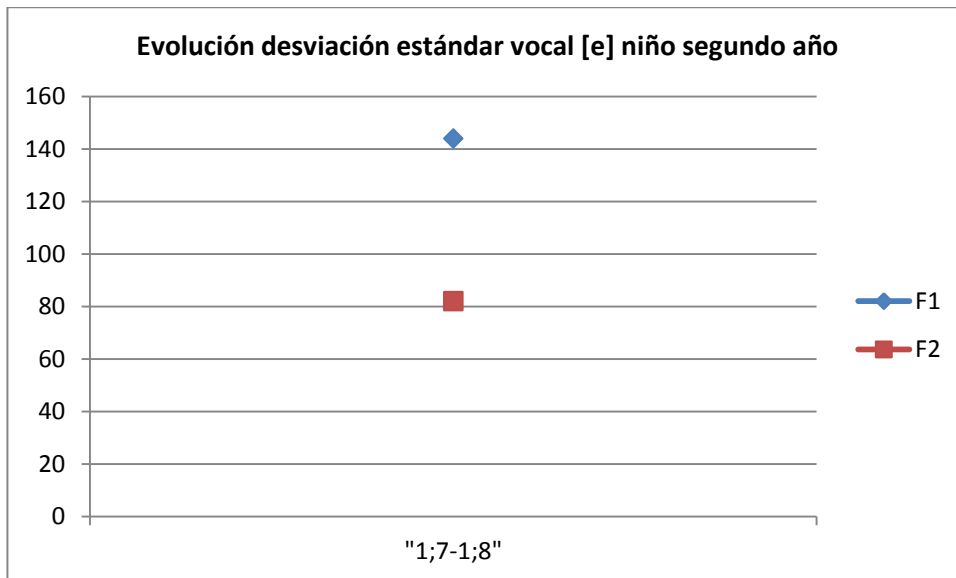


**Gráfico 7.49. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año).**



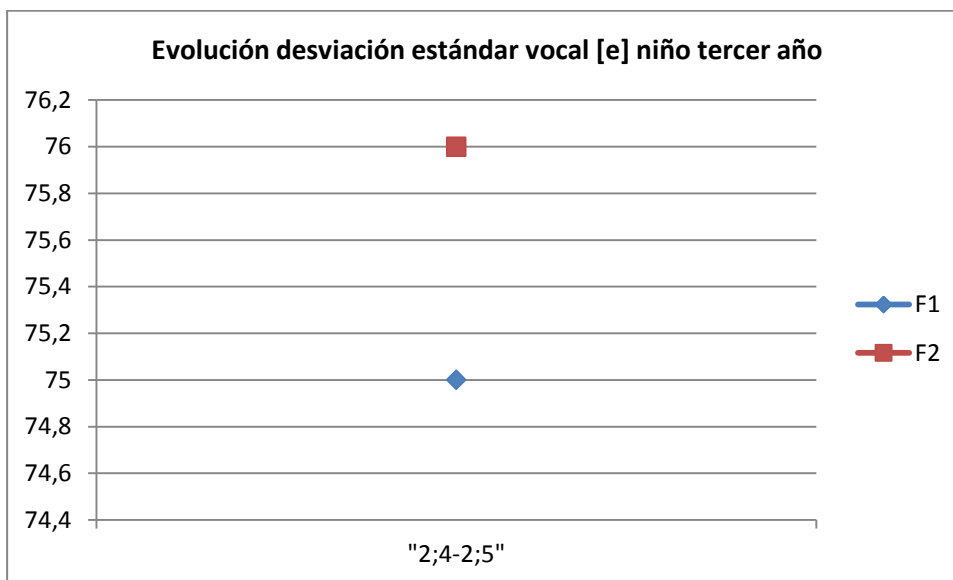
A diferencia de lo que ocurría en el primer año de vida, tan solo hallamos un mes del segundo periodo en el que calcular la desviación estándar (gráfico 7.50). Respecto al último mes del primer año de vida, la desviación típica del F1 aumentó 123 Hz mientras que la del F2 descendió 151 Hz.

**Gráfico 7.50. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años).**



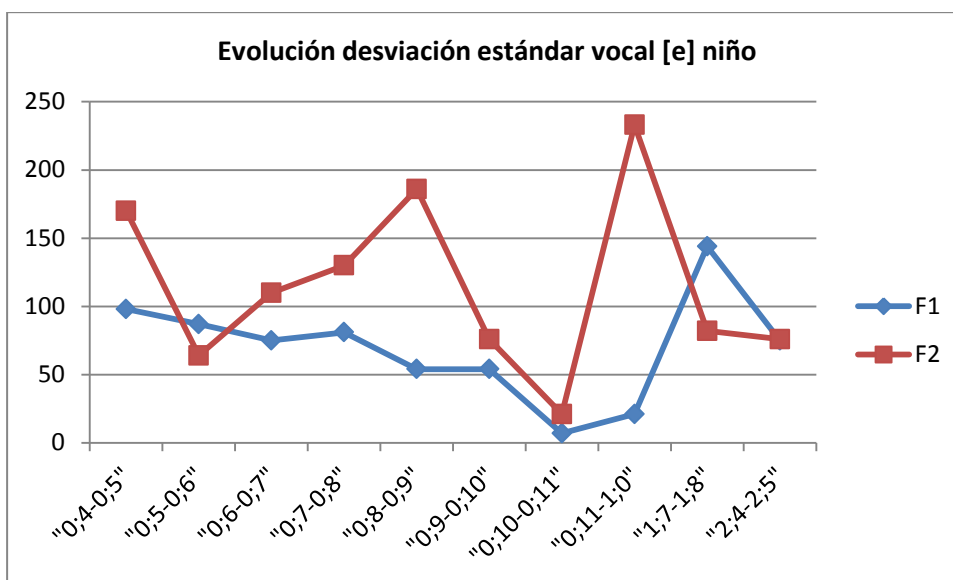
En el tercer año de vida y al igual que ocurría en el segundo, tan solo pudimos calcular la desviación estándar de un mes (gráfico 7.51). En él y con respecto al último mes analizado del periodo anterior, la desviación típica del F1 descendió 69 Hz y la del F2 lo hizo en 6 Hz.

**Gráfico 7.51. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años).**



El gráfico siguiente (gráfico 7.52) expone la evolución de la desviación estándar de ambos formantes en los tres periodos analizados. Como podemos observar, las desviaciones típicas de ambos formantes tienden a descender a lo largo del primer año registrándose los valores mínimos entre los 0;10 y 0;11 años. A partir de este mes la desviación estándar del F2 aumenta bruscamente para alcanzar su máximo en el mes siguiente (233 Hz) y después experimenta dos descensos, uno notorio y otro más suave. Respecto al F1, este experimenta un ascenso hasta los 1;7 y 1;8 años, mes en el que alcanza la desviación máxima (144 Hz), y luego desciende en el transcurso al tercer año de vida. Los incrementos y las disminuciones de la desviación del segundo formante fueron más pronunciados que los del primero a lo largo de los periodos estudiados.

**Gráfico 7.52. Evolución de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en los tres primeros años de vida (0;4-3;0 años).**



Los valores exactos de la desviación estándar en los meses representados en los gráficos anteriores así como la desviación típica anual se presentan en las tablas siguientes:

**Tabla 7.37. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;4-0;5	98	170
0;5-0;6	87	64
0;6-0;7	75	110
0;7-0;8	81	130
0;8-0;9	54	186
0;9-0;10	54	76
0;10-0;11	7	21
0;11-1;0	21	233
Desviación estándar anual	115	170

**Tabla 7.38. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;7-1;8	144	82
Desviación estándar anual	117	189

**Tabla 7.39. Valores de la desviación estándar de los dos primeros formantes de la vocal [e] del niño en el tercer año de vida (2;0-3;0 años).**

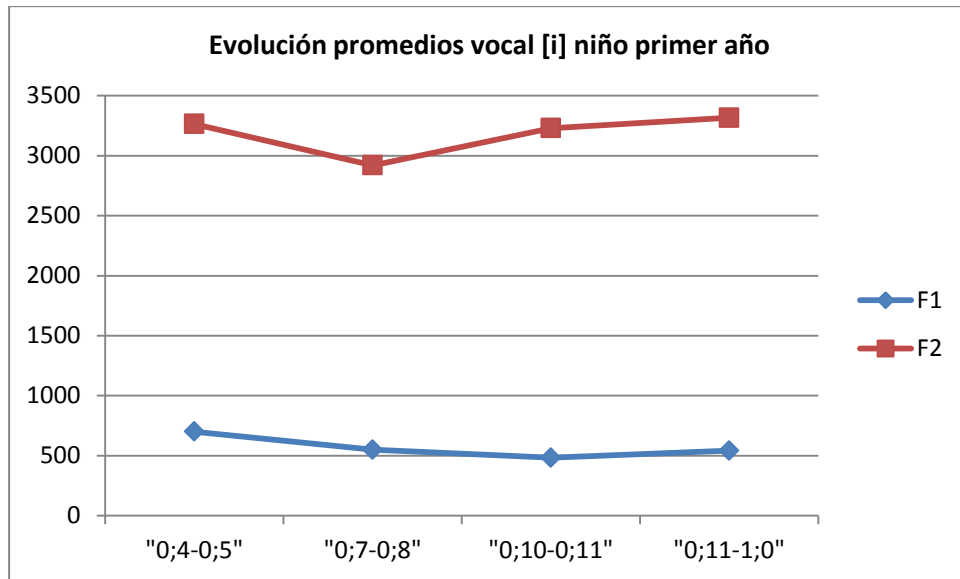
Edad	Desviación estándar F1	Desviación estándar F2
2;4-2;5	75	76
Desviación estándar anual	58	107

El grado de apartamiento de los valores del F1 y del F2 con respecto a la media aumentó ligeramente entre el primer y el segundo año de vida (2 Hz en el F1 y 19 Hz en el F2) y descendió entre el segundo y el tercero (59 Hz en el F1 y 82 Hz en el F2) así como entre el primero y el tercero (57 Hz en el F1 y 63 Hz en el F2). La desviación típica del F1 y del F2 de la vocal anterior semialta se comporta como la desviación estándar de los dos primeros formantes de la [a], es decir, aumenta entre el primer y el segundo año de vida.

### **7.1.2.3. Vocal anterior alta, [i].**

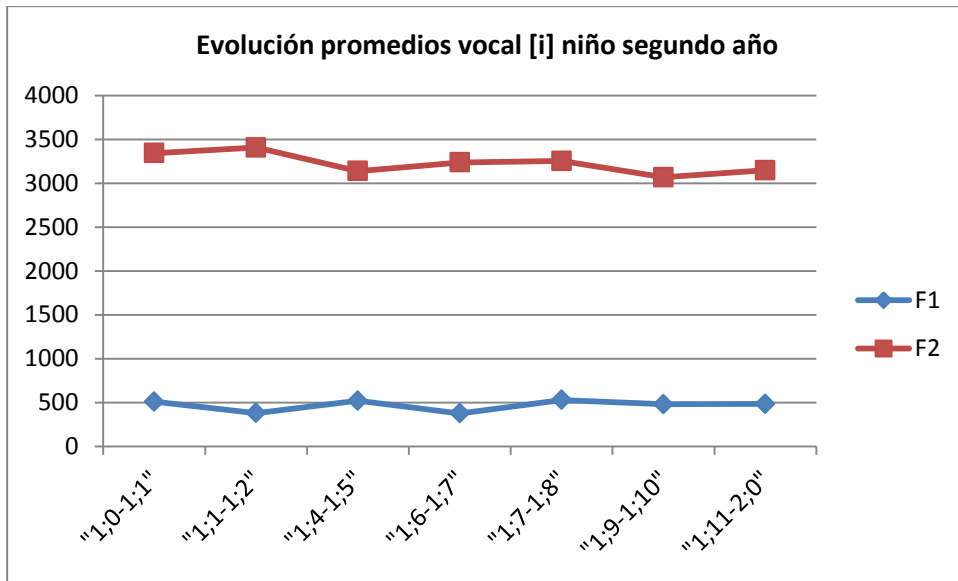
Después de calcular el promedio de los valores en los cuatro meses del primer periodo temporal en los que la vocal [i] fue reconocida por cuatro de los transcritores y tras representar esa media en un gráfico lineal (gráfico 7.53), observamos que la media de ambos formantes descende entre el primer y el segundo mes analizado siendo más notorio el descenso del F2. A partir de los 0;8 años el promedio del F1 continúa descendiendo y aumenta ligeramente en el último mes mientras que el del F2 se incrementa hasta alcanzar el máximo en el último mes analizado (a diferencia del F1 que presentó su promedio máximo en el primer mes del periodo).

**Gráfico 7.53. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año).**



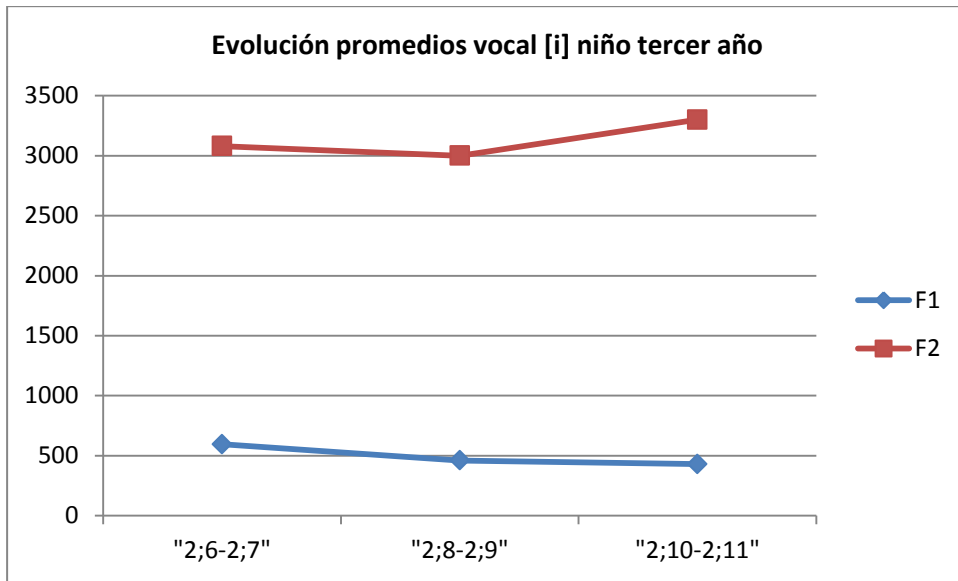
El promedio del F1 en el segundo periodo (gráfico 7.54) presentó una progresión bastante lineal puesto que los valores se situaron entre los trescientos y pico y los quinientos hertzios durante todo el periodo con cambios poco notorios. Entre los sutiles descensos y aumentos hallamos dos meses, el de los 1;7 años, mes en el que se registró el valor mínimo (377 Hz), y el de los 1;8 años, mes en el que se alcanzó el máximo (530 Hz). En cuanto al F2, los promedios mensuales volvieron a situarse en torno a los 3000 y 3500 Hz, alcanzándose el máximo en el segundo mes analizado (3410 Hz) y el mínimo en el penúltimo (3070 Hz). La evolución de la media de ambos formantes no manifestó descensos ni aumentos bruscos.

**Gráfico 7.54. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años).**



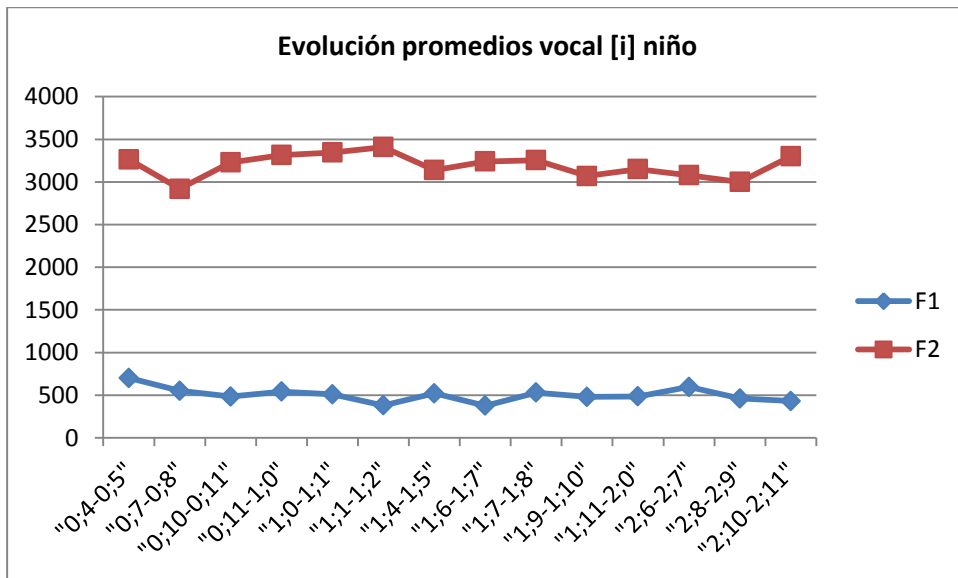
Como en los años anteriores, la media de los dos primeros formantes se movió alrededor de los mismos valores en el tercer periodo estudiado (gráfico 7.55). Respecto al curso evolutivo del promedio de ambos formantes debemos decir que el F1, que en el primer mes analizado de este tercer periodo aumentó 110 Hz respecto al último mes del periodo anterior, descendió progresivamente a lo largo de los tres meses analizados en este tercer año de vida logrando su máximo (de 595 Hz) en el primer mes analizado y su mínimo en el último (de 430 Hz). El F2 por su parte (y cuya diferencia con el último mes analizado en el periodo anterior fue de 70 Hz), también descendió entre el primer y el segundo mes examinado pero ascendió en el transcurso hacia el último mes analizado alcanzando su máximo (3300 Hz).

**Gráfico 7.55. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años).**



El gráfico siguiente (gráfico 7.56) viene a resumir lo observado anteriormente en cada uno de los periodos temporales representados por separado: evolución prácticamente lineal del promedio del F1a lo largo de todo el periodo y ligero descenso de los valores del F2 a lo largo del segundo con un repunte de los mismos a finales del tercero.

**Gráfico 7.56. Evolución de los promedios de los dos primeros formantes de la vocal [i] del niño en los tres primeros años de vida (0;4-3;0 años).**



Una vez más exponemos las tablas que recogen los valores concretos de los promedios hallados en los distintos meses de los tres periodos estudiados:

**Tabla 7.40. Valores de los promedios de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;4-0;5	700	3265
0;7-0;8	550	2920
0;10-0;11	483	3230
0;11-1;0	541	3316
Media anual	541	3246

**Tabla 7.41. Valores de los promedios de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;0-1;1	510	3345
1;1.1;2	380	3410
1;4-1;5	520	3140
1;6-1;7	377	3240
1;7-1;8	530	3255
1;9-1;10	480	3070
1;11-2;0	485	3150
Media anual	465	3224

**Tabla 7.42. Valores de los promedios de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;6-2;7	595	3080
2;8-2;9	460	3000
2;10-2;11	430	3300
Media anual	520	3115

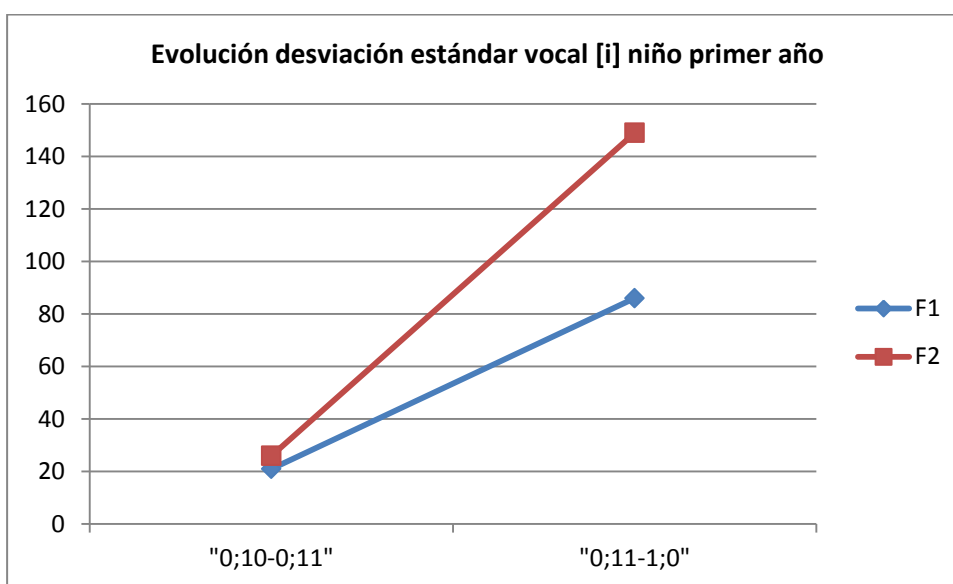
Tal y como muestran las tablas el promedio del F1 descendió 76 Hz entre el primer y el segundo año de vida y aumentó 55 Hz entre el segundo y el tercero (produciéndose un descenso de 21 Hz entre el primer y el tercer año). La tendencia del promedio del F2 fue hacia la



reducción en todos los periodos cayendo 22 Hz entre el primer y el segundo periodo y 109 Hz entre el segundo y el tercero (habiendo una diferencia de 131 Hz entre el primero y el tercero). Como era de suponer a partir de estos datos, el ANOVA realizado confirmó que no había diferencias significativas en los valores de los dos primeros formantes entre los periodos estudiados (ni a lo largo del primer año, ni entre el primer y el segundo año de vida, ni entre el segundo y el tercero ni entre el primero y el tercero). La evolución del promedio de esta vocal se asemeja más a la de la otra vocal anterior analizada, la anterior semialta [e], que a la de la vocal central baja [a] cuyo F2 experimentó un descenso significativo entre el primer y el segundo año de vida.

El segundo análisis realizado, el que midió el grado de apartamiento de los valores con respecto a la media, solo pudo calcularse en dos de los cuatro meses analizados para la [i] en el primer año de vida (gráfico 7.57). Tal y como muestra el gráfico inferior la desviación estándar aumentó para el F1 y el F2 entre el penúltimo y el último mes del periodo.

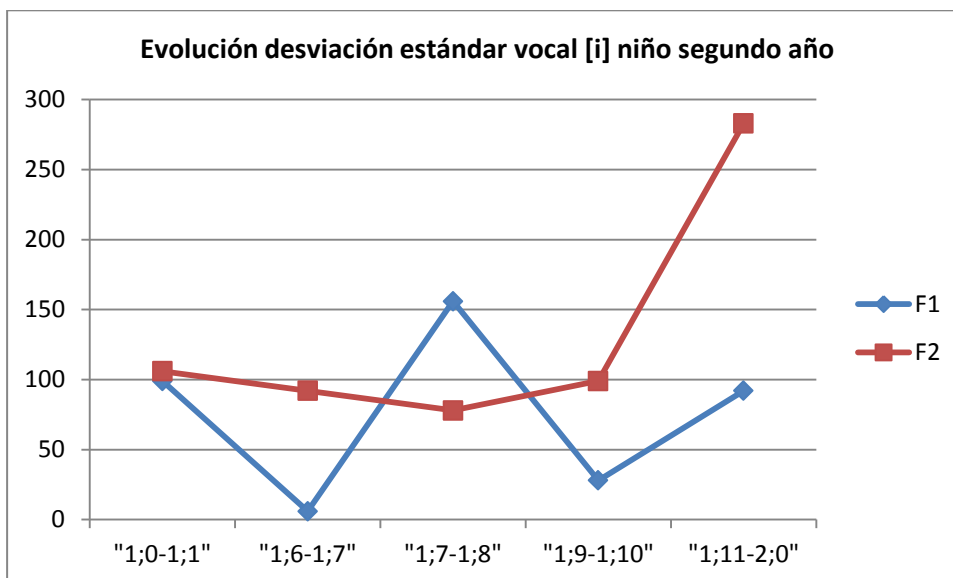
**Gráfico 7.57. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año).**



La evolución de la desviación típica del primer formante en el segundo periodo estudiado se asemejó a la silueta de una montaña rusa (gráfico 7.58). Dicho formante partió de una desviación estándar media (habiendo aumentado 13 Hz con respecto al último mes del periodo anterior), registró el valor más bajo en el mes siguiente (de 6 Hz), aumentó en el mes sucesivo alcanzando el valor máximo (156 Hz), descendió en el transcurso al mes próximo y volvió a incrementarse en el último mes del periodo. Respecto al F2, y aunque este también comenzó con una desviación estándar media, esta descendió progresivamente hasta alcanzar el

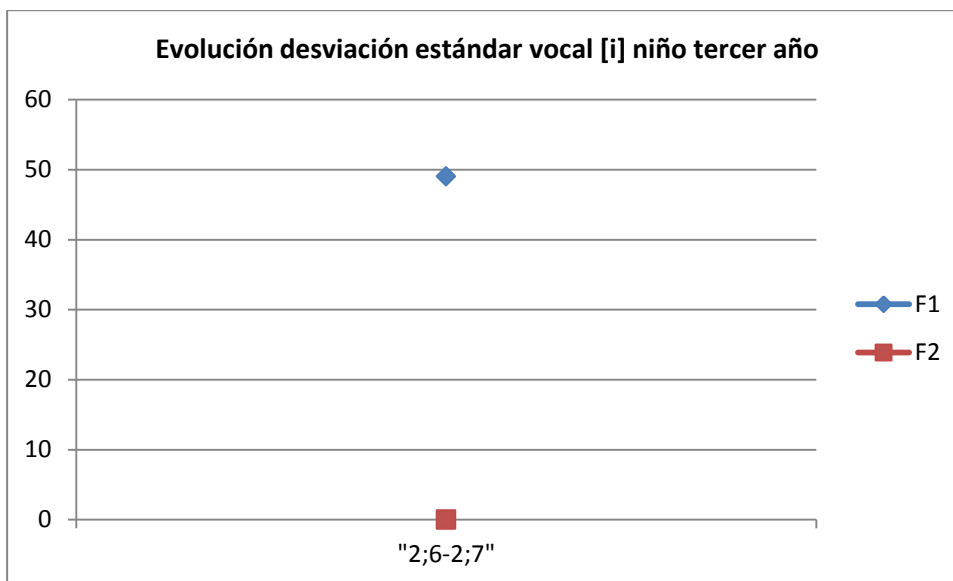
mínimo a los 1;8 años (78 Hz), momento a partir del cual comenzó a ascender produciéndose el incremento más abrupto entre los 1;10 y los 2;0 años, mes en el que alcanzó el máximo (283 Hz). La desviación estándar de este formante experimentó un descenso de 43 Hz entre el último mes analizado del primer periodo y el primero del segundo.

**Gráfico 7.58. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años).**



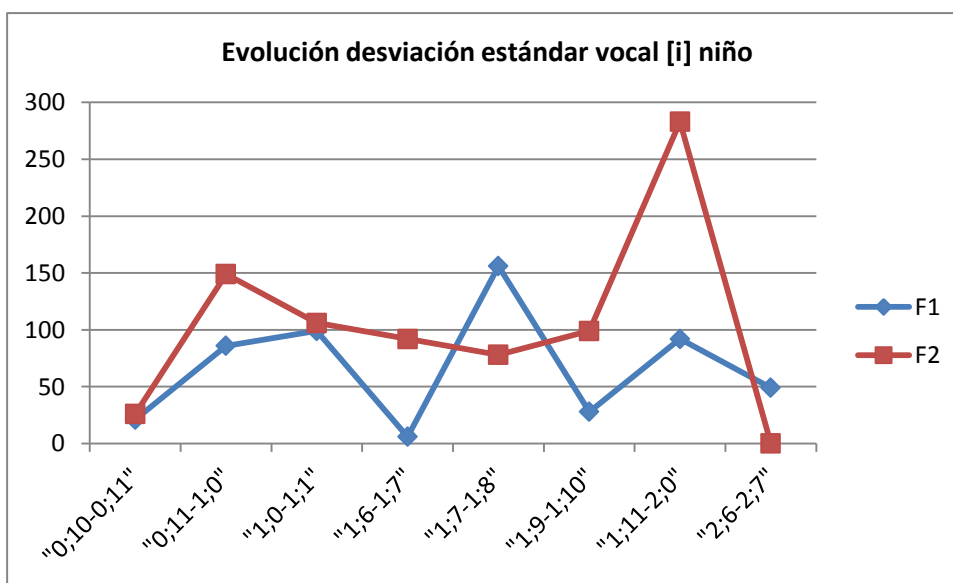
Tan solo fue posible calcular la desviación estándar en uno de los tres meses del tercer año de vida en los que la vocal [i] fue identificada (gráfico 7.59). Las desviaciones de ambos formantes descendieron cuando se compararon con las del último mes del segundo año de vida si bien la del F2 lo hizo en mayor medida puesto que disminuyó 283 Hz frente a los 43 Hz del F1.

**Gráfico 7.59. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años).**



El gráfico siguiente refleja la evolución completa que presentó la desviación de los dos primeros formantes a lo largo de todo el periodo estudiado (gráfico 7.60). Puestos en común los cursos evolutivos de las desviaciones de ambos formantes hallamos algunas diferencias tales como el hecho de que las variaciones más notables en la desviación estándar se produzcan en el F1 en los dos primeros años analizados mientras que en el F2 estas tienen lugar entre el segundo y el tercer periodo presentando los dos primeros periodos aumentos y descensos más escalonados.

**Gráfico 7.60. Evolución de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en los tres primeros años de vida (0;4-3;0 años).**



Presentamos a continuación las tablas que recogen los valores exactos obtenidos tras el cálculo de la desviación estándar en cada uno de los meses y de los años de nuestro estudio longitudinal:

**Tabla 7.43. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;10-0;11	21	26
0;11-1;0	86	149
Desviación estándar anual	85	157

**Tabla 7.44. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;0-1;1	99	106
1;6-1;7	6	92
1;7-1;8	156	78
1;9-1;10	28	99
1;11-2;0	92	283
Desviación estándar anual	87	146

**Tabla 7.45. Valores de la desviación estándar de los dos primeros formantes de la vocal [i] del niño en el tercer año de vida (2;0-3;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
2;6-2;7	49	0
Desviación estándar anual	92	129

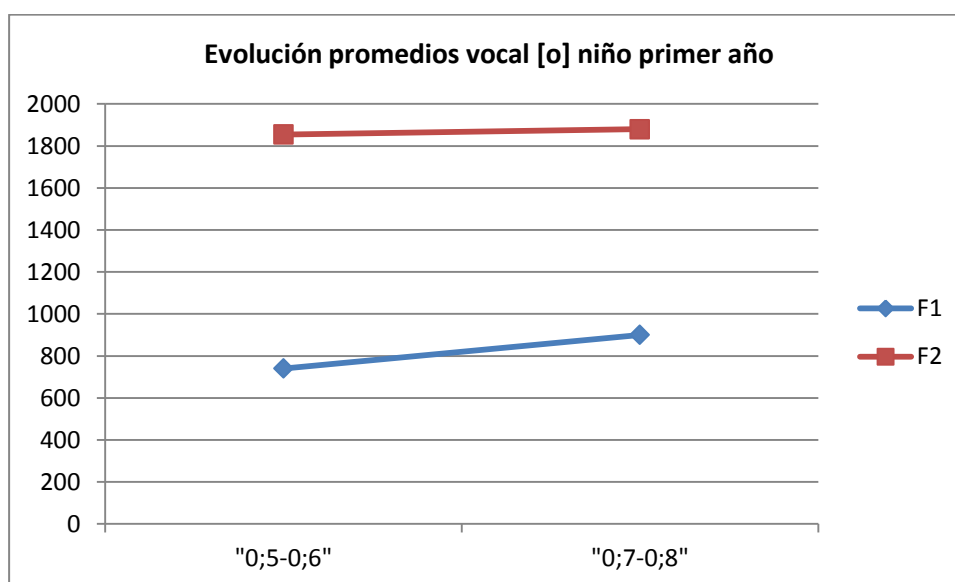
La desviación estándar del primer formante aumentó 2 Hz entre el primer y el segundo año de vida y 5 Hz entre el segundo y el tercero (incremento de 7 Hz entre el primer y el tercer año). Por el contrario, el F2 disminuyó con el paso del tiempo y se redujo 11 Hz entre el primer y el segundo año de vida y 17 Hz entre el segundo y el tercero (el descenso entre el primer y el tercer año fue de 28 Hz). Tras equiparar estos resultados con los obtenidos en otras vocales observamos que al igual que ocurría en la vocal central baja [a] y en la anterior semialta [e], al

menos durante los dos primeros años de vida, la desviación estándar del F1 tendió a aumentar ligeramente con el paso del tiempo mientras que la tendencia de la desviación típica del F2 de la vocal anterior alta [i] solo coincide con la de la otra vocal anterior, la semialta [e], en cuanto a que en ambas tiende a descender mientras que en la central baja [a] aumenta.

#### 7.1.2.4. Vocal posterior semialta, [o].

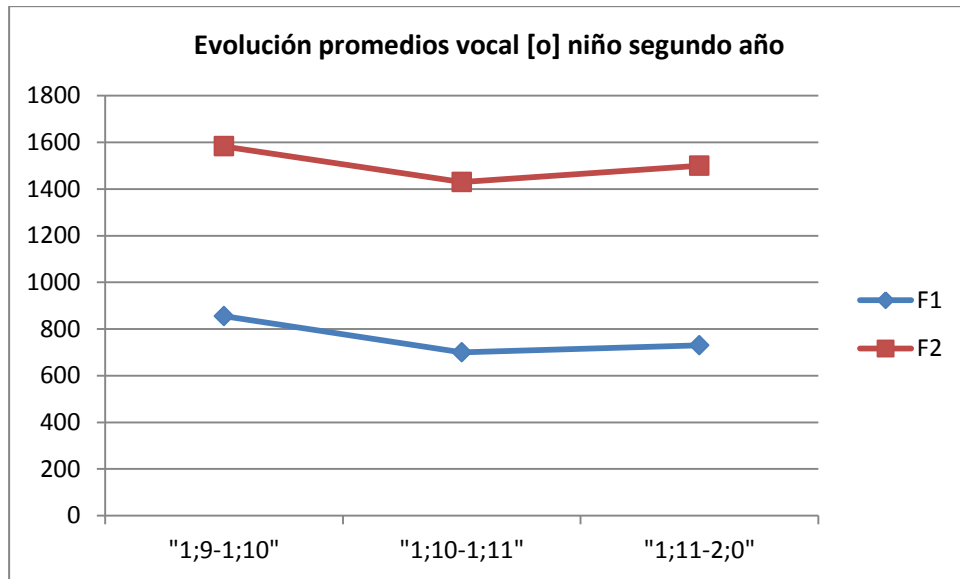
El promedio de los valores de los dos primeros formantes durante el primer año de vida solo pudo calcularse entre los 0;5 y los 0;6 años y entre los 0;7 y los 0;8 años debido a que tan solo en estos dos meses se identificó la [o] por cuatro de los cinco oyentes. Tal y como refleja el gráfico 7.61, la tendencia del promedio de ambos formantes fue hacia el aumento incrementándose 160 Hz el F1 y 25 Hz el F2.

**Gráfico 7.61. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en el primer año de vida (0;4-1;0 año).**



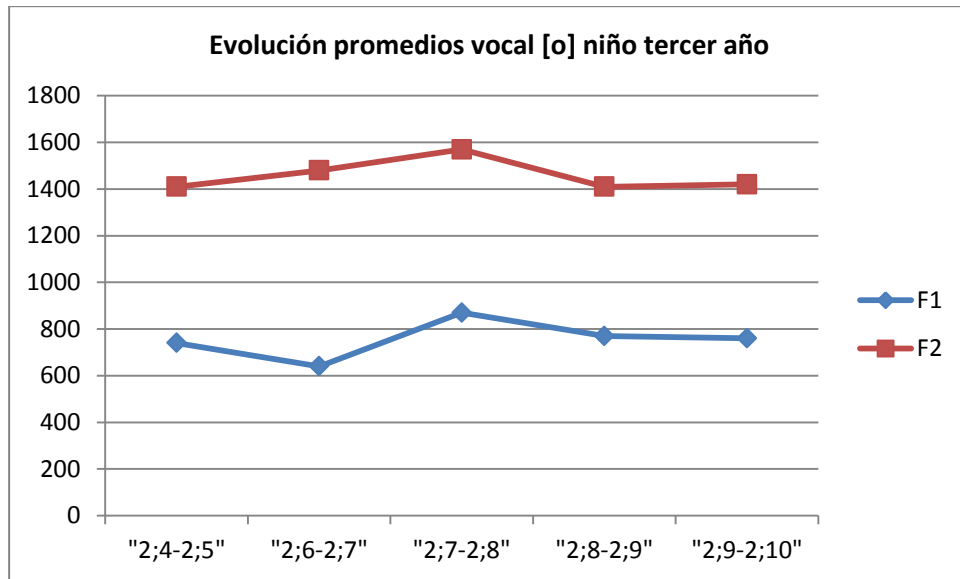
En el segundo periodo estudiado el promedio de los dos primeros formantes presentó una evolución muy similar tal y como muestra el gráfico 7.62. Ambos formantes comenzaron el periodo con el promedio más alto (de 855 Hz en el F1 y de 1583 Hz en el F2), descendieron sutilmente en el mes intermedio para registrar el valor mínimo (de 700 Hz en el F1 y de 1430 Hz en el F2) y volvieron a ascender ligeramente en el transcurso hacia el último mes del periodo. Llama la atención el descenso notorio del promedio del segundo formante que cae 297 Hz entre el último mes analizado del primer año de vida y el primero del segundo mientras que el F1 solo desciende 45 Hz.

**Gráfico 7.62. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en el segundo año de vida (1;0-2;0 años).**



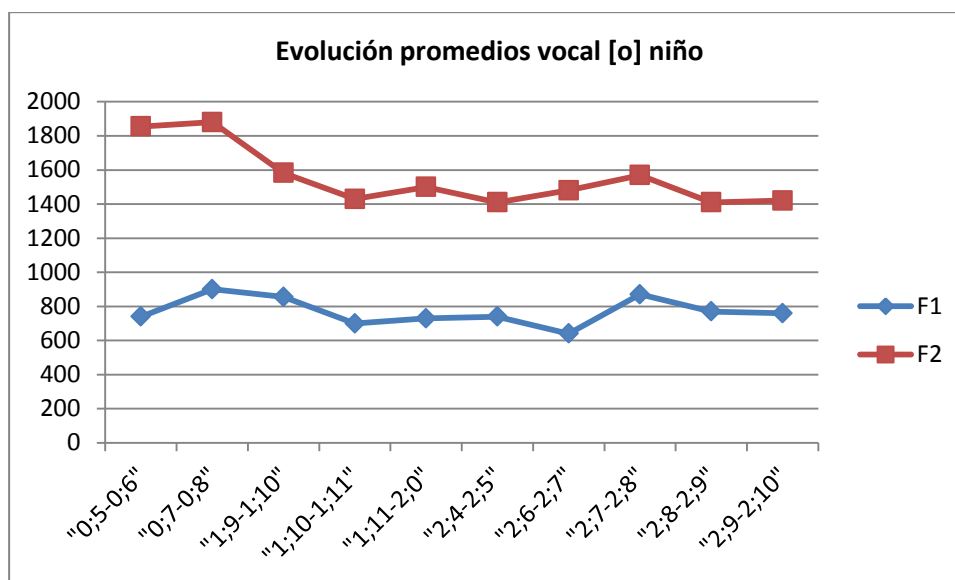
La diferencia hallada en la media del segundo formante entre el último mes analizado del primer año de vida y el primero del segundo fue mucho mayor que la encontrada en el mismo formante entre el último mes del segundo año y el primero del tercero donde dicha media tan solo se redujo 90 Hz tal y como manifiesta el gráfico inferior (gráfico 7.63). En cuanto a la evolución de este, dicho formante manifestó ligeros aumentos y descensos situándose los valores a lo largo de todo el periodo entre los 1400 Hz y los 1500 Hz. Los valores mínimos y máximos se registraron entre los 2;4 y 2;5 años (1410 Hz) y los 2;7 y 2;8 años (1570 Hz), respectivamente. En relación al F1, este comenzó el periodo incrementando en 10 Hz la media mensual con respecto a la del último mes del periodo anterior. A continuación descendió registrándose el valor mínimo a los 2;7 años (640 Hz), mes a partir del cual el promedio fue en aumento alcanzando el promedio máximo a los 2;8 años (870 Hz), y volvió a disminuir ligeramente en los meses siguientes.

**Gráfico 7.63. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en el tercer año de vida (2;0-3;0 años).**



En el gráfico inferior (gráfico 7.64), que refleja la evolución del promedio de los dos primeros formantes a lo largo de los tres años de vida, puede observarse, quizás con mayor claridad que en los gráficos en los que se representaban los periodos aislados, el notorio descenso de valores que tiene lugar en el F2 entre el primer y el segundo periodo. Entre el segundo y el tercero tan solo encontramos pequeñas variaciones y la media del F2 se situó entre los 1400 Hz y los 1500 Hz en lo que restó de etapa. Por otro lado, aunque la media del F1 mostró su descenso más notable también entre el primer y el segundo año de vida, el cambio fue mucho menos drástico que el ocurrido en el segundo formante. Finalmente, entre el segundo y tercer periodo las medias del F1 se asemejaron en gran medida situándose en todo momento entre los 600 y 800 Hz.

**Gráfico 7.64. Evolución de los promedios de los dos primeros formantes de la vocal [o] del niño en los tres primeros años de vida (0;4-3;0 años).**



Un seguimiento más detallado de los promedios obtenidos en cada uno de los meses y la media final de los tres periodos estudiados puede efectuarse a partir de las siguientes tablas:

**Tabla 7.46. Valores de los promedios de los dos primeros formantes de la vocal [o] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;5-0;6	740	1855
0;7-0;8	900	1880
Media anual	793	1863

**Tabla 7.47. Valores de los promedios de los dos primeros formantes de la vocal [o] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;9-1;10	855	1583
1;10-1;11	700	1430
1;11-2;0	730	1500
Media anual	799	1536



**Tabla 7.48. Valores de los promedios de los dos primeros formantes de la vocal [o] del niño en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;4-2;5	740	1410
2;6-2;7	640	1480
2;7-2;8	870	1570
2;8-2;9	770	1410
2;9-2;10	760	1420
Media anual	774	1467

Como podemos observar, ambos formantes tendieron a descender con el paso del tiempo si bien el primero experimentó un leve aumento entre el primer y el segundo año de vida. De esta forma, el F1 aumentó 6 Hz entre el primer y el segundo año de vida y disminuyó 25 Hz entre el segundo y el tercero (habiendo una diferencia entre el primer y el tercer año de 19 Hz). El F2 por otra parte, se redujo 327 Hz entre el primer y el segundo periodo y 69 Hz entre el segundo y el tercero produciéndose una reducción de 396 Hz entre la primera y la tercera etapa.

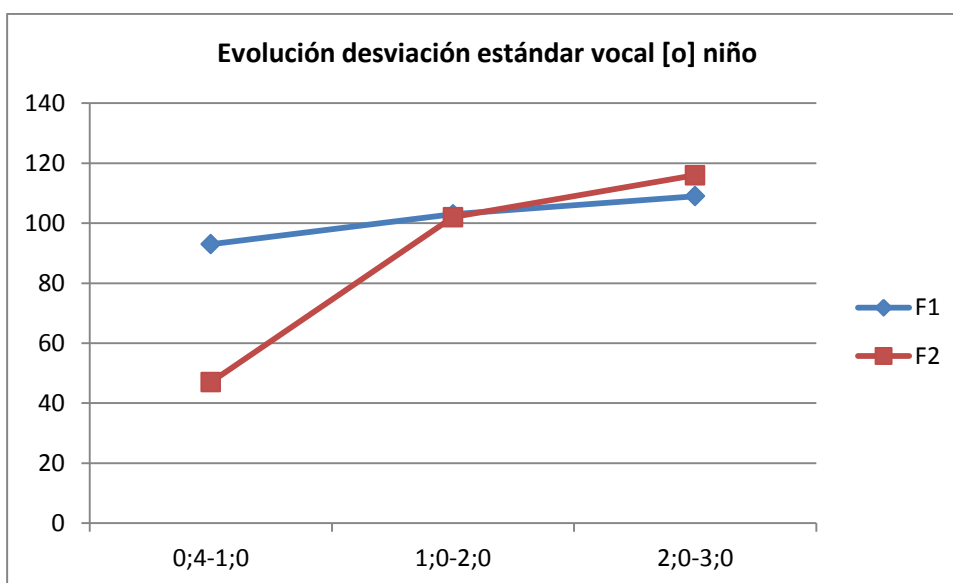
Una vez realizados los análisis estadísticos se comprobó que los valores del F1 hallados en el primer año de vida eran muy similares a los encontrados en el segundo y en el tercero (la semejanza se repitió al comparar los valores del segundo con los del tercer año). En el F2 sin embargo, el ANOVA reveló que un descenso significativo de los valores tenía lugar entre el primer y el segundo año de vida y que este se producía en concreto entre los 1;9 y los 1;10 años. Ningún descenso significativo fue hallado al comparar los valores del segundo año con los del tercero y sí lo hubo al equiparar los del primero con los del tercero.

Si comparamos estos resultados con los obtenidos para otras vocales, encontramos que hay una coincidencia con la vocal central baja [a] en cuanto al descenso significativo del F2 entre el primer y el segundo año de vida, descenso que no ocurre en las vocales anteriores alta y semialta. En cuanto al F1, este se mantiene sin cambios significativos como en el resto de vocales analizadas.

El siguiente gráfico refleja la evolución de la desviación estándar anual de la vocal [o] en los tres primeros años de vida y no en cada uno de los años ya que esta solo pudo calcularse en un mes del primer año, en uno del segundo año y en dos del tercero. Por eso y para evitar incluir gráficos con un solo valor (sin que estos puedan compararse con gráficos que reflejen

más valores como ocurría en el caso de la vocal [e]) presentamos a continuación un solo gráfico con la evolución anual de la desviación típica en los tres periodos estudiados (gráfico 7.65). Como podemos observar, la desviación estándar de ambos formantes aumentó con el paso del tiempo si bien en ninguno de ellos el aumento fue muy notorio (incremento de 16 Hz en el F1 y de 69 Hz en el F2).

**Gráfico 7.65. Evolución de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en los tres primeros años de vida (0;4-3;0 años).**



Los datos de la desviación estándar mensual y anual para esta vocal se recogen en las siguientes tablas:

**Tabla 7.49. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Desviación estándar F1	Desviación estándar F2
0;5-0;6	14	64
Desviación estándar anual	93	47

**Tabla 7.50. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
1;9-1;10	96	105
Desviación estándar anual	103	102

**Tabla 7.51. Valores de la desviación estándar de los dos primeros formantes de la vocal [o] del niño en el tercer año de vida (2;0-3;0 años).**

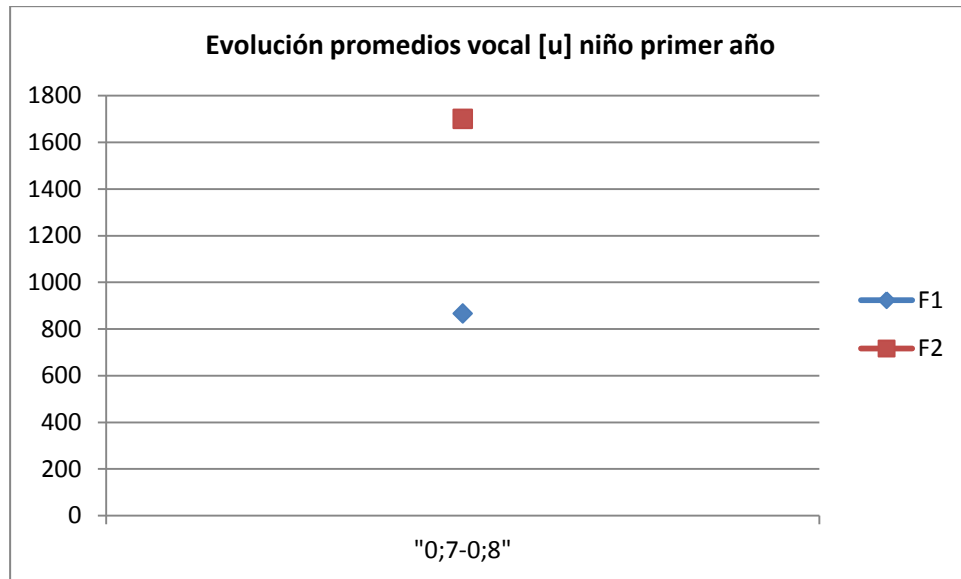
Edad	Desviación estándar F1	Desviación estándar F2
2;7-2;8	99	42
2;8-2;9	156	212
Desviación estándar anual	109	116

Si nos fijamos en los valores anuales y tal y como hemos comentado en el gráfico 7.65, observamos que la desviación estándar de ambos formantes aumenta con el paso del tiempo aunque de manera leve. De esta forma el F1 se incrementa 10 Hz entre el primer y el segundo año de vida y 6 Hz entre el segundo y el tercero (aumento de 16 Hz entre el primer y el tercer año). Por otro lado, el F2 aumenta 55 Hz entre el primer y el segundo año de vida y 14 Hz entre el segundo y el tercero (incremento de 69 Hz entre el primer y el tercer año). En relación con la desviación estándar del resto de vocales, la desviación típica de la vocal posterior semialta [o] se asemeja más a la desviación de la vocal central baja [a] puesto que en ambas tanto la desviación del F1 como la del F2 aumentaron levemente con el paso del tiempo (lo mismo ocurrió con la desviación estándar del F1 de la vocal anterior alta [i]).

#### **7.1.2.5. Vocal posterior alta, [u].**

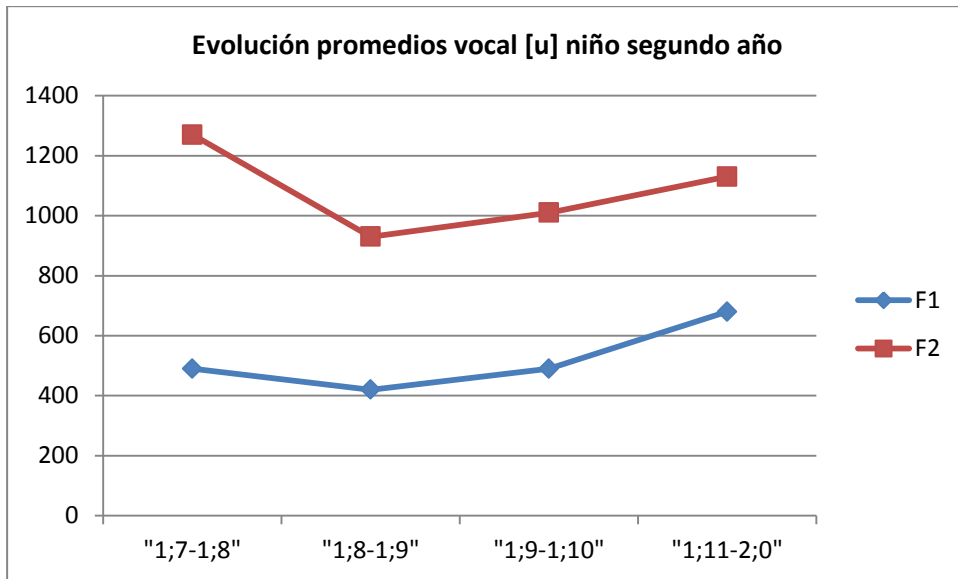
Al haber sido reconocida por cuatro de los cinco oyentes una única vocal posterior alta redondeada en el primer año de vida, el promedio del F1 y F2 se redujo a los valores formánticos de la misma sin que otros valores pudieran utilizarse para realizar dicho cálculo (gráfico 7.66).

**Gráfico 7.66. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el primer año de vida (0;4-1;0 año).**



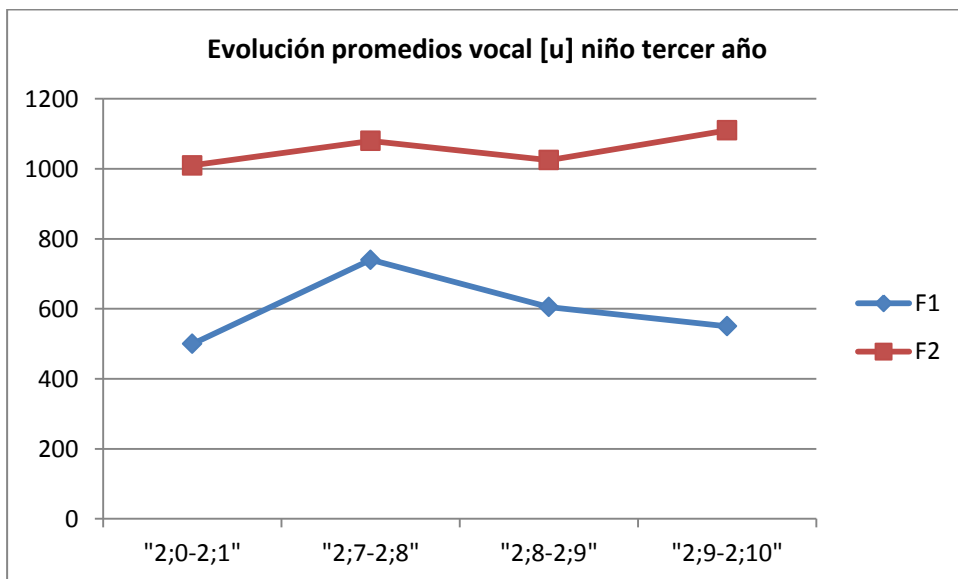
En el segundo año de vida cuatro vocales posteriores altas fueron identificadas por los oyentes (gráfico 7.67). La pertenencia de estas cuatro vocales a cuatro meses diferentes provocó que la media mensual se redujera como en el primer periodo a los valores de las mismas. Tal y como podemos observar los promedios de ambos formantes muestran una evolución prácticamente idéntica puesto que ambos comienzan descendiendo, registran el valor mínimo en el segundo mes analizado (420 Hz en el F1 y 930 en el F2), y vuelven a ascender de manera progresiva en los meses siguientes (alcanzando el F1 su valor máximo en el último mes analizado frente al F2 que lo registró en el primero). La diferencia en el F1 entre el primer mes analizado del segundo año de vida y el último del primero fue de 375 Hz (descenso entre el primer y el segundo año). Respecto al F2, este descendió 430 Hz en el mismo periodo.

**Gráfico 7.67. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el segundo año de vida (1;0-2;0 años).**



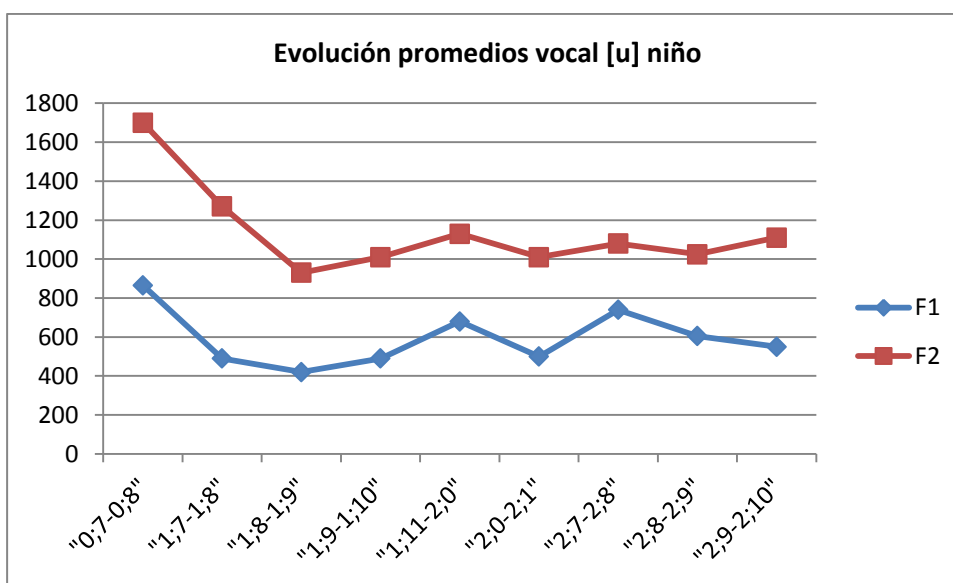
De nuevo un patrón similar vuelve a acercarse a la evolución de los promedios de los dos primeros formantes tal y como muestra el gráfico 7.68 que representa el tercer año de vida. En ambos formantes el promedio aumenta entre el primer y el segundo mes analizado y desciende en el transcurso hacia el tercero si bien la media del F1 sigue descendiendo hasta el final del periodo mientras que la del F2 aumenta. Respecto al último mes del periodo anterior, el promedio del F1 es 180 Hz menor y 120 Hz menor el del segundo.

**Gráfico 7.68. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años).**



La evolución de los promedios de ambos formantes explicada por separado en cada uno de los periodos del estudio longitudinal puede observarse de manera conjunta en el gráfico inferior (gráfico 7.69). En este destaca el drástico descenso producido tanto en el F1 como en el F2 entre los 0;8 y 1;8 años frente a los aumentos y descensos más escalonados que tienen lugar en los otros periodos.

**Gráfico 7.69. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en los tres primeros años de vida (0;4-3;0 años).**



El grado de descenso de los promedios de ambos formantes puede cuantificarse una vez revisados los valores concretos obtenidos en cada uno de los periodos que se exponen en las siguientes tablas.

**Tabla 7.52. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el primer año de vida (0;4-1;0 año).**

Edad	Media F1	Media F2
0;7-0;8	865	1700
Media anual	865	1700

**Tabla 7.53. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Media F1	Media F2
1;7-1;8	490	1270
1;8-1;9	420	930

1;9-1;10	490	1010
1;11-2;0	680	1130
Media anual	520	1085

**Tabla 7.54. Evolución de los promedios de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años).**

Edad	Media F1	Media F2
2;0-2;1	500	1010
2;7-2;8	740	1080
2;8-2;9	605	1025
2;9-2;10	550	1110
Media anual	592	1060

Como puede deducirse a partir de los valores aportados en la tabla, el promedio de los dos primeros formantes se redujo notoriamente entre el primer y el segundo año de vida (345 Hz el F1 y 615 Hz el F2). Entre la segunda y la tercera etapa la media del F1 aumentó 72 Hz mientras que la del F2 disminuyó 25 Hz. La diferencia en el F1 entre el primer y el tercer periodo fue de 273 Hz mientras que en el F2 fue de 640 Hz. Con el objetivo de averiguar si el descenso producido entre el primer y el segundo periodo estudiados era significativo, decidimos realizar un análisis de varianza que sin embargo presentó en esta ocasión una dificultad. El inconveniente residió en que el número de señales identificadas en el primer año de vida (una) si bien no impidió que un análisis de varianza pudiera ser realizado sí que pudo enmascararlo puesto que a pesar del notable descenso hallado entre el primer y el segundo año de vida tanto en el F1 como en el F2 (345 Hz y 615 Hz respectivamente), el análisis de varianza no mostró que un descenso significativo de los valores tuviera lugar entre esos periodos. No obstante, la importante reducción de 345 Hz en el primer formante y de 615 Hz en el segundo parece indicar que un descenso significativo de los valores formánticos tuvo lugar entre el primer y el segundo año de vida, y que este pudo haberse producido en el primer mes analizado del segundo periodo, es decir entre los 1;7 y los 1;8 años, ya que la diferencia entre este mes y el periodo anterior fue de 375 Hz para el F1 y de 430 Hz para el F2.

La posibilidad de que se trate de una reducción significativa puede inferirse también si tenemos en cuenta la disminución en los promedios anuales de otras vocales que presentaron más de una muestra anual para calcular el ANOVA. Regresando así a los puntos anteriores de

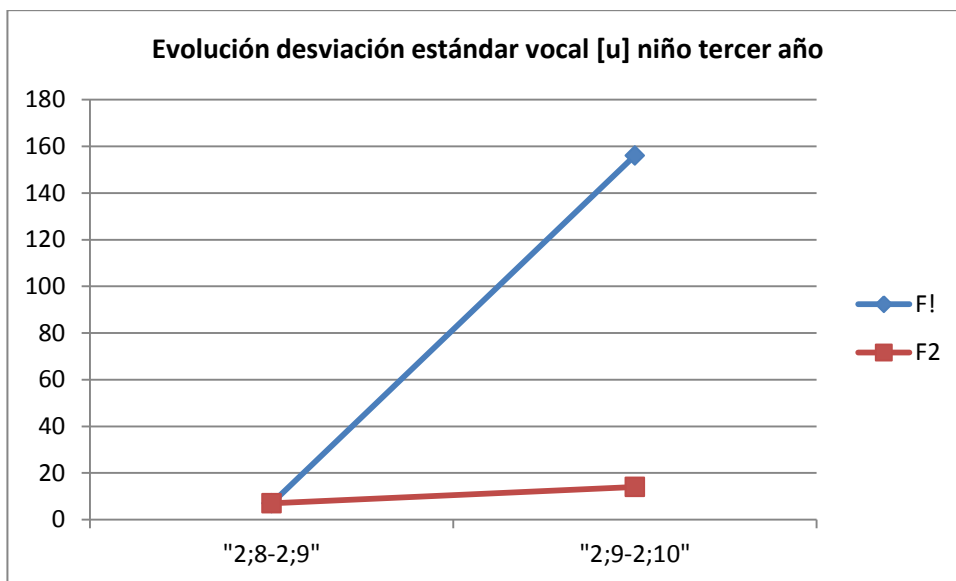
este apartado observamos que en el descenso significativo de la media del F2 entre el primer y el segundo año de vida en el caso de la vocal [a] (descenso que se produjo concretamente entre los 1;7 y los 1;8 años), la bajada en los valores fue de 438 Hz, y en el de la vocal [o] de 327 Hz (produciéndose este último entre los 1;9 y los 1;10 años). De este modo, si la reducción en el promedio del F2 en ambas vocales fue menor que la encontrada en la vocal actual [u] (de 615 Hz) y ello conllevó un descenso significativo de los valores, todo apunta a que el mismo descenso significativo podría producirse para la vocal posterior alta al manifestar esta una reducción aún mayor en los promedios anuales.

En el caso del F1, el hecho de que este primer formante no se haya reducido de manera significativa en ninguna de las vocales anteriores nos impide comparar el descenso del promedio del F1 de la vocal posterior alta entre el primer y el segundo año de vida con los datos de las otras vocales. No obstante, la reducción de 345 Hz entre el primer año de vida y el segundo, y la reducción de 375 Hz entre la media del único mes del primer año y el promedio del primer mes del segundo, parecen indicar que un descenso significativo tuvo lugar entre estos dos periodos.

Respecto al segundo análisis estadístico, el referente a la evolución de la desviación estándar a lo largo de los periodos estudiados, y debido a que este no pudo realizarse ni en el primer año de vida por identificarse en este tan solo una señal, ni en el segundo periodo por identificarse en este cuatro señales pero en meses diferentes, presentamos el gráfico que refleja la evolución de este cálculo estadístico en el tercer año de vida, periodo en el que pudo realizarse en al menos dos meses (gráfico 7.70). Tal y como podemos observar aunque ambos formantes parten de la misma desviación (7 Hz), el F1 aumenta de manera considerable (149 Hz) mientras que el F2 apenas lo hace (7 Hz).

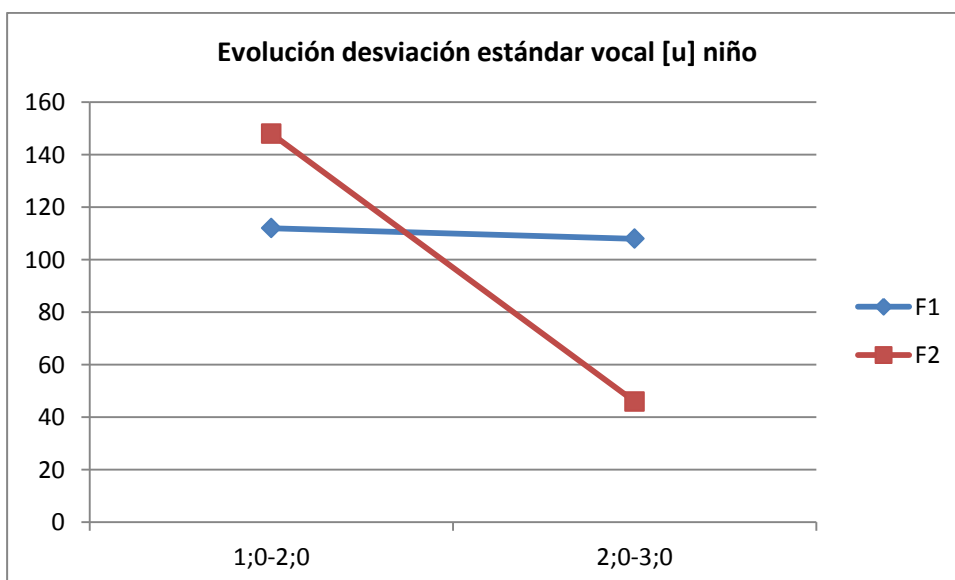


**Gráfico 7.70. Evolución de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años).**



A continuación exponemos la evolución de la desviación estándar anual calculada en el segundo y el tercer año de vida (gráfico 7.71). Aunque la evolución de las desviaciones de ambos formantes es similar puesto que ambas tienden al descenso, este descenso es mucho más notorio en el caso del F2 que disminuye de los 148 Hz a los 46 Hz, que en el del F1 que lo hace de los 112 Hz a los 108 Hz.

**Gráfico 7.71. Evolución de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en los dos últimos años de vida (1;0-3;0 años).**



Las tablas inferiores recogen los valores exactos obtenidos tras este segundo análisis:

**Tabla 7.55. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en el segundo año de vida (1;0-2;0 años).**

Edad	Desviación estándar F1	Desviación estándar F2
Desviación estándar anual	112	148

**Tabla 7.56. Valores de la desviación estándar de los dos primeros formantes de la vocal [u] del niño en el tercer año de vida (2;0-3;0 años).**

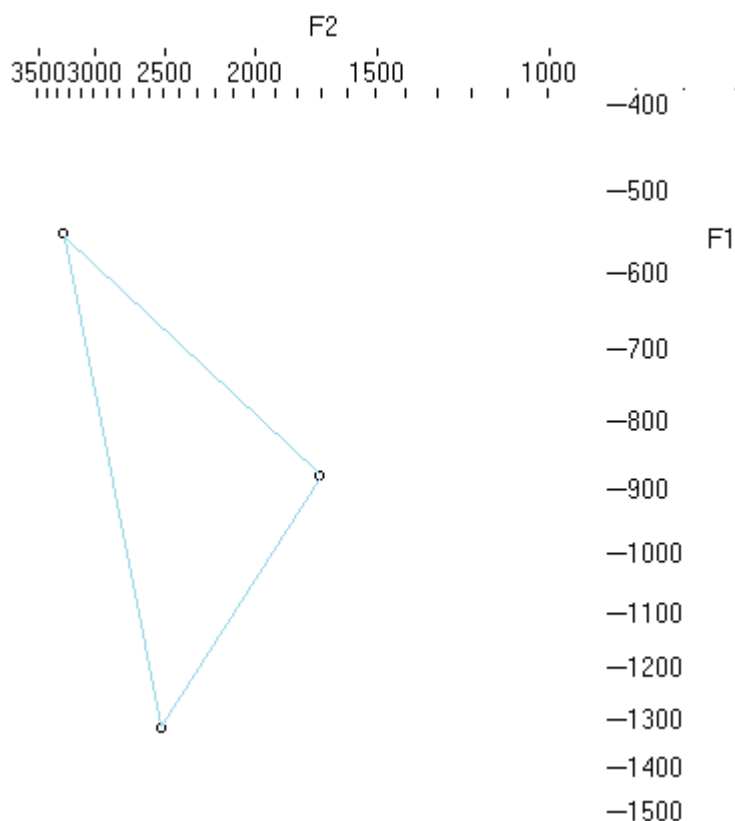
Edad	Desviación estándar F1	Desviación estándar F2
2;8-2;9	7	7
2;9-2;10	156	14
Desviación estándar anual	108	46

En cuanto a la tendencia evolutiva de la desviación, esta es en ambos formantes hacia el descenso si bien como ya hemos comentado este es mucho más acentuado en el F2 que en el F1. En relación a otras vocales, observamos que tanto la desviación estándar del F1 como la del F2 son similares a las de las vocales anteriores aunque la desviación estándar del F2 en el tercer año de vida es la más baja de todas.

#### ***7.1.2.6. Evolución del área vocálica a partir de los valores obtenidos en las señales del niño.***

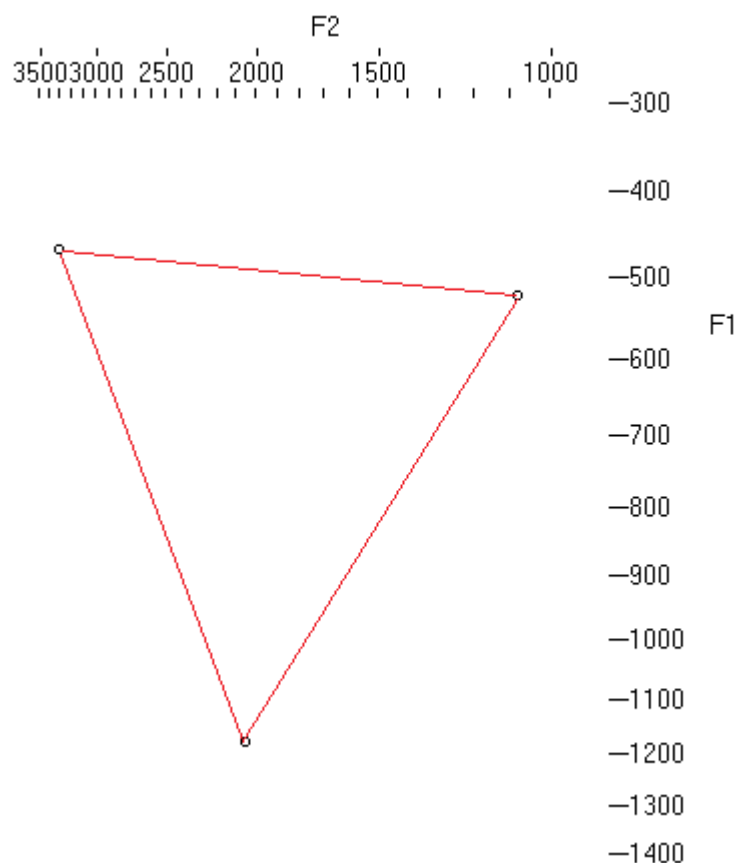
El triángulo vocálico en este primer año de vida analizado en el niño se situó a la izquierda y en la parte baja del gráfico debido a los elevados valores del F1 y del F2 de las tres vocales representadas (gráfico 7.72):

**Gráfico 7.72. Área vocálica del niño en el primer año de vida (0;4-1;0 año).**



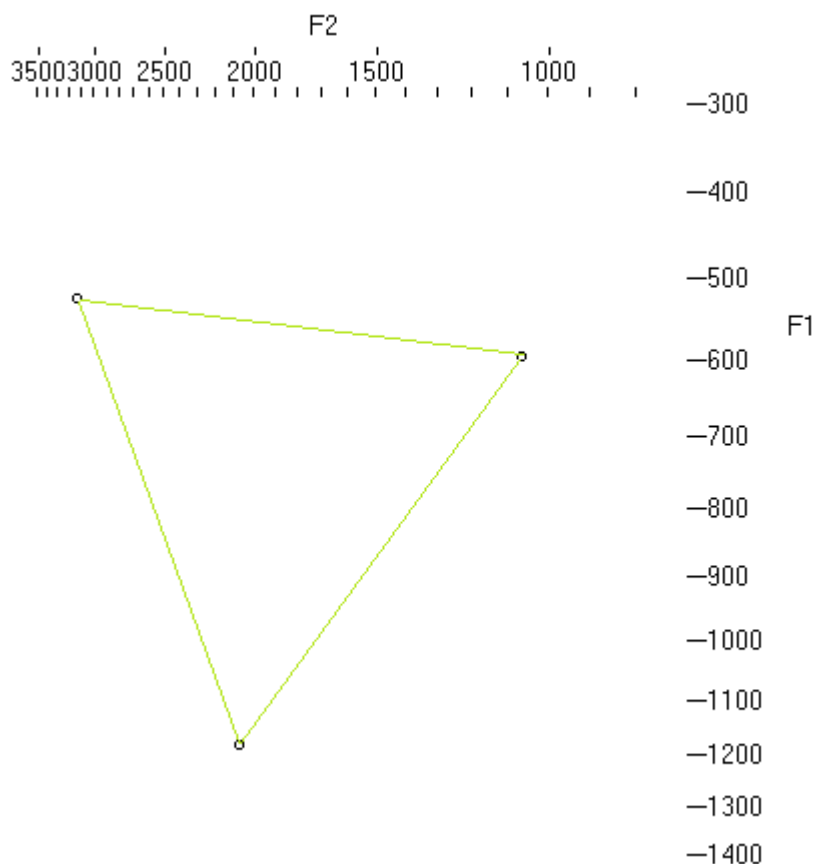
La notable reducción de los valores del F2 en la vocal central baja [a] y en la posterior alta [u] provocó que el triángulo vocálico en el segundo año de vida se desplazara un tanto hacia la derecha como muestra el gráfico inferior (gráfico 7.73). Sin embargo, quizás más que desplazarse y tal y como ocurrió con el área vocálica de la niña del segundo periodo, el descenso del segundo formante de estas vocales provocó un alargamiento o estiramiento del triángulo hacia la parte derecha del gráfico manteniéndose los valores del F2 de la [i] en la parte izquierda del mismo. La reducción significativa en los promedios del F2 de la [a] y la [u], y la falta de esta en los de la [i] conllevó por otro lado un aumento del área vocálica en este segundo año de vida con respecto a la del primero (área de 464758 y de 725196 respectivamente). En relación a la posición del triángulo vocálico en el eje vertical, el descenso en este segundo año de vida de los valores del primer formante en todas las vocales supuso que dicho triángulo se localizara en una posición más elevada con respecto al periodo anterior.

**Gráfico 7.73. Área vocálica del niño en el segundo año de vida (1;0-2;0 años).**



La falta de acuerdo entre los oyentes a la hora de reconocer la vocal baja central [a] en el tercer año de vida hizo como vimos en el apartado 6.2.2.3 que no contásemos con muestras de esta vocal en el tercer periodo analizado. La imposibilidad de extraer el promedio de esta vocal y la necesidad de trazar el área vocálica del niño en este intervalo temporal nos llevaron a utilizar los promedios formánticos hallados para dicha vocal en el segundo año de vida. De esta forma, si bien el gráfico siguiente no reflejó el área vocálica real, sí permitió que nos formásemos una idea de los cambios sufridos en esta con respecto al segundo y al primer año de vida al menos para las vocales [i] e [u]. Como puede observarse en el gráfico 7.74, la similitud en los valores de los promedios de ambos formantes entre el segundo y el tercer año de vida hizo que la figura del triángulo del tercer periodo fuera muy similar a la del segundo aunque una pequeña reducción del área vocálica se detectó tras comparar el área del segundo año de vida con la del tercero (725196 y 632978 respectivamente). En relación a la altura, al aumentar levemente los promedios del F1 de la [i] y de la [u], el área vocálica se situó ligeramente por debajo de la del año anterior.

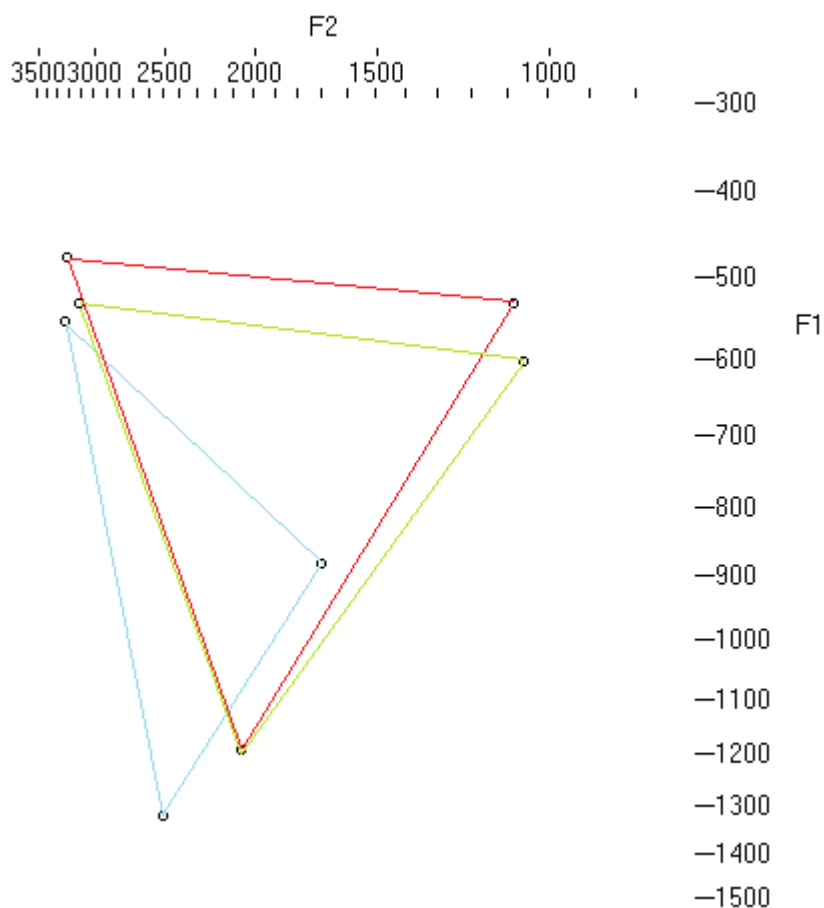
**Gráfico 7.74. Área vocálica del niño en el tercer año de vida (2;0-3;0 años).**



El último gráfico (7.75) de este punto recoge la evolución del área vocálica calculada sobre los valores obtenidos en las señales del niño a lo largo de los tres periodos del estudio. Como podemos observar, la ubicación a la izquierda del gráfico que presentó el área del primer periodo analizado disminuyó con el paso del tiempo quedando los triángulos vocálicos posteriores más centrados en el gráfico debido al alargamiento de dicha área por la falta de descenso significativo del F2 de la vocal anterior alta [i]. Respecto al segundo periodo analizado, encontramos que la ya comentada reducción significativa de los valores del F2 en las vocales [a] y [u] y la ausencia de esta en la [i] provocaron que el área vocálica aumentara con respecto a la del primer periodo analizado tal y como se observa en el gráfico. En cuanto al tercer año de vida y debido como comentábamos a la ausencia de cambios significativos en los promedios de ambos formantes, el triángulo que se observa es muy similar al del segundo año aunque se sitúa un poquito más hacia la derecha del gráfico debido al leve descenso del F2 de la [i] y de la [u] que pasaron de 3224 Hz a 3115 Hz y de 1085 Hz a 1060 Hz respectivamente. En cuanto a la posición en el eje vertical, el triángulo vocálico del primer año de vida fue el que ocupó el lugar más bajo debido a los valores elevados del primer formante en todas las vocales. Dichos promedios provocaron al disminuir que los triángulos vocálicos del segundo y del tercer

año de vida se situaran más arriba en el gráfico. La ausencia de la vocal central baja en el tercer periodo hizo que tomásemos los valores de esta vocal calculados en el segundo año de vida, circunstancia que nos impidió observar si también el ángulo correspondiente a dicha vocal se reducía o aumentaba en el último año del estudio.

**Gráfico 7.75. Evolución del área vocálica del niño en los tres primeros años de vida (0;4-3;0 años).**



### 7.1.3. Comparación de los resultados obtenidos en cada uno de los sujetos.

Dedicaremos este punto a la comparación de los resultados obtenidos tras la realización de los diferentes análisis estadísticos sobre las señales identificadas con el mismo timbre vocálico por los oyentes, en cada uno de los meses analizados del estudio longitudinal. Para este apartado incluimos nuevas tablas que recogen el promedio formántico y la desviación estándar anual para cada una de las vocales analizadas en ambos niños (datos que han sido ofrecidos a lo largo de este capítulo en diferentes secciones), sin otra finalidad que la de facilitar la equiparación de los resultados. Y siguiendo con la estructuración de los apartados anteriores,

comenzamos mostrando los resultados anuales obtenidos en cada uno de los niños para la vocal central baja [a] (tablas 7.57 y 7.58):

**Tabla 7.57. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [a] de la niña.**

Niña				
Vocal [a]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	1103	2359	141	128
1;0-2;0	1120	2128	76	108
2;0-3;0	970	1915	184	120

**Tabla 7.58. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [a] del niño.**

Niño				
Vocal [a]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	1299	2498	98	21
1;0-2;0	1173	2060	170	191

Respecto al promedio del primer formante, este fue muy similar en ambos niños en el segundo año de vida mientras que en el primero el promedio hallado en el niño superó en 196 Hz al calculado en la niña (resultando imposible la comparación de los valores en el tercer periodo debido a la ausencia de aes en el tercer año de vida del niño). Quizás entre las causas se encuentre el diferente número de señales identificadas en ambos sujetos (veinte en la niña y tan solo cuatro en el niño). En cuanto al análisis de varianza, el ANOVA reveló que en ninguno de los sujetos se produjo un descenso significativo del F1 con el paso del tiempo.

En relación a la media del F2, los promedios hallados en ambos niños fueron muy similares. En lo que toca a la evolución, los análisis de varianza revelaron que el descenso del segundo formante que tuvo lugar con el paso del tiempo en ambos sujetos fue significativo. La única diferencia hallada a este respecto residió en el punto temporal del segundo año de vida a partir del cual los valores del segundo formante descendieron significativamente: entre los 1;5 y 1;6 años en la niña y entre los 1;7 y 1;8 años en el niño.

En cuanto al segundo análisis estadístico, los resultados obtenidos fueron muy similares tanto para el F1 como para el F2 encontrándose una diferencia de alrededor de 100 Hz en cada uno de los periodos. La evolución de dicha desviación fue similar ya que aumentó levemente con el paso del tiempo en ambos sujetos, excepto en el F2 de la niña donde disminuyó ligeramente.

En relación a la primera vocal anterior analizada, la vocal anterior semialta [e], las siguientes tablas resumen los resultados obtenidos (tablas 7.59 y 7.60):

**Tabla 7.59. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [e] de la niña.**

Niña				
Vocal [e]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	795	2876	143	181
1;0-2;0	768	2873	139	169
2;0-3;0	703	2914	99	112

**Tabla 7.60. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [e] del niño.**

Niño				
Vocal [e]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	751	2906	115	170
1;0-2;0	787	2860	117	189
2;0-3;0	660	2716	58	107

Si nos fijamos en los promedios formánticos de las tablas anteriores (tablas 7.59 y 7.60), observaremos que una similitud en los valores se dio en ambos niños tanto en el F1 como en el F2 a lo largo de los tres periodos estudiados excepto en el promedio del segundo formante en el tercer año de vida que fue 198 Hz superior en la niña. Por otro lado, los valores de los dos primeros formantes tendieron a disminuir en ambos sujetos con el paso del tiempo (excepto los del F2 de la niña que aumentaron ligeramente entre el segundo y el tercer año de vida, y los del F1 del niño que si bien descendieron a lo largo de los tres periodos analizados aumentaron



sutilmente entre el primer y el segundo año de vida), pero este descenso no fue acusado en ninguno de los sujetos a excepción del F2 del niño que disminuyó 190 Hz entre el primer y el tercer año de vida.

Así lo demuestran los análisis de varianza que no encontraron ninguna disminución significativa del F1 a lo largo del tiempo en ninguno de los sujetos salvo en las circunstancias ya explicadas en el apartado 7.1.1.2, donde a pesar de que el ANOVA detectó que un descenso significativo se producía en los promedios de la niña entre el primer y el tercer año de vida, la diferencia inferior a cien hertzios entre ambos periodos parecía indicar que dicho descenso pudo producirse por el diferente número de señales entre uno y otro periodo, y en las circunstancias señaladas en el apartado 7.1.2.2, donde el descenso significativo que se produjo entre los cuatro primeros y los cuatro últimos meses del primer año de vida del niño pudo deberse a la agrupación de señales con un elevado F1 en la primera parte del periodo ya que en el segundo año de vida el promedio del F1 no solo no fue igual o inferior al del primer año sino que fue superior.

Respecto a la evolución del F2, el ANOVA no halló ningún descenso significativo de este formante en ninguno de los sujetos ni en ninguno de los periodos equiparados excepto entre el primer y el tercer año de vida del niño donde el análisis de varianza reveló que una reducción significativa de este formante tenía lugar. No obstante, y tal y como apuntábamos en el apartado 7.1.2.2, la notable diferencia en el número de señales entre el primer y el tercer año de vida (veintisiete frente a cinco) y la coincidencia en el tercer año de algunas vocales anteriores semialtas con un F2 algo inferior que se impusieron en la media al ser tan reducido el número de señales de este periodo, pudieron quizás influir en este resultado.

La tendencia hacia el descenso también se produjo en la desviación estándar de ambos formantes y de ambos niños. Los valores formánticos de la vocal anterior semialta tendieron por tanto a homogeneizarse en ambos sujetos con el paso del tiempo.

Muy similares fueron las conclusiones extraídas para la vocal anterior alta [i] tal y como se desprende de las siguientes tablas (7.61 y 7.62):

**Tabla 7.61. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [i] de la niña.**

Niña				
Vocal [i]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	609	3304	136	201
1;0-2;0	629	3304	49	186
2;0-3;0	536	3160	104	150

**Tabla 7.62. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [i] del niño.**

Niño				
Vocal [i]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	541	3246	85	157
1;0-2;0	465	3224	87	146
2;0-3;0	520	3115	92	129

Los promedios formánticos de ambos niños se situaron alrededor de los mismos valores tanto en el primer como en el segundo formante a lo largo de los tres periodos analizados. Puede apreciarse además en ambos sujetos un sutil descenso de los promedios de ambos formantes tras comparar los valores del primer año con los del tercero (excepto el leve aumento del F1 de la niña entre el primer y el segundo año). No obstante, y tal y como muestran los ANOVA, el descenso no fue significativo en ninguno de los formantes de forma que los resultados hallados para la vocal anterior alta [i] son similares a los encontrados en la vocal anterior semialta [e] en ambos niños.

Respecto a la desviación estándar, el grado de apartamiento de la media de los valores de la vocal [i] fue muy similar en ambos niños y se redujo con el paso del tiempo salvo por el leve ascenso que experimentó dicha desviación entre el segundo y el tercer año de vida del niño.

Los valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal posterior semialta [o], recogidos en las tablas inferiores 7.63 y 7.64,

fueron muy similares en ambos sujetos aunque el promedio del F1 de la niña en el primer año de vida fue 159 Hz superior al del niño.

**Tabla 7.63. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [o] de la niña.**

Niña				
Vocal [o]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	952	1878	116	142
1;0-2;0	794	1582	64	140
2;0-3;0	772	1502	78	97

**Tabla 7.64. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [o] del niño.**

Niño				
Vocal [o]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	793	1863	93	47
1;0-2;0	799	1536	103	102
2;0-3;0	774	1467	109	116

En relación al primer análisis estadístico y tras equiparar los resultados obtenidos en cada uno de los niños, nos percatamos de dos circunstancias: por un lado, advertimos que las medias anuales extraídas en cada niño fueron muy similares en ambos formantes y en todos los periodos analizados excepto el promedio del primer formante en el primer año de vida que como ya hemos comentado fue superior en la niña (159 Hz); por otro lado, observamos que el promedio del segundo formante se redujo de manera notoria en ambos niños entre el primer y el segundo año de vida.

Respecto a la primera circunstancia, el análisis de varianza concluyó que ningún descenso significativo del F1 tuvo lugar en ninguno de los niños a lo largo del tiempo salvo en la excepción comentada, esto es, entre el primer y el segundo año de vida de la niña. En este periodo de tiempo el ANOVA manifestó que el F1 descendía de manera significativa aunque al intentar establecer en qué mes concreto del segundo año de vida se producía este descenso, la

escasez de señales analizadas en cada uno de los meses del segundo periodo impidió que este análisis indicara el momento temporal en que este se producía. No obstante, la diferencia de 232 Hz entre el promedio del primer año de vida y el obtenido entre los 1;9 y 1;10 años parece indicar que es en este mes cuando tiene lugar el descenso del formante. En el niño sin embargo, el promedio del F1 no descendió de manera significativa y quizás entre las causas de esta disimilitud se encuentren el diferente número de señales identificadas en cada uno de los sujetos en el primer año de vida (doce en la niña y tres en el niño), y el hecho de que en el caso de la niña algunas de estas señales se apartaran más de la media y presentaran un F1 más alto que elevó la media mensual.

En cuanto al F2, este descendió de manera significativa entre el primer y el segundo año de vida en ambos niños produciéndose este descenso entre los 1;8 y los 1;9 años en la niña y entre los 1;9 y 1;10 años en el niño.

Después de calcular la desviación estándar observamos que los resultados obtenidos fueron similares en ambos sujetos (la diferencia fue menor de 100 Hz en cada uno de los periodos analizados) aunque la evolución de los mismos fue diferente ya que en la niña la desviación típica de los dos primeros formantes disminuyó levemente con el paso del tiempo mientras que en el niño esta aumentó en el transcurso de un periodo a otro pero también de forma sutil.

Por lo que respecta a la última vocal analizada, la posterior alta [u], estas son las tablas que recogen los resultados (tablas 7.65 y 7.66):

**Tabla 7.65. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [u] de la niña.**

Niña				
Vocal [u]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	915	1800	134	184
1;0-2;0	640	1170		
2;0-3;0	611	1251	102	176

**Tabla 7.66. Valores anuales de los promedios y de la desviación estándar de los dos primeros formantes de la vocal [u] del niño.**

Niño				
Vocal [u]				
Edad	Promedio F1	Promedio F2	DS F1	DS F2
0;4-1;0	865	1700		
1;0-2;0	520	1085	112	148
2;0-3;0	592	1060	108	46

Si observamos las tablas anteriores encontramos que tanto los valores como las evoluciones de los promedios de los dos primeros formantes son similares en ambos niños. Respecto a la media de los dos primeros formantes, esta descendió de manera notoria entre el primer y el segundo año de vida en ambos sujetos de modo que a pesar de la imposibilidad en el caso de la niña de realizar el ANOVA debido a la escasez de señales (apartado 7.1.1.5) y a las dificultades que este análisis ofreció en el caso del niño debido también al reducido número de muestras (apartado 7.1.2.5), el descenso en el primer formante de 275 Hz en el caso de la niña y de 345 Hz en el del niño, y el descenso en el segundo formante de 630 Hz en el caso de la niña y de 615 Hz en el caso del niño, parecen indicar que el descenso que tuvo lugar fue significativo en ambos sujetos.

Los resultados obtenidos tras el cálculo de la desviación estándar fueron muy similares en ambos sujetos así como la tendencia de esta a reducirse con el paso del tiempo.

Finalmente, hemos de comentar la evolución de las áreas vocálicas tras la equiparación de las mismas en los dos niños. En el primer periodo temporal el triángulo vocálico se situó en ambos sujetos en la parte izquierda del gráfico debido a los elevados valores del F2 de todas las vocales en este primer año de vida. Con el paso del tiempo, el segundo formante de las vocales [a] y [u] se desplazó hacia la derecha del gráfico reflejando el descenso significativo de los valores de este formante y ocasionando el consiguiente alargamiento del triángulo vocálico. Esta reducción formántica notoria para la vocal central baja y para la vocal posterior alta, y la ausencia de la misma para la vocal anterior alta conllevó que en ambos niños el área vocálica aumentara entre el primer y el tercer periodo estudiados (constituyéndose como punto de anclaje el promedio del F2 de la vocal [i] que “tiraba” del área vocálica ante el desplazamiento de las vocales [a] y [u] hacia la derecha del gráfico por la ya comentada reducción del segundo formante).

Diferentes fueron por otro lado las evoluciones del área vocálica respecto al eje vertical de la carta de formantes. En el caso de la niña el leve aumento del F1 de la [i] y de la [a] entre el primer y el segundo año de vida provocó que el triángulo vocálico descendiera sutilmente en el gráfico con respecto al periodo anterior, mientras que en el del niño la reducción de los promedios formánticos de todas las vocales en este intervalo temporal originó que el triángulo vocálico ascendiera ligeramente en la carta de formantes (no obstante y tal y como se pudo comprobar en las cartas de formantes, el aumento de los valores en la niña y el descenso de los mismos en el niño fue muy poco notorio de forma que las áreas vocálicas de ambos sujetos se situaron en posiciones muy cercanas). Entre el segundo y el tercer año de vida el promedio del F1 descendió en la niña en las tres vocales analizadas lo que supuso una elevación del triángulo vocálico en el gráfico formántico de este periodo. En el niño por el contrario, el F1 aumentó en las vocales [i] e [u] (manteniéndose en el caso de la [a] por las causas ya explicadas) por lo que el triángulo vocálico del tercer periodo se desplazó levemente hacia la parte inferior del gráfico en comparación con el triángulo de la segunda etapa.

## 8. DISCUSIÓN.

Dedicaremos este capítulo a la revisión de las hipótesis de partida que planteábamos al comienzo del estudio para verificar si quedan confirmadas a partir de nuestros resultados de la misma forma que fueron corroboradas por distintos autores en estudios de adquisición de otras lenguas. Recuperaremos por tanto algunos de los múltiples trabajos a los que aludimos en el capítulo del estado de la cuestión para relacionar nuestra investigación por un lado con los estudios en los que dichas hipótesis se corroboraron, y para establecer comparativas por otro entre nuestros datos y los hallados en dichos trabajos. Es este por tanto un capítulo en el que quedarán relacionados la mayor parte de los contenidos expuestos a lo largo de la tesis constituyéndose de este modo como la cohesión más que conveniente en todo estudio de considerables dimensiones. No obstante, en este apartado no volveremos a exponer las distintas investigaciones de manera exhaustiva sino que tomaremos los datos de estas que nos permitan establecer el parangón con nuestros resultados y corroborar las hipótesis. Comprobamos ahora una a una cada una de las conjeturas:

La primera hipótesis que planteábamos sostenía que los valores de los formantes de los bebés superaban con creces los valores de los formantes de los adultos. Para comprobarlo presentamos a continuación dos tablas en las que los profesores Martínez Celdrán y Fernández Planas (2007: 175-177) recogieron los valores formánticos para cada una de las vocales en la voz adulta femenina y masculina después de haberlos medido en diez jóvenes universitarios hablantes de español estándar, cinco de cada sexo, cuyas edades estaban entre los 20 y los 30 años (el experimento consistió en la pronunciación de cada una de las vocales del español en un contexto tónico, siempre el mismo, cuya única variación estaba en la consonante que precedía a la vocal y que podía ser oclusiva sorda o sonora). He aquí los valores obtenidos en dicho experimento y los hallados tras nuestro análisis (tablas 8.1-8.8):

**Tabla 8.1. Promedios de los valores formánticos de las vocales de un adulto mujer (Martínez Celdrán y Fernández Planas, 2007).**

Valores de las vocales de un adulto mujer		
Vocales	F1	F2
a	886	1712
e	576	2367
i	369	2685
o	586	1201
u	390	937

**Tabla 8.2. Promedios de los valores formánticos de las vocales de la niña entre los 0;4-1;0 año.**

Valores de las vocales de la niña/ 0;4-1;0 años		
Vocales	F1	F2
a	1103	2359
e	795	2876
i	609	3304
o	952	1878
u	915	1800

**Tabla 8.3. Promedios de los valores formánticos de las vocales de la niña entre los 1;0-2;0 años.**

Valores de las vocales de la niña /1;0-2;0 años		
Vocales	F1	F2
a	1120	2128
e	768	2873
i	629	3304
o	794	1582
u	640	1170

**Tabla 8.4. Promedios de los valores formánticos de las vocales de la niña entre los 2;0-3;0 años.**

Valores de las vocales de la niña /2;0-3;0 años		
Vocales	F1	F2
a	970	1915
e	703	2914
i	536	3160
o	772	1502
u	611	1251



**Tabla 8.5. Promedios de los valores formánticos de las vocales de un adulto varón (Martínez Celdrán y Fernández Planas, 2007).**

Valores de las vocales de un adulto varón		
Vocales	F1	F2
a	699	1471
e	457	1926
i	313	2200
o	495	1070
u	349	877

**Tabla 8.6. Promedios de los valores formánticos de las vocales del niño entre los 0;4-1;0 año.**

Valores de las vocales del niño / 0;4-1;0 años		
Vocales	F1	F2
a	1299	2498
e	751	2906
i	541	3246
o	793	1863
u	865	1700

**Tabla 8.7. Promedios de los valores formánticos de las vocales del niño entre los 1;0-2;0 años.**

Valores de las vocales del niño / 1;0-2;0 años		
Vocales	F1	F2
a	1173	2060
e	787	2860
i	465	3224
o	799	1536
u	520	1085

**Tabla 8.8. Promedios de los valores formánticos de las vocales del niño entre los 2;0-3;0 años.**

Valores de las vocales del niño / 2;0-3;0 años		
Vocales	F1	F2
a	No hay aes	No hay aes
e	660	2716
i	520	3115
o	774	1467
u	592	1060

Como podemos observar, los valores de los dos primeros formantes de ambos bebés superan a los valores de los formantes de los adultos en todas las vocales tanto si hacemos distinción de sexo, y comparamos los obtenidos en nuestro sujeto de sexo femenino con los del adulto mujer y los de nuestro sujeto de sexo masculino con los del adulto hombre, como si no lo hacemos. Entre las razones encontramos la configuración particular del tracto vocal del bebé que expusimos en el apartado 2.1.1 del estado de la cuestión y que resumimos a continuación con el fin de justificar esta primera hipótesis planteada.

Las principales características anatómicas del aparato fonador del bebé que asemejan dicho aparato al tracto vocal de un primate no humano más que al tracto vocal de un adulto humano y que son las responsables de esta diferencia de valores formánticos se encuentran tanto entre los 0 y los 0;4 años como después de esta edad y una vez producida la reestructuración anatómica. Estas son principalmente: el diferente tamaño del tracto vocal entre los bebés y los adultos como expuso en su día Fant (1960) y como fue refrendado después por otros autores tales como Stevens y House (1961) según los cuales al aumentar la longitud del tracto vocal tienden a reducirse las frecuencias de todas las resonancias mientras que el acortamiento de este último aumenta las frecuencias de todas las resonancias (recordemos que según Ménard y Boë [2004] el tracto vocal mide unos 8 cm en el recién nacido frente a los 17 cm del adulto varón); una alta posición de la laringe (según Fried [1983] esta se encuentra entre la primera y la tercera vértebra cervical y según Sapienza et al. [2004] en su desarrollo va pasando de una forma cónica a una más cilíndrica); un acoplamiento entre esta y la naso-faringe y una aproximación de la velo-faringe y la epiglotis que condicionan al bebé como un respirador y vocalizador nasal (además es esta particular disposición de la laringe y del emparejamiento epiglótico uno de los factores que más asemejan el tracto vocal del infante al de otros mamíferos no humanos como el del chimpancé); las cuerdas vocales del recién nacido miden 2.5-3.0 mm según Sapienza et al.

(2004) y 4 mm según Tizte (1989), frente a los aproximadamente 17-21 mm que miden en los adultos masculinos y los 11-15 mm que miden en los femeninos según Hirano et al. (1983) (según Tizte [1989] estas miden 18 mm en los varones y 12 mm en las mujeres). Estas además tienen una inclinación más hacia abajo desde la parte posterior a la anterior; entre los 0;4 y los 0;6 años tiene lugar la ya explicada reestructuración anatómica que supone según Eckel et al. (2000), y entre otros cambios, un descenso en la laringe, el desacoplamiento entre esta y la naso-faringe (que permite al bebé respirar por la nariz y reducir el número de vocalizaciones nasales), la separación del velo y la epiglotis, ajustes de tensión, longitud y forma de las cuerdas vocales y el aumento progresivo de la porción cartilaginosa frente a la prominencia de la parte membranosa en las primeras semanas de vida.

Todas estas características junto con las diferencias halladas en la función y en las vías respiratorias (Polgar y Weng [1979]) hacen según Kent y Bauer (1985) y Kent y Vorperian (2006) que el tracto vocal del bebé no sea simplemente una miniatura del tracto vocal del adulto sino que sea además, funcionalmente diferente. Finalmente y tal y como advertimos en el estado de la cuestión, según Hirano et al. (1983), Kent (1984), Polger y Weng (1979) y Kent y Miolo (1995) las modificaciones de los tejidos óseos, de la lengua, de los cartílagos y las dimensiones del conducto continúan desarrollándose y acercándose a las del adulto durante el primer decenio de vida aunque no se alcanzará el tamaño adulto hasta la edad de entre los siete y los dieciocho años según Kent y Vorperian (1995) y Lieberman et al. (2001).

Por otro lado, y antes de continuar con la segunda hipótesis de partida, creemos conveniente detenernos brevemente en verificar si los valores dados para el primer y el segundo formante en las vocales del español y en las de otras lenguas superan también a los de los adultos, y detenernos además en comprobar si estos son similares a los nuestros.

Comenzamos pues presentando los promedios de los dos primeros formantes de las cinco vocales españolas hallados en el único estudio que ha realizado este análisis en lengua española con sujetos de tan corta edad. Se trata del trabajo de Casal et al. (2002) en el que se estableció una comparación entre los valores de los formantes de veintidós niños con un desarrollo fisiológico y lingüístico normal y veintidós pacientes pediátricos que se dividían en cuatro grupos: seis niños que presentaban el labio partido, siete que presentaban el paladar partido, siete pacientes con el paladar y el labio partido de manera unilateral y dos pacientes que mostraban el paladar y el labio partido de un modo completo. La media de edad de los niños fue de 2;3 años y todos habían sido operados habiéndose producido la última intervención casi a los 0;7 años. Además todos los pacientes recibieron apoyo psicológico y se sometieron a una terapia de habla para mejorar la producción y la coordinación de los gestos de esta. Las vocales

analizadas acústicamente fueron las cinco vocales del español que aparecieron en la mayoría de las ocasiones entre dos consonantes y cuando no se encontraron en esta posición se tomaron las vocales iniciales o finales. Recogemos en las tablas siguientes (tablas 8.9-8.12) el promedio de los dos primeros formantes que presentaron los niños con desarrollo normal a la edad media de 2;3 años (los valores de Casal et al. [2002] aparecen redondeados tal y como hicimos con los nuestros):

**Tabla 8.9. Promedios de los valores formánticos de las cinco vocales del español en los seis sujetos de desarrollo normal cuyos formantes se compararon con los de los niños con labio partido (Casal et al., 2002).**

Vocal	F1	F2
a	880	1912
e	707	2307
i	593	2967
o	653	1487
u	370	893

**Tabla 8.10. Promedios de los valores formánticos de las cinco vocales del español en los siete sujetos de desarrollo normal cuyos formantes se compararon con los de los niños con paladar partido (Casal et al., 2002).**

Vocal	F1	F2
a	754	1554
e	813	2107
i	507	2053
o	660	1360
u	540	1320

**Tabla 8.11. Promedios de los valores formánticos de las cinco vocales del español en los siete sujetos de desarrollo normal cuyos formantes se compararon con los de los niños que presentaron labio y paladar partido de manera unilateral (Casal et al., 2002).**

Vocal	F1	F2
a	913	1493
e	696	2616
i	488	2920
o	616	1360
u	488	800

**Tabla 8.12. Promedios de los valores formánticos de las cinco vocales del español en los dos sujetos de desarrollo normal cuyos formantes se compararon con los de los niños que presentaron labio y paladar partido de manera completa (Casal et al., 2002).**

Vocal	F1	F2
a	820	1480
e	640	2420
i	520	3480
o	480	1520
u	No hay úes	No hay úes

Si comparamos las tablas que recogen los valores de los dos primeros formantes en los análisis de Casal et al. (2002) (tablas de 8.9 a 8.12) con las que presentan los valores medios hallados por nosotros (tablas de 8.2 a 8.8) observamos que los resultados para cada una de las vocales son muy similares. En cuanto a las disimilitudes más notorias, estas quizás puedan explicarse por la diferencia de valores formánticos entre los distintos sujetos de cada una de las comparaciones tal y como refleja en esos casos la alta desviación estándar (en nuestro estudio estas son más reducidas al haber solo un sujeto en cada grupo). De este modo, la diferencia de al menos 400 Hz en el F2 de la vocal [e] hallada entre nuestros resultados (los obtenidos entre el segundo y el tercer año de vida) y aquellos en los que se comparó a los niños con desarrollo normal con aquellos que presentaron el labio partido (tabla 8.9) puede explicarse si observamos la desviación estándar que presentó este formante en el grupo de niños con desarrollo normal de Casal et al. (2002) y que fue de 640 Hz, valor que puede indicar que hubo niños que sí presentaron un segundo formante mucho más alto del que se dio como valor medio (2307 Hz) y posiblemente más cercano a los valores de nuestros sujetos.

La misma explicación puede aducirse para la diferencia entre los valores del segundo formante de las vocales [a], [e] e [i] que presentaron los niños con desarrollo normal del estudio de Casal et al. (2002), cuando se compararon sus valores con los de los sujetos que presentaron el paladar partido, y los valores encontrados en los sujetos de nuestro estudio, valores estos últimos mucho más elevados (desviación estándar de 421 Hz para la [a], de 625 Hz para la [e] y de 981 Hz para la [i]). Quizá sea el mismo razonamiento el que esté detrás de las diferencias encontradas entre los valores de los sujetos de nuestro estudio y los de los niños con desarrollo normal que se compararon con los sujetos que presentaron labio y paladar partido de manera completa.

A continuación ofrecemos una comparación de nuestros resultados con los obtenidos en otros trabajos cuyos sujetos adquirirían una lengua diferente al idioma español. En el estudio de Buhr (1980) si bien no se nos dan los valores concretos de cada una de estas vocales, estos pueden deducirse de los gráficos en los que se presenta la evolución de los tres primeros formantes de las vocales extraídas en un bebé que adquiriría la lengua inglesa entre los 0;4 y los 1;4 años. El tiempo en el que se midieron los formantes de las distintas vocales puede coincidir o diferir en algunas semanas debido a que no todas las vocales estuvieron presentes en el mismo momento temporal. No obstante, los valores presentados son los deducidos a lo largo del primer año de vida. Aquí están los valores aproximados para cada una de las vocales compartidas con el español (tablas 8.13-8.17):

**Tabla 8.13. Promedios de los valores formánticos de la vocal [a] de un niño (Buhr, 1980).**

Valores de [a] según Buhr	F1	F2
0;4 años y una semana	1000	2400
0;7 años y una semana	1200	2500
0;9 años	1200	2100
1;0 año	1300	2400

**Tabla 8.14. Promedios de los valores formánticos de la vocal [e] de un niño (Buhr, 1980).**

Valores de [e] según Buhr	F1	F2
0;4 años y una semana	900	3000
0;6 años	500	2500
0;9 años y una semana	400	2700
1;0 año menos una semana	700	2900

**Tabla 8.15. Promedios de los valores formánticos de la vocal [i] de un niño (Buhr, 1980).**

Valores de [i] según Buhr	F1	F2
0;4 años y dos semanas	400	4000
0;6 años	800	3800
0;9 y una semana	500	4100
1;0 año menos una semana	700	4300

**Tabla 8.16. Promedios de los valores formánticos de la vocal [o] de un niño (Buhr, 1980).**

Valores de [o] según Buhr	F1	F2
0;9 años y una semana	1200	2200
0;11 años menos una semana	900	2300
1;0 año	900	1900
1;0 año y tres semanas	1000	2100

**Tabla 8.17. Promedios de los valores formánticos de la vocal [u] de un niño (Buhr, 1980).**

Valores de [u] según Buhr	F1	F2
0;6 años	500	1200
0;6 años y dos semanas	600	1600
1;1 años y dos semanas	900	1600
1;2 años y una semana	800	2000

Como podemos observar, todos los valores formánticos calculados en las vocales de Buhr superan con creces a los hallados para los dos primeros formantes de un adulto femenino o masculino. Por otro lado, si volvemos a las tablas de los apartados 7.1.1 (en el caso de la niña) y 7.1.2 (en el caso del niño) en las que se presentaba el promedio mensual de los dos primeros formantes para cada una de las vocales a lo largo de los tres primeros años de vida, encontraremos rápidamente las semejanzas con los valores dados aquí por Buhr (1980).

No obstante, llama la atención frente a esta similitud de valores la diferencia de hertzios hallada cuando comparamos en ambos estudios el segundo formante de la vocal [i]. Encontramos así una desigualdad de 920 Hz entre el valor más alto del segundo formante de cualquier promedio mensual de la [i] en el primer año de vida de la niña (3380 Hz) y el segundo formante más elevado para esta vocal presentado por Buhr (4300 Hz). La disimilitud entre el valor más alto hallado en la vocal [i] de nuestro infante niño (3316 Hz) y el presentado por Buhr (4300 Hz) es de 984 Hz. Esta disconformidad puede ser debida a factores tan distintos como al diferente número de vocales [i] extraídas en cada bebé, o a la pronunciación más anterior de la vocal [i] por parte de los bebés que están adquiriendo el inglés americano (puesto que una diferencia similar aparece cuando comparamos los resultados de Buhr [1980] con los del estudio de Casal et al. [2002]). Para el resto de las vocales las medias que presentan los gráficos de Buhr (1980) son muy similares a las halladas por nosotros superando en algunos meses las nuestras a las de Buhr y viceversa.

Los siguientes estudios, el de Eguchi y Hirsh (1969) y el de Kuhl y Meltzoff (1996), nos permitieron comparar los valores de los dos primeros formantes tan solo para las vocales de los extremos del español, esto es para [i, a, u]. Si comenzamos con el primer trabajo encontramos que en la investigación de sujetos que adquirían inglés americano entre las edades de 3;0 y 13;0 años, Eguchi y Hirsh (1969) hallaron los siguientes valores a la edad de 3;0 años para dichas vocales (tabla 8.18):

**Tabla 8.18. Promedios de los valores formánticos de las vocales de los extremos (Eguchi y Hirsh, 1969).**

Niños de 3;0 años (Eguchi y Hirsh, 1969)		
Vocales	F1	F2
i	484	3318
a	986	1982
u	568	1664

Al igual que sucedía con el estudio de Buhr (1980), los valores de los dos primeros formantes que dan Eguchi y Hirsh (1969) son además de superiores a los de los adultos muy similares a los nuestros habiendo una mayor diferencia en el segundo formante de la vocal [u] (cuya media es de 1664 Hz en Eguchi y Hirsh [1969] frente a la hallada por nosotros que fue de 1251 Hz en la niña y de 1060 Hz en el niño). Entre las causas pueden aducirse, como ya hicimos para la disimilitud encontrada entre los valores del segundo formante de la vocal anterior alta de Buhr (1980) y la nuestra, el diferente número de muestras extraídas en cada bebé así como la distinta articulación de la misma vocal por hablantes de distintas lenguas.

Por otro lado, Kuhl y Meltzoff (1996) después de analizar las vocales articuladas por bebés de entre 0;3 y 0;5 años tras escuchar las palabras *hop*, *heap* y *hoop* emitidas por un hablante femenino a través de una máquina, registraron los siguientes valores (tabla 8.19):

**Tabla 8.19. Promedios de los valores formánticos de las vocales de los extremos (Kuhl y Meltzoff, 1996).**

Valores para las vocales cardinales dados por Kuhl y Meltzoff (1996)			
Edad	Vocal	F1	F2
0;3	i	781	3121
	a	933	2606
	u	731	2198



0;4	i	739	2887
	a	1044	2499
	u	757	2156
0;5	i	778	2947
	a	945	2393
	u	675	2335

Una vez más y como sucedía en los estudios anteriores, los valores para las vocales de los extremos dados por Kuhl y Meltzoff (1996) fueron superiores a los de los adultos y muy similares a los hallados por nosotros, aunque como ocurría en el trabajo de Eguchi y Hirsh (1969) la vocal [u] presentó un segundo formante mucho más alto en el estudio inglés que en el español. Entre las razones pueden barajarse de nuevo las de un distinto número de señales analizadas en cada bebé así como la diferente articulación de esta vocal por parte de hablantes de distintas lenguas.

Este último argumento, el que defiende la existencia de diferencias en los valores formánticos de vocales comunes a dos lenguas debido a que dichas vocales no comparten una ubicación exacta en la carta formántica de estas lenguas al estar influidas por el ambiente lingüístico, y el cual hemos venido esgrimiendo desde que hallamos la diferencia formántica entre los valores obtenidos en nuestro estudio y los de Buhr (1980), Eguchi y Hirsh (1969) y Kuhl y Meltzoff (1996), ha sido expuesto en algunos trabajos de análisis acústico.

Así pues en la investigación de Lyakso et al. (2006) tras comparar el habla de niños que adquirirían la lengua rusa con el habla de otros que adquirirían húngaro y holandés a la edad de 2;0 años, se comprobó que las vocales del holandés variaban más y se situaban más periféricamente en el espacio acústico que las vocales de los niños húngaros y rusos. Lo mismo sucedió en el estudio de Chung et al. (2008) tras equiparar las vocales compartidas por el inglés, el coreano, el griego y el cantonés. También en Rvachew et al. (2006) encontramos ejemplo de ello al citarse en este los trabajos de Escudero y Polka (2003), LaCharite y Paradis, (1997) y Martin (2002) según los cuales las vocales del francés canadiense y del inglés de Canadá están caracterizadas por diferencias acústicas y fonéticas incluso cuando hay un solapamiento fonológico (estos trabajos hallaron que la [i] de los niños que adquirirían francés canadiense era más difusa en relación a la [i] de los que adquirirían inglés, que la [u] de los niños hablantes del francés de

Canadá era más grave en relación a la [u] de los hablantes de inglés, y que la [a] de los sujetos que adquirirían el francés de Canadá era ligeramente menos compacta que la [a] de los que adquirirían inglés). Estas diferencias formánticas en las vocales compartidas por lenguas distintas pueden explicar las diferencias en los valores formánticos registrados entre las emisiones de nuestros niños y las de los sujetos que adquirirían otras lenguas maternas.

Finalmente, y a pesar de ser este último análisis (el que calcula los promedios formánticos de las señales agrupadas por timbre vocálico) el análisis que aporta datos más concretos y reales de las medias que presentan las distintas vocales a lo largo del tiempo así como de su evolución, muchos de los estudios que analizan acústicamente el habla infantil no hacen diferenciación de timbre y realizan los cálculos estadísticos sobre el total de señales identificadas en un mes y en un año. Esta circunstancia nos ha llevado a incluir en nuestra tesis un análisis del promedio formántico de todas las señales comprendidas en cada uno de los meses y años sin que estas aparezcan agrupadas por timbre para poder establecer la comparativa entre nuestro trabajo y otros estudios anteriores y poder observar así qué similitudes y diferencias ocurren.

En este otro parangón comenzamos con el trabajo de Kent y Murray (1982) en el que se analizaron los dos primeros formantes de los sonidos vocálicos aislados emitidos por veintidós niños que adquirirían el inglés americano a los 0;3, 0;6 y 0;9 años de edad. Las tablas siguientes muestran los resultados y en ellas aparecen los dos valores en torno a los cuales se movieron las dos primeras resonancias vocálicas tanto en el estudio de Kent y Murray (1982) (tabla 8.20) como en el nuestro (tabla 8.21 y 8.22):

**Tabla 8.20. Valores formánticos mínimos y máximos de las vocales de un niño (Kent y Murray, 1982).**

Valores Kent y Murray (1982)		
Edad	F1	F2
0;3	500-1500	1800-3800
0;6	500-1700	1600-3800
0;9	500-1800	1400-4100

**Tabla 8.21. Valores formánticos mínimos y máximos de las vocales de la niña de nuestro estudio.**

Valores niña		
Edad	F1	F2
0;3		
0;6	600-1290	1680-3190
0;9	400-1250	1780-3590

**Tabla 8.22. Valores formánticos mínimos y máximos de las vocales del niño de nuestro estudio.**

Valores niño		
Edad	F1	F2
0;3		
0;6	850-1230	2400-3260
0;9	400-1390	2500-3160

A pesar de la ausencia de valores en el tercer mes de vida puesto que nuestro estudio comenzó en el cuarto, un parangón pudo establecerse entre los valores de los otros dos meses. Si nos fijamos en el F1, observamos la similitud de valores mínimos en los bebés de las dos lenguas, aunque en el valor máximo los datos ofrecidos por Kent y Murray (1982) son mayores que los hallados en nuestro trabajo. Algo distinto ocurrió con el F2. En este formante los valores mínimos encontrados en nuestro estudio superaron con creces el valor más bajo ofrecido por Kent y Murray (1982) mientras que respecto al valor más alto del segundo formante, fue la media ofrecida por estos últimos autores la que superó en gran medida el promedio calculado sobre los sonidos vocálicos de nuestros niños. Antes de intentar dar una explicación veamos si estas diferencias se repiten en otros trabajos.

Un año más tarde, Gilbert et al. (1997) analizaron las vocales de cuatro niños que adquirirían inglés americano a los 1;3, 1;6, 1;9, 2;0 y 3;0 años de edad. Los autores extrajeron los sonidos vocálicos que podían aparecer aislados o acompañados de consonantes y tras medir los dos primeros formantes en cada uno de los niños calcularon la media de grupo para cada una de las edades. Los resultados fueron los siguientes (tabla 8.23):

**Tabla 8.23. Promedios de los valores formánticos de las vocales de cuatro niños (Gilbert et al., 1997).**

Valores medios hallados en el estudio de Gilbert et al. (1997)								
	Niño 1		Niño 2		Niño 3		Niño 4	
Edad	F1	F2	F1	F2	F1	F2	F1	F2
1;3	1029	2733	910	2675	1090	2584	1321	2201
1;6	1083	2652	879	2334	1268	2475	1141	2505
1;9	1116	2692	1170	2685	1081	2751	896	2336
2;0	1115	2747	1129	2622	1043	2785	685	1993
3;0	630	1899	900	2134	560	1858	528	1897

Los valores obtenidos por nosotros tras el cálculo del promedio de los dos primeros formantes de todos los sonidos comprendidos en los periodos temporales observados por Gilbert et al. (1997) son los que siguen (tabla 8.24):

**Tabla 8.24. Promedios de los valores formánticos de las vocales de los sujetos de nuestro estudio.**

Valores medios de los sujetos de nuestro estudio				
	Niña		Niño	
Edad	F1	F2	F1	F2
1;3	880	2960		
1;6	947	2159	870	2800
1;9	812	2310	800	1445
2;0	817	2937	605	2198
3;0	703	2547		

A partir de las tablas anteriores dos circunstancias pueden ser señaladas. Por un lado, las semejanzas y diferencias entre los valores obtenidos en las emisiones de los niños estudiados por Gilbert et al. (1997) y los nuestros para alguno de los formantes (a veces para los dos) y en determinados periodos temporales. Entre las similitudes, encontramos una semejanza entre la niña de nuestro estudio y el primero de los niños de Gilbert et al. (1997) puesto que los valores medios de los dos primeros formantes fueron similares. En las diferencias, una disimilitud considerable es la que hallamos si comparamos el F2 del niño de nuestro estudio a los 1;9 años,

que presentó un valor muy bajo, con el F2 de cualquiera de los niños observados por Gilbert et al. (1997) cuyos valores fueron mucho más altos.

Por otro lado, llaman la atención las diferencias en ocasiones importantes en los valores del F1 y del F2 de los niños de la misma edad que adquieren la misma lengua. Si volvemos a la primera tabla (tabla 8.23) comprobamos que la diferencia en el primer formante entre el niño segundo y el niño cuarto a la edad de 1;3 años fue de 411 Hz o que la desigualdad en el mismo formante y en los mismos niños a la edad de 2;0 años fue de 444 Hz (estas diferencias también fueron halladas en el segundo formante y a esta se añade la que aparece a los 1;3 años entre el niño primero y el cuarto cuyos segundos formantes distan 532 Hz). Disimilitudes notables fueron también las encontradas en los valores de los dos primeros formantes de los niños de nuestro estudio. Así a los 1;6 años de vida y frente a la escasa diferencia encontrada en el F1 de ambos bebés, el F2 presentó una distancia de 641 Hz. A los 1;9 años la diferencia en el segundo formante se incrementó alcanzando los 865 Hz.

Finalmente, presentamos los valores formánticos ofrecidos por Robb et al. (1997) tras analizar los dos primeros formantes en los sonidos vocálicos que ocurrieron dentro de una exhalación acústica, de veinte niños que adquirirían la lengua inglesa a las edades de 0;4, 0;8, 1;3, 1;6, y 2;1 años. La tabla inferior resume los resultados (tabla 8.25):

**Tabla 8.25. Promedios de los valores formánticos de las vocales de veinte niños (Robb et al., 1997).**

Valores medios hallados por Robb et al. (1997)		
Edad	F1	F2
0;4	768	2407
0;8	775	2494
1;3	815	2367
1;6	881	2417
2;1	821	2347

En la tabla siguiente aparecen los resultados del promedio del F1 y F2 de los sujetos de nuestro estudio para las edades indicadas. No obstante, debemos señalar que a los 1;3 años tan solo obtuvimos datos en la niña y que de manera contraria a los 2;1 años únicamente pudimos analizar los dos formantes del niño (tabla 8.26):

**Tabla 8.26. Promedios de los valores formánticos de las vocales de los dos sujetos de nuestro estudio.**

Valores medios de los sujetos de nuestro estudio				
	Niña		Niño	
Edad	F1	F2	F1	F2
0;4 (0;4-0;5)	865	2303	906	2992
0;8	927	2523	747	2544
1;3	880	2960		
1;6	947	2159	870	2800
2;1			500	1010

Como ocurría con otros estudios, podemos establecer semejanzas y diferencias entre los valores ofrecidos por Robb et al. (1997) y los nuestros. Así pues a los 0;8 años los valores de ambos estudios y para ambos formantes fueron muy similares mientras que a los 1;3 años nuestros resultados para el segundo formante superaron con creces a los de Robb et al. (1997), produciéndose la situación inversa a los 2;1 años, periodo en el que los valores del primer y del segundo formante fueron mucho más elevados en el estudio de Robb et al. (1997) que en el nuestro.

Una vez mostrados los resultados de los diferentes estudios, nos detenemos a explicar las dos circunstancias que se han repetido a lo largo de todos ellos. Por un lado el hecho de que existan diferencias en los valores formánticos de niños que poseen la misma edad pero que están rodeados de lenguas ambiente diferentes puede achacarse según el estudio de Boysson-Bardies (1989) a la diferente lengua materna que están adquiriendo los infantes. Esta fue la conclusión a la que llegaron los autores después de medir los dos primeros formantes en niños que adquirirían cuatro lenguas diferentes (inglés, francés, argelino y cantonés) y observar de este modo que la media del F1 de todas las vocales producidas por un niño francés fue de 758 Hz frente a los 1173 Hz que presentó un niño hablante de cantonés, o que la media del F2 fue de 2061 Hz para un niño de Hong Kong mientras que la de un niño hablante de inglés fue de 2805 Hz. Según los autores estos resultados pueden reflejar la preferencia de los niños de una lengua concreta por una vocal determinada y ello puede elevar o reducir los dos primeros formantes causando diferencias en la media mensual de los valores formánticos de los niños procedentes de distintas lenguas (como ya vimos en el apartado de la influencia de la lengua ambiente los niños hablantes de cantonés prefirieron las vocales con un elevado F1 pero con un bajo F2, repitiéndose esta tendencia también en el habla adulta, mientras que los niños argelinos

presentaron valores intermedios de F1 y valores bajos de F2, reflejándose esta tendencia también en el habla adulta). Este razonamiento explicaría la diferencia hallada entre nuestros valores y los que dan Kent y Murray (1982), Gilbert et al. (1997) y Robb et al. (1997), al tratarse el inglés y el español de dos lenguas diferentes.

Una segunda circunstancia, que se desprende de los trabajos de Gilbert et al. (1997), de Boysson-Bardies (1989) y de nuestro estudio, trata de las desigualdades en los valores de los formantes de niños que poseen la misma edad y que están aprendiendo además la misma lengua. En el estudio de Boysson-Bardies (1989) podemos observar, a partir de la tabla que resume los valores formánticos de todos los niños, que dentro del conjunto de niños que adquirirían el argelino, una de las niñas presentó un segundo formante de 2599 Hz mientras que el de otra fue tan solo de 2170 Hz (diferencia de 429 Hz). Lo mismo sucedía en el grupo de niños que adquiriría cantonés donde el F2 de una de ellos alcanzó los 2566 Hz mientras que el de la otra fue de 2061 Hz (diferencia de 505 Hz). En nuestro estudio estas diferencias se observan en las tablas presentadas anteriormente (8.21, 8.22, 8.24 y 8.26) y entre las que destacó la disimilitud de 739 Hz entre el F2 de la niña y el del niño en el segundo año de vida (siendo superior el de la niña). También en el trabajo de Rvachew et al. (1996), estudio comentado en el apartado 2.3.3 de revisión de estudios acústicos y en el que los autores describieron las vocales producidas por nueve niños que adquirirían inglés entre los 0;6 y los 1;6 años de edad, se encontraron diferencias en los valores formánticos con respecto a otros estudios de análisis acústico que analizaron las emisiones de sujetos que adquirirían inglés a las mismas edades. Un ejemplo residió en el promedio del F1 hallado por Rvachew et al. (1996) que fue similar al obtenido en los estudios de Kent y Murray (1982) mientras que el del F2 fue mucho más reducido (unos 2400 Hz menor).

En todos estos trabajos los niños no solo compartieron la edad sino la lengua a aprender por lo que la desigualdad no puede ser atribuida al diferente ambiente lingüístico sino a la preferencia individual por una determinada vocal. De este modo, la reducida media hallada en los promedios mensuales del tercer año de vida de nuestro sujeto varón así como el promedio anual final, podrían ser explicados si consideramos una inclinación por parte del niño en este tercer año a la articulación de vocales posteriores caracterizadas por poseer un segundo formante bajo. Los gráficos que simulan las cartas de formantes del apartado 6.1 y los resultados del test de percepción de apartado 6.2, corroboran esta predicción puesto que en los primeros pudo observarse un mayor agrupamiento de señales en el ángulo posterior de las cartas de formantes a lo largo del tercer año de vida, siendo también mayor la cantidad de vocales posteriores semialtas identificadas por los oyentes en el último año del estudio longitudinal. La propensión del sujeto varón a la emisión de vocales posteriores (no solamente orales sino

nasales) fue por tanto la responsable del descenso de la media mensual y anual de este tercer año.

No obstante, y a pesar de estas pequeñas discrepancias entre los valores formánticos obtenidos en las vocales de otras lenguas y el idioma español, el hecho es que tal y como reza esta primera hipótesis corroborada, los valores formánticos de los niños son mayores que los de los adultos independientemente de la lengua que hable el sujeto y de la preferencia que este tenga por unas u otras vocales, y ello es debido fundamentalmente a la configuración anatómica que caracteriza el tracto vocal del bebé y niño pequeño y no el del adulto (Buhr [1980], Clement y Wijnen [1994], Kuhl y Meltzoff [1996], Palethorpe et al. [1996] y Gilbert et al. [1997]).

En referencia a la segunda hipótesis de partida, esta será revisada junto a la tercera debido a que gran parte de las explicaciones relacionadas con la configuración y el desarrollo anatómico, con la percepción y con la influencia de la lengua ambiente, de las que echaremos mano para corroborarla, son compartidas por ambos supuestos.

La primera de estas dos conjeturas, según la cual las vocales que articulan en primer lugar los bebés y las más frecuentes al menos en el primer año de vida son las vocales centrales bajas y medias y las vocales anteriores, queda ratificada también en la lengua española según los resultados obtenidos en esta tesis. Así lo corroboran las cartas de formantes presentadas en los apartados 6.1.1.1 y 6.1.2.1 que ofrecían los resultados del análisis acústico de las emisiones de ambos niños en el primer año de vida y en las que podíamos observar como las señales analizadas se agrupaban en su mayoría en la parte central e izquierda del gráfico en posiciones medias y bajas.

De la misma forma, las vocales identificadas un mayor número de veces en el primer año de vida en ambos sujetos y por cuatro de los oyentes, tanto en los sonidos modales como en los nasales (apartados 6.2.1.1 y 6.2.2.1), fueron las vocales centrales y anteriores. Estos fueron los porcentajes obtenidos: en los sonidos modales analizados en el primer año de vida de la niña la vocal anterior semialta [e] supuso el 44,7% de la muestra y la vocal central baja [a] el 23,5%; en los nasales, esta última vocal alcanzó el 51,7% del conjunto y la vocal anterior semialta [ẽ] el 31% (en ambos tipos de señales el porcentaje que quedó para el total de los sonidos posteriores fue muy inferior). Por otro lado, en el primer año de vida analizado en el sujeto varón y en el caso de las señales modales, los porcentajes fueron del 60% para la vocal anterior semialta y del 8,8% para la central baja; en los sonidos nasales, la vocal anterior semialta representó el 65,3% y la central baja el 20,4 % de la totalidad de vocales identificadas en ese año (como ocurría en la niña, la suma de los porcentajes de estos tipos vocálicos superó con creces el porcentaje restante en el que quedaron representadas las vocales posteriores). A partir del primer año de vida y



hasta el tercero, el número de vocales posteriores aumentó en ambos niños hallándose la misma circunstancia en los estudios de Irwin (1948), que registró el habla de niños que adquirían inglés desde el nacimiento hasta los 2;6 años y encontró un aumento de las vocales posteriores entre los 0;11 y los 2;6 años, y en el de Landberg y Lundberg (1989) quienes observaron un aumento de las vocales posteriores a partir de 1;0 año en niños que adquirían la lengua sueca y que fueron grabados entre los 0;1 y los 1;5 años. Es la configuración particular del tracto vocal del recién nacido la que favorece según Ménard et al. (2004) la producción de aquellas vocales percibidas como centrales bajas y anteriores tal y como reflejan los resultados obtenidos tras los análisis acústico y perceptivo de las señales de ambos niños.

Estos mismos resultados, tanto los obtenidos tras el análisis acústico como los derivados del test de percepción, sirven para confirmar la tercera hipótesis postulada según la cual el bebé puede articular determinadas vocales en los meses inmediatos al nacimiento mientras que otras como la posterior alta [u] no aparecen hasta pasado un tiempo. Si volvemos al capítulo 6, observaremos la ausencia casi total de sonidos situados en una posición elevada del ángulo derecho o posterior de las cartas de formantes en el primer año de vida de los dos niños. Del mismo modo, en las transcripciones realizadas tras la audición de las señales del primer año de vida de ambos niños llama la atención la escasez de vocales altas posteriores identificadas por los oyentes: en el caso de la niña la [u] únicamente fue reconocida en los sonidos modales y en tan solo dos ocasiones (entre los 0;6-0;7 y entre los 0;7 y 0;8 años); en el del niño, esta apareció una sola vez en las señales modales (entre los 0;7 y los 0;8 años) ausentándose en las nasales.

Como habrá podido observar el lector, además de la escasez de este timbre vocálico a lo largo del primer año de vida destaca su ausencia antes de un momento temporal determinado: el primer semestre de vida. A pesar de que nuestro análisis comenzó a los 0;4 años y no a los 0;1, la [u] no apareció en las emisiones de ninguno de los niños antes de los 0;6 años mientras que entre los 0;5 y 0;6 años ya estaban presentes el resto de vocales del español tanto en las vocalizaciones de la niña como en las del niño. A partir del primer año de vida, y si bien su número siguió siendo muy reducido, la presencia de esta vocal aumenta ligeramente en las emisiones de ambos niños.

Factores de tipo anatómico y relacionados con el control motor de los órganos articulatorios pueden explicar estos hallazgos dando cuenta, al mismo tiempo, de la segunda y de la tercera hipótesis propuestas según las cuales las vocales más frecuentes en las emisiones de los bebés son las anteriores y centrales medias y bajas por un lado, y por otro el bebé puede articular determinados sonidos nada más nacer, como ocurre con las vocales que acabamos de mencionar, pero no otros como sucede con la vocal posterior alta [u], debido a

condicionamientos de esta índole. A continuación ofrecemos una revisión de los trabajos en los que se explica con detalle el desarrollo anatómico y el control motor responsables de estos resultados.

Si comenzamos con las limitaciones anatómicas y volvemos a lo explicado en el apartado 2.1.1, recordaremos que la configuración fisiológica que presentaba el bebé inmediatamente después a su nacimiento permitía únicamente la producción de algunos tipos vocálicos. Entre estas características encontramos según Lieberman et al. (1972), Fletcher (1973), Bosma (1975), Sasaki et al. (1977), Goldstein (1979), Kent (1981), Kent y Murray (1982) y Kent (1992) que el tracto vocal del bebé formaba un ángulo gradual en vez de uno orientado hacia la derecha creándose un solo tubo resonador frente a los dos del adulto que según MacNeilage (1980) y Carré et al. (1995) aumentaba la capacidad para producir sonidos diferentes. Podemos hablar por tanto de que el tracto vocal del bebé está constituido por un solo tubo resonador que no presenta la curva a la derecha ni las discontinuidades que posee el tracto vocal adulto necesarias para la producción de [i] y de [u] (en el caso de la vocal anterior alta [i] la división del tubo resonador en dos partes, una recta y otra orientada hacia la derecha, que tiene lugar a partir del cuarto mes de vida, así como el mayor control articulatorio de los órganos que intervienen en su articulación, permiten como veremos más adelante su realización por parte del niño).

Esta configuración lleva a hablar a algunos especialistas en este tipo de estudios, como De Boysson-Bardies et al. (1986), de una evolución desde un espacio vocálico centralizado en los meses inmediatos al nacimiento hasta un espacio vocálico más extendido al final del primer año (nótese que nuestro estudio comienza a los 0;4 años con lo que ya ha comenzado a producirse la reestructuración anatómica y el grado de centralización de los sonidos en la carta formántica es menor si bien puede apreciarse en los primeros meses analizados en la niña).

En cuanto a la lengua, según Clement (2004) esta es relativamente grande y ocupa casi por completo el espacio oral limitándose así sus movimientos dentro de esta cavidad. Según Kent y Murray (1982) la masa de la lengua está en la parte posterior de la cavidad oral y no en la anterior ocupando así el lugar de la faringe. Esto hace que la lengua del bebé sea más corta y más ancha en su parte posterior y que, por tanto, tenga menos capacidad de movimiento para llegar a los distintos puntos de la cavidad bucal donde se articulan las distintas vocales: zona alta anterior [i] y zona alta posterior [u].

Tras producirse la reestructuración entre los 0;4 y los 0;6 años, la capacidad del bebé de producir vocales diferentes aumenta en gran medida debido causas tales como: el descenso de la laringe que según Clement (2004) y Nishimura et al. (2003) contribuye a una mayor

independencia de los procesos de fonación y articulación de las vocalizaciones; además, el descenso de la laringe y el de la lengua facilitan la creación de la pared anterior de la faringe (Bosma [1975a, 1975b, 1976, 1985], Crelin [1973, 1976], Laitman y Crelin [1976], Lieberman [1977, 1984], Sasaki et al. [1977], Kent [1981], Fried et al. [1982], Westhorpe [1987] y Kent y Vorperian [1995]); la división del tubo resonador (pasando como ya hemos explicado de ser un solo tubo en una posición lineal a ser dos tubos al inclinarse la parte superior de este hacia la derecha) en la región nasofaríngea que permite la diferenciación en dos tubos resonadores y la articulación por tanto de señales diferenciadas (Lieberman [1984] y Carre, et al. [1995]); el crecimiento de la mandíbula hacia abajo otorga a la lengua de un mayor espacio para moverse posibilitándose según Clement (2004) la pronunciación de vocales diferentes. Diez años antes, Clement y Wijnen (1994) ya opinaban que antes de producirse esta reestructuración, el movimiento de la lengua en la dirección anterior-posterior estaba restringido debido a la falta de espacio y que esto resultaba en rangos restringidos del F2 en comparación con aquellos de los hablantes mayores. Esta posibilidad parece ser apoyada por los hallazgos de Bond et al. (1982) en opinión de los cuales según el niño se hace mayor, la parte dorsal de la lengua y la laringe descienden produciéndose un alargamiento de la faringe. El rango de movimiento de la lengua a lo largo del eje anterior-posterior aumentaría con esta nueva disposición incrementándose también la habilidad del niño de producir contrastes en el eje del F2. También Bloom (1998) cree que el hecho de que la lengua sea ahora relativamente pequeña dentro de la cavidad oral y el que goce de más espacio para moverse debido a esta reestructuración anatómica, contribuye a un mayor control de los articuladores y resulta en diferentes tipos de vocalización.

No obstante, no solo son las restricciones anatómicas las responsables de la prevalencia de determinadas vocales frente a otras y las que impiden que determinados sonidos vocálicos no aparezcan hasta un momento determinado sino que tal y como apunta Feld (1999:6) el nivel del control de los órganos articuladores delimita el tipo de sonido que se va a producir:

“[...] El orden de adquisición de los fonemas parece depender de su nivel de dificultad desde el punto de vista sensoriomotor: ciertos sonidos exigen la actuación de un mayor número de músculos, un control más preciso de los movimientos y una coordinación más afinada”.

A la misma conclusión llegaron los trabajos que han estudiado el desarrollo de las vocalizaciones con ayuda de trectos vocales artificiales y según los cuales la tardía articulación de determinados sonidos no radica en diferencias de tipo anatómico entre el niño y el adulto sino más bien en el diferente control motor de los órganos necesarios para su producción.

Este es el caso de los ya mencionados estudios de Nordström (1975, 1977) en los que tras reproducirse el tracto vocal de una mujer y de un niño a partir del tracto vocal de un hombre, se concluyó que las diferencias anatómicas solo explican parte de las diferencias formánticas entre hombres, mujeres y niños pudiéndose inferir que además de la anatomía existen diferencias funcionales entre estos grupos de edad. Otro ejemplo lo encontramos en el trabajo de Ménard y Boé (2004) que como comentamos en el capítulo del estado de la cuestión utilizó el *Variable Linear Articulatory Model* (VLAM) creado por Boë (1999) para reproducir siete variables que permitieran estudiar la producción de las vocales orales francesas [i y u e ø o ε œ]. Tras el estudio se concluyó que es la falta de control motor, y no la configuración anatómica que presenta el niño al nacer, la que determina qué sonidos puede producir el bebé y cuáles no (conclusión a la que también llegaron Mowrer [1980], Ménard [2003] y Ménard et al. [2004] en sus respectivos trabajos). Serkhane et al. (2007) partieron también del modelo VLAM de Boë (1999) para determinar el grado de control motor que el niño posee para producir los sonidos de su lengua a los 0;4 y 0;7 años de edad. Los resultados demostraron que a los 0;4 años las vocalizaciones de los corpus reales no cubrían todos los sonidos que se hubieran producido si el bebé hubiera utilizado el rango completo de configuraciones articulatorias que le ofrecía el VLAM.

Esta insuficiencia de control articulatorio radica en parte en la falta de madurez de las células nerviosas que controlan la actividad de los músculos y por tanto el habla. Tal y como exponía Smith (2010) al séptimo mes de vida aunque es alto el número de neuronas que ha alcanzado la madurez, los patrones de la conexión entre células necesarios para adquirir las habilidades motoras no se han desarrollado aún a esta edad. Esta escasa actividad neuronal es aún más patente según Ploog (1979) en la etapa del llanto y de sonidos reflejos pero poco a poco las conexiones van aumentando en la etapa del susurro entre los 0;1 y 0;3 años y se intensifican como veíamos entre los 0;6 y los 0;9 años de vida. Con el balbuceo según Whitaker (1976) el control neuronal que el niño posee sobre la musculatura del tracto vocal aumenta incrementándose también el tipo de vocales producidas.

Una vez contemplada una de las razones principales de la falta de control articulatorio por parte del niño que redundo en un limitado rango de vocalizaciones, es necesario subrayar que existen diferencias en dicho control según el articulador del que se trate. Según Smith (2010) hay diferencias en el desarrollo del control motor de la mandíbula, lengua, labios y velo de forma que el control de un articulador puede ser más parecido al del adulto mucho antes que el control de otro (esta es la conclusión a la que llega tras detectar diferencias pronunciadas en la inervación anatómica, biomecánica y neural de varios articuladores). La misma conclusión puede ser extraída del trabajo de Green et al. (2000) quienes tras analizar la coordinación entre

la mandíbula y los labios en las producciones de sílabas con consonantes bilabiales en niños de 1;0, 2;0, y 6;0 años y en adultos jóvenes, observaron que en los niños de 1;0 año la mandíbula era el principal articulador frente los labios superior e inferior. Entre los 2;0 y los 6;0 años, la contribución del labio inferior aumentó significativamente (también lo hizo el del labio superior aunque no de manera significativa). Dos años más tardes y después de grabar el movimiento del labio superior, inferior y de la mandíbula en niños de 1;0, 2;0, 6;0 y 7;0 años y en adultos de 35 a 37 años en la producción de las palabras de *baba*, *papa* y *mama*, Green et al. (2002) demostraron que el control de la mandíbula de los infantes fue más parecido al del adulto de lo que lo fue el de los labios (este dominio mandibular es el responsable de que las vocales más frecuentemente emitidas por los bebés sean vocales centrales bajas y anteriores).

Como comentábamos en el estado de la cuestión, el mayor control sobre la mandíbula puede deberse a múltiples razones tales como: a que el feto de ocho semanas ya es capaz de abrir la mandíbula según Humphrey (1964) mientras que la musculatura del labio está aún desarrollándose a esta edad (Gasser [1967]); a que los labios tienen grados superiores de libertad de movimiento en comparación con la mandíbula según Smith (2010); a que los movimientos de la mandíbula para el habla tienen lugar en la dimensión vertical y no ocupan un gran porcentaje del trabajo potencial de la mandíbula (Ostry et al. [1997]); o porque según Honda et al. (1995), Honda (1996) y Gerard et al. (2003) las contracciones de los músculos que producen los movimientos de los labios (y los de la lengua) son mucho más complejas que las necesarias para el movimiento mandibular. La simplicidad del movimiento mandibular pudo hacer según el estudio de Green et al. (2000) que el movimiento de otros articuladores como el del labio inferior se asimilara al patrón establecido por la mandíbula habiendo después de esta fase de integración una de diferenciación y otra de refinamiento hacia el estado adulto.

La relativa facilidad de movimiento mandibular y la mayor estabilidad que esta presentaba debido a su gran tamaño llevaron a McAllilster (2009) a hablar de un movimiento similar al de Green et al. (2000) al considerar el movimiento en bloque de los órganos articulatorios por parte del niño el cual emparejaría el movimiento de la lengua y de los labios al de la mandíbula para formar una unidad siendo esta última el apoyo fundamental a la hora de realizar cualquier movimiento articulatorio. El apoyo mandibular y la integración de gestos o el movimiento en bloque permitirán la articulación de determinadas vocales, como las anteriores o centrales (vocales que en opinión de Lindblom y Sundberg [1969] pueden ser producidas únicamente por la manipulación de la posición de la mandíbula siendo el desarrollo de esta así como el de la parte frontal de la lengua mucho más rápido que el de la parte de atrás de la misma necesario para la producción de los sonidos posteriores) , y dificultarán la articulación de otras como es el caso de las posteriores.

Todos estos condicionantes anatómicos y de control motor de los distintos articuladores que caracterizan al bebé en los primeros meses de vida son los responsables de que sean los sonidos vocálicos centrales y anteriores bajos y medios (como [a] y [e]) los más producidos por los niños en la infancia temprana quedando las vocales posteriores, para las que el control de un mayor número de estructuras es necesario, relegadas a un segundo plano en cuanto a frecuencia se refiere. Dichas limitaciones en el control motor de determinados órganos articulatorios pero no en el de otros pueden también estar detrás del hecho de que la vocal anterior alta [i] esté presente desde el primer mes analizado en ambos niños mientras que la vocal posterior alta [u] aparece más tarde.

No obstante y dentro del ángulo posterior, la vocal alta redondeada [u] fue la que más tardó en aparecer y la que menos veces se repitió a lo largo del periodo de estudio frente a la vocal [o] cuya presencia fue anterior y cuya frecuencia fue más elevada. A continuación ofrecemos las razones que aportan algunos estudiosos sobre el porqué de esta circunstancia.

En lo que se refiere a la [u], las pocas muestras que se hallaron de esta vocal no solo en los trabajos citados cuando postulábamos la tercera hipótesis de partida sino también en nuestros resultados (corroborándose por tanto dicha hipótesis), pueden explicarse tal y como indican algunos autores si tenemos en cuenta que la coordinación del movimiento de los labios, la lengua y la mandíbula necesaria para su producción no se ha desarrollado de una manera consistente en los primeros meses de vida. Buhr (1980) añade que además de esta falta de coordinación posnatal indispensable para la producción de dicha vocal, el bebé cuando nace está desprovisto del control neuromuscular necesario para la protusión de los labios como el que implica la articulación de este sonido (más concretamente quizás sea la falta de control de los músculos faciales que inervan el *Orbicularis Oris Superior* y que son los responsables de que los labios se plieguen, como veíamos en el punto 2.1.3 del estado de la cuestión, produciéndose el redondeamiento). Así, aunque en el instante de nacer, tal y como reivindican Ruch y Zimbardo (1971) (citados en Mowrer [1980]), el bebé es capaz de flexionar algunos músculos de los labios y es capaz de apretar el pezón de la madre para succionar la leche, puede que muchos de estos movimientos sean puramente reflejos.

Por otro lado, es importante también el grado de constricción producido por la lengua necesario para la articulación de dicha vocal y que resulta además de vital importancia en la teoría de Stevens y House (1955). Para estos autores, que miden las áreas creadas a partir del punto de constricción de la lengua, es necesario que la constricción esté bien definida en la cavidad oral y esto conlleva un desarrollo del sistema neuromuscular que inerve el movimiento de la lengua y de la mandíbula el cual no está en el nacimiento sino que se desarrolla en los

meses posteriores (esto hace que el bebé no pueda pronunciar sonidos al nacer con una constricción muy específica como la requerida para la [u], pero sí otros sin una constricción muy delimitada como la necesaria para la [a]). De la misma opinión son Ménard et al. (2004) quienes mostraron que la producción de esta vocal por los niños requiere un gesto articulatorio relativamente preciso (una constricción velo-palatal) y que ligeras variaciones en la mandíbula y en los parámetros del dorso de la lengua resultarían en una vocal que no sería percibida como [u]. Esta exactitud en la articulación, no solo de [u] sino también de [o], fue igualmente expuesta en los trabajos que se citan en el estudio de Boë et al. (1992), que presentamos en el capítulo segundo a propósito de la no unicidad del tracto vocal y de la articulación compensatoria, en los que el análisis vocálico a partir de modelos de trectos vocales artificiales reveló que la precisión del control de cada parámetro (localización de la constricción, área transversal de la constricción, área de la abertura del labio, etc.) que interviene en la articulación depende de la vocal que se considere, siendo imposible para la [o] y la [u] cualquier compensación articulatoria debido a que requieren un control muy preciso o exacto de la abertura del labio (frente a la compensación permitida para otras vocales como [i] como demostró la investigación de Oller y MacNeilage [1983]).

La frecuente presencia de vocales centrales y anteriores medias y bajas hallada en nuestro estudio así como en otras investigaciones sobre la adquisición de las vocales en distintas lenguas maternas, parece confirmar la hipótesis de la independencia de la que hablábamos en el punto 2.2.2 y según la cual la semejanza en las producciones hallada en niños de diferente ambiente lingüístico se debe a que son restricciones de tipo universal, como las anatómicas o las que tienen que ver con el control motor de los órganos articulatorios, las que rigen las emisiones de los bebés independientemente de la lengua que estos hablen. No obstante, los resultados obtenidos tanto en nuestro trabajo como en los que estudian el desarrollo vocálico en otras lenguas indican que la hipótesis de la interacción, según la cual las emisiones de los bebés están influenciadas por la lengua ambiente, también está presente. Veamos de qué manera se deja sentir la influencia de la lengua ambiente en las emisiones de nuestros sujetos.

En la presente investigación y dentro de la mayor frecuencia de vocales anteriores y centrales, es la vocal anterior semialta [e] la que aparece en un mayor número de ocasiones por encima incluso de la vocal central baja [a] en determinados periodos temporales. De la misma forma y al margen de ser la última vocal en aparecer debido a su dificultad articulatoria, la vocal posterior alta [u] fue la menos articulada por ambos niños en los distintos periodos que comprendió el estudio. ¿Reflejan estos resultados la mayor o menor frecuencia de estas vocales en la lengua meta? Veámoslo.

La presencia dominante de la [e] así como la menor articulación de la [u] fueron también halladas en el estudio de Quilis y Esgueva (1980)<sup>7</sup> tras computar el número de veces que cada una de las cinco vocales del español aparecía en un corpus basado en la lengua hablada y en el que se utilizaron ocho hombres y ocho mujeres pertenecientes a cuatro generaciones diferentes: de entre 18 y 19 años, de entre 33 y 35 años, de entre 52 y 54 años y de entre 62 y 65 años. Según el recuento, la vocal anterior semialta [e] superó a la central baja [a] tanto en hombres como en mujeres en las cuatro generaciones en las que se analizó. A continuación presentamos una tabla (tabla 8.27) con los porcentajes de aparición que los autores hallaron para las tres vocales de las que se ha hablado:

**Tabla 8.27. Frecuencia de los fonemas vocálicos en español (Quilis y Esgueva, 1980).**

		Hombres	Mujeres	Por generación	En las cuatro
Vocal	Generación	Porcentaje	Porcentaje	Porcentaje	Porcentaje
e	Primera	15,69	14,81	15,25	14,6725
	Segunda	14,49	14,16	14,325	
	Tercera	15,10	14,425	14,7625	
	Cuarta	14,79	13,915	14,3525	
a	Primera	11,265	11,765	11,515	12,1931
	Segunda	12,375	12,675	12,525	
	Tercera	11,765	11,88	11,8225	
	Cuarta	12,85	12,97	12,91	
o	Primera	10,065	10,905	10,485	9,985

<sup>7</sup> Nos decantamos por el estudio de Quilis y Esgueva a pesar de la presencia en la literatura de otros trabajos en los que se llevó a cabo un recuento de las vocales del español, por ser este el que más se ajustó a nuestro criterio de “español hablado conversacional”. Las otras investigaciones halladas, todas recogidas en Quilis y Esgueva (1980), partían de un recuento vocálico en corpus de habla distintos al nuestro. Algunos ejemplos son: el trabajo de Zipf y Rogers (1939), el de Navarro Tomás (1946), el de Alarcos Llorach (1961) o el de Lloyd y Schnitzer (1967), cuyos corpus se basaban en la lengua escrita; el de Delattre (1965) cuyo recuento se hacía a partir de un corpus en el que se mezclaban distintos tipos de lengua oral, a saber, el estilo narrativo con el coloquial; y otros como el de Guirao y Borzone (1972) en el que se mezclaba la lengua escrita con la lengua oral.



	Segunda	8,765	9,68	9,2225	
	Tercera	9,755	10,155	9,955	
	Cuarta	10,065	10,49	10,2775	
i	Primera	6,995	7,105	7,05	7,3812
	Segunda	7,73	7,445	7,5875	
	Tercera	6,97	7,49	7,23	
	Cuarta	7,195	8,12	7,6575	
u	Primera	3,245	3,505	3,375	3,3343
	Segunda	3,34	3,27	3,305	
	Tercera	3,53	3,525	3,5275	
	Cuarta	3,23	3,03	3,13	

Como puede desprenderse de la tabla, en el español hablado en situación conversacional (misma condición en la que fueron realizadas las grabaciones de nuestros sujetos) los porcentajes de la vocal anterior semialta [e] superaron a los de la vocal central [a] tanto en los hombres como en las mujeres por separado, en la combinación de las generaciones de los hombres y de las mujeres así como en la mezcla de las cuatro generaciones (no obstante, esta última ocupó el segundo puesto en dicho estudio circunstancia que se repitió también en determinados periodos del nuestro). Respecto a la [u], los porcentajes obtenidos tanto en hombres como en mujeres, los hallados en la combinación de las generaciones masculinas y femeninas y en los de la unión de todas las generaciones, están muy por debajo de los de la otra vocal que menos veces se repitió, y que es la vocal anterior alta [i], cuyo porcentaje final es de 7,3812 % frente a los 3,3343 % de la vocal [u].

Observamos por tanto que la mayor presencia de la [e] dentro de la mayoría de vocales anteriores y centrales en los primeros meses de vida, así como las pocas muestras de la [u] y el último lugar que esta ocupa dentro de las vocales posteriores, parecen estar influenciados además de por los condicionamientos anatómicos y de control motor apuntados anteriormente, por los patrones fonéticos de la lengua española evidenciándose pues que una conjunción de la hipótesis de la independencia y de la interaccional debe emplearse a la hora de explicar las producciones vocálicas de los infantes. Revisamos a continuación estudios de otras lenguas en

los que el influjo de la lengua meta se dejó sentir en las tempranas vocalizaciones de los bebés y niños pequeños.

Un ejemplo es el estudio de Oller y Eilers (1981) en el que tras grabar las emisiones de ocho niños que adquirirían español y ocho que adquirirían inglés a la edad de 1;0 año, se encontró que los niños hablantes de español producían un mayor número de las vocales que se dan en esta lengua ([a, e, i, o, u]) y una menor cantidad de las vocales características del inglés (entre ellas [æ, ɪ, ε, ʊ, ɔ, α, ʌ, ə]).

Otra muestra la encontramos en el trabajo de De Boysson-Bardies et al. (1989) en el que se analizaron las vocales extraídas del balbuceo canónico y variado (que siguiera el patrón CV) de veinte niños de 0;10 años hablantes de francés, de inglés, de cantonés y de árabe (cinco niños por comunidad). Los resultados mostraron que los niños ingleses producían más vocales anteriores, que los infantes franceses emitían una mayor cantidad de vocales centrales medias, que los hablantes de cantonés articulaban un mayor número de centrales bajas, y que los niños árabes producían más vocales posteriores medias hallándose las mismas tendencias en las respectivas lenguas adultas.

La influencia de la lengua meta se deja ver también en el estudio de Tse (1991) en el que tras registrar el comportamiento vocálico de niños que adquirirían cantonés, se observó que la vocal alta del extremo posterior no apareció de manera significativa en las primeras etapas del desarrollo. Este hecho podría explicarse por la influencia que la lengua ambiente ejerció en el repertorio vocálico del niño ya que, según Fok (1979), la [u] es la vocal menos frecuente del cantonés con lo que los niños tenían menos oportunidades de practicar su producción (lo mismo ocurriría en el caso del español ya que al ser como hemos visto una vocal poco frecuente, los niños la articularon en pocas ocasiones).

El mismo argumento se esgrime en el artículo de Otomo y Stoel-Gammon (1992) en el que se nos dice que más allá del factor de la precisión requerida para la producción de la [u] (lo que complica su articulación), su escasa emisión por parte del bebé se debe al limitado número de palabras que contienen esta vocal en el área del vocabulario del niño más pequeño (y no solo en el vocabulario del pequeño sino también en el del adulto como ocurre con el español).

En el estudio del balbuceo y de la primera palabra de un sujeto que aprendía italiano a los 0;10 años, Zmarich (1997) observó que la vocal baja [a] era una de las vocales más articuladas reflejando este hecho la alta presencia de dicha vocal en la lengua italiana.

Los investigadores Davis y Davis (2002) compararon las emisiones vocálicas de dos niños que adquirirían portugués de Brasil a la edad de entre 1;0 y 3;0 años con las producciones

de niños ingleses de la misma edad. Los resultados mostraron que, si bien todos los niños preferían las vocales centrales y anteriores, estos diferían en la altura de las mismas decantándose los niños que adquirían portugués por las vocales altas frente a los ingleses que mostraban una predilección por las medias y las bajas reflejando así los patrones de sus respectivos ambientes lingüísticos.

Rvachew et al. (2006) después de un estudio comparativo informaron de que los niños que adquirían inglés producían más vocales altas y posteriores que los niños que adquirían francés entre los 0;8 y los 1;6 años. Dos años más tarde los mismos autores llevaron a cabo un estudio longitudinal en el que se examinaron las producciones de veinticuatro niños que adquirían inglés canadiense y veintisiete que adquirían el francés de Canadá entre los 0;8 y los 1;6 años de vida (expuestos a una sola de las lenguas y cuyos padres eran hablantes nativos también de un solo idioma). El ANOVA realizado mostró que los niños ingleses emitían una mayor cantidad de vocales posteriores altas redondeadas en comparación con los niños de habla francesa reflejando la influencia de la lengua input en el niño a edades muy tempranas.

Las emisiones de niños de entre 0;8 y 2;1 años procedentes de cinco ambientes lingüísticos distintos (turco, francés, rumano, holandés y árabe tunecino) fueron comparadas en el trabajo de Kern y Davis (2009). Los resultados indicaron, como ya comentamos en el estado de la cuestión, que una combinación de patrones universales y específicos de la lengua ambiente tenía lugar en las emisiones de los sujetos. De esta forma, si bien predominó en todas las lenguas el cuadrante inferior izquierdo del espacio vocálico, diferencias específicas de la lengua meta fueron halladas: los niños hablantes de francés produjeron más vocales altas que los niños que adquirían árabe tunecino, y las vocales medias de los primeros junto con las vocales medias de los sujetos hablantes de turco fueron más altas que en el resto de comunidades lingüísticas. Finalmente, los niños rumanos presentaron la mayor frecuencia de vocales altas centrales en comparación con el resto de comunidades.

La combinación de la hipótesis de la independencia y de la interaccional en las emisiones tempranas de los bebés, hallada en nuestro estudio y en los trabajos de los párrafos anteriores, estuvo también presente en la investigación de Lee et al. (2009) en la que se compararon las vocales extraídas del balbuceo canónico (concretamente del patrón CV) de niños que adquirían coreano entre los 0;8 y los 1;0 años de edad, con la frecuencia de sonidos vocálicos presentes en el habla dirigida a ellos por sus madres (habla que reflejaba las características de lengua ambiente). Al comparar los resultados con los obtenidos en niños de la misma edad que adquirían inglés se halló que si bien en las emisiones de los infantes de ambos ambientes lingüísticos predominaban las vocales posteriores bajas y centrales altas, incluso

aunque las frecuencias de esas vocales eran diferentes en el habla dirigida al niño inglés y al coreano, los niños que adquirían inglés producían más vocales anteriores altas y bajas, y más vocales centrales medias que los niños que adquirían coreano quienes emitían más vocales centrales bajas.

Las restricciones universales y las específicas de la lengua ambiente también se recogen en el trabajo de Chen y Kent (2010) en el que se estudiaron las vocalizaciones espontáneas en la estructura CV de veinticuatro niños que adquirían el mandarín entre 0;7 y 1;0 año (grupo 1) y entre los 1;1 y 1;6 años (grupo 2). Los resultados mostraron que los niños producían de forma más frecuente las vocales anteriores medias semibajas y centrales medias [ɛ, ə] y que en sus emisiones además predominaban las vocales bajas y medias sobre las vocales altas, circunstancia similar a la hallada en estudios con otras lenguas como el inglés. No obstante, la producción de la vocal [a] en el 40% de las ocasiones, porcentaje no encontrado en otros estudios longitudinales, reflejó la influencia de la lengua input en las emisiones de los niños pertenecientes incluso al grupo de menor edad<sup>8</sup>.

Una vez corroboradas la segunda y la tercera hipótesis planteadas queda claro por tanto que el predominio de determinadas vocales sobre otras así como la aparición tardía de algunas de ellas se debe principalmente a las restricciones anatómicas y de control motor que irán abandonando al niño a lo largo de su desarrollo, así como a la influencia que los patrones de la lengua ambiente tienen en las emisiones de los infantes. Se descarta por tanto que la mayor o menor frecuencia de una vocal o la aparición tardía de algunas de ellas se deba a factores de tipo perceptivo como el de la falta de diferenciación entre los sonidos vocálicos de una lengua debido a la dificultad de los mismos o a su escasez en la lengua específica ya que tal y como veíamos en el apartado 2.2.1 del estado de la cuestión, el feto (Mattock et al. [2010] y Zimmer et al. [1993]) y el bebé (Trehub [1973] y Kuhl y Miller [1982], entre otros) son capaces de

---

<sup>8</sup> Remitimos al lector en este punto a los apartados 2.2.2 y 2.2.3 para recordar los mecanismos que según algunos autores intervienen o ayudan en el aprendizaje del niño de los patrones fonéticos de su lengua ambiente (además del perceptivo, claro) y que eran tales como el *vocal learning* o aprendizaje vocal según Kuhl y Meltzoff (1996) (herramienta también hallada en el experimento de Meltzoff y Moore [1989]), el *feedback* auditivo o sensaciones que el niño recibe de sus propias emisiones según De Boysson-Bardies *et al.* (1986) y Vihman (1996), así como el mecanismo de la imitación visual que lleva a cabo el bebé de las unidades fonéticas de su lengua meta (afirmado en el trabajo de Jung [2009]) y según el cual el niño puede utilizar la información visible para la adquisición de determinados rasgos como es el caso del rasgo [alto]/[bajo] atribuible a determinados sonidos del habla.

diferenciar las vocales de su lengua y en el caso de este último la distinción se extiende a vocales de lenguas a las que nunca ha estado expuesto (Trehub [1976]).

Por otro lado, la anatomía del sistema auditivo así como las habilidades perceptivas se desarrollan antes que las articulatorias, tal y como veíamos en el mencionado apartado, de forma que la percepción va un paso por delante de la producción al comienzo del desarrollo independientemente de las estrategias que utilice el pequeño para la distinción auditiva (teoría de las asimetrías direccionales de Polka y Bohn [2011]), de la parte del continuum en la que se fije, esto es, si en la porción estable o en las transiciones (Borzorne de Manrique [1977] vs Nittrouer y Miller [1997], Morrongiello et al. [1984], Nittrouer y Studdert-Kenney [1987] y Nittrouer [1992] y [2005]), de que la percepción se realice de manera categorial o continua (Pisoni [1975] vs González Álvarez [1997] y Swoboda et al. [1976]), de si esta se basa en las frecuencias absolutas de las vocales o en el intervalo formántico (Mattock [2010] vs Lloyd [1890a], [1891], [1892]), del hablante que realice la emisión (normalización lingüística de Mearan et al. [1992] y Walton et al. [1991]) o de que la habilidad universal de distinguir contrastes no nativos comience a desaparecer a partir de los 0;6 años según reportaba el experimento de Polka y Werker (1994).

En cuanto a la cuarta hipótesis de partida que postulaba que una superposición de sonidos vocálicos se producía en los primeros meses de vida en las emisiones del bebé, es esta la hipótesis que quedó corroborada en menor medida en los resultados obtenidos tras el análisis acústico de las señales de ambos sujetos sin que descartemos que dicha circunstancia pueda deberse al escaso número de señales extraídas en los primeros meses examinados.

Como podemos observar en las cartas de formantes del apartado 6.1.1.1, las vocales extraídas de las emisiones de la niña se agrupan en el centro y ligeramente en la parte anterior de la carta de formantes colocándose unas muy juntas a las otras y en algunos casos superponiéndose como ocurre entre los 0;4 y 0;5 años, y entre los 0;5 y los 0;6 años excepto para una vocal alta anterior, y en menor medida entre los 0;6 y 0;7 años, tiempo este en el que las señales comienzan a separarse moviéndose en torno a las categorías vocálicas y evidenciándose así una pequeña expansión del espacio vocálico que irá aumentando en los meses siguientes delimitándose poco a poco el triángulo vocálico del español.

Respecto al niño y aunque la superposición es quizás menor, las señales aparecen muy cercanas en la parte anterior del gráfico correspondiente a los 0;4 y 0;5 años, y a los 0;5 y 0;6 años de vida, sin que haya un agrupamiento en la zona central de la carta de formantes. Entre los 0;6 y 0;7 años, podemos observar una mayor diferenciación de las vocales que parecen haberse agrupado en torno a sus respectivas categorías produciéndose por tanto una expansión del

espacio vocálico. La ausencia de una superposición central de las vocales del niño puede ser debida en parte al escaso número de señales extraídas de las emisiones del sujeto en los primeros meses analizados ya que las primeras cartas de formantes del niño sí presentaron una leve superposición de los sonidos aunque esta fuera en la parte izquierda del gráfico (circunstancia que hace pensar que de haberse extraído más señales éstas podrían haber sido centrales y haber estado superpuestas).

Por otro lado, la superposición en ambos niños quizás hubiera sido más notoria en los meses anteriores a nuestro análisis, esto es entre los 0;1 y los 0;4 años, debido a que en esa etapa inicial las restricciones anatómicas son mayores puesto que aún no se ha producido la reestructuración anatómica, y el control motor de los bebés es más limitado. A partir de los 0;7 años, la superposición en ambos niños va dando lugar a una diferenciación vocálica. A continuación recuperamos parte de los trabajos expuestos en el estado de la cuestión en los que esta hipótesis quedó confirmada con más contundencia.

En el trabajo de Lieberman (1980) en el que se midieron el F1 y el F2 de distintas vocales aisladas (vocales del balbuceo e insertas en palabras con significado) en cuatro niños que adquirirían la lengua inglesa entre las edades de 0;4 y 5;0 años (recordemos que no se tomó el mismo periodo de estudio en cada uno de los niños) se observó que en el sujeto L. S. tenía lugar una clara superposición de las vocales [ɪ, ɛ, æ, ʊ] a los 0;4 años mientras que a la edad de 0;6 años parecía como si dichas vocales se estuvieran moviendo hacia posiciones acústicas que dibujaban un sutil triángulo vocálico (entre los 0;10 y los 1;3 años el espacio vocálico siguió expandiéndose en este sujeto). También el sujeto R. C. mostró una expansión del espacio vocálico entre los 0;9 y los 1;5 años de la misma manera que esa expansión fue hallada en el sujeto J. B. entre los 1;4 y los 3;0 años de edad y en el niño F. R. entre los 2;7 y los 3;6 años continuando esta expansión más allá del balbuceo según Lieberman (1980) y acercándose cada vez más al espacio acústico adulto.

Una superposición de todas las vocales excepto de la [ɪ], circunstancia también hallada en la carta de formantes de nuestro sujeto femenino entre los 0;5 y 0;6 años, tuvo lugar en las vocales analizadas por Buhr (1980) cuando el niño contaba con 0;4 años. A los 0;6 años surgió un rudimentario espacio vocálico delimitado por las vocales [u,æ,ɪ].

De las cartas de formantes realizadas en el trabajo de Kent y Murray (1982) en el que se analizaron las vocales de veintiún niños de 0;3, 0;6 y 0;9 años que adquirirían el inglés americano, se desprende también que una superposición vocálica tuvo lugar a los 0;3 años reduciéndose en parte a los 0;6 y a los 0;9 años. En el mismo año Bond et al. (1982) recogieron las emisiones de una niña hablante de inglés americano a lo largo de tres etapas: la etapa I fue

de 1;5 a 1;10 años, la etapa II comprendió los 2;2 años y la etapa III los 2;5 años. Tras analizar los dos primeros formantes los resultados revelaron que una superposición de todas las vocales, incluyendo las de los extremos ([i], [æ], [ɑ] y [u]), tenía lugar en las etapas primera y segunda descendiendo notoriamente en la etapa tercera (etapa esta última en la que además las vocales se hicieron más distintas). Según Fletcher (1973), especialista al que recuperan los autores, la superposición se debe a los gestos articulatorios imprecisos que poseen los bebés en las dos primeras etapas. Dicha imprecisión se irá reduciendo con el paso del tiempo ganando el bebé un mayor control articulatorio que se reflejará en un espacio vocálico más definido.

En el experimento de Kuhl y Meltzoff (1996) en el que analizaron las vocales articuladas por bebés de entre 0;3 y 0;5 años después de que estos escucharan las palabras *hop*, *heap* y *hoop* emitidas por un hablante femenino a través de una máquina, se observó que los espacios vocálicos ocupados por las vocales de los bebés [a, i, u] se iban separando progresivamente entre los 0;3 y los 0;5 años debido a que un estrecho agrupamiento de las vocales en cada categoría tenía lugar con el paso del tiempo. Para las autoras este cambio puede ser debido a dos factores: por un lado a una estabilización de los movimientos articulatorios del niño con el paso del tiempo (como se apuntaba en el trabajo de Bond et al. [1982]), y por otro al ya mencionado *vocal learning* puesto que quizás los bebés que están escuchando las vocales producidas por los adultos se esfuerzan en producir sonidos semejantes a los que los adultos emiten resultando ello en un mayor agrupamiento de las vocales en sus respectivas categorías.

Un resultado similar fue hallado en el estudio de Rvachew et al. (2006) en el que, después de analizar los formantes de las vocales aisladas y las contenidas en la sílaba CV de veintitrés niños expuestos al francés del Canadá y de veinte expuestos al inglés canadiense entre las edades de 0;10 y 1;6 años, se observó que una expansión del espacio vocálico tenía lugar en el habla de los niños pertenecientes a ambas comunidades lingüísticas (reduciéndose así la superposición vocálica inicial) si bien dicha expansión fue más acusada en la esquina difusa en el caso de los niños franceses y en la esquina grave en la de los ingleses (reflejando una vez más la influencia de la lengua ambiente en las producciones tempranas de los bebés).

También en el caso de la lengua japonesa se encontró una superposición vocálica en las primeras etapas del desarrollo tal y como manifiesta el trabajo de Ishizuka et al. (2007) quienes hallaron una superposición de las vocales [a] y [i], [o] e [i], y de [o] y [e] en los primeros meses en los que se analizaron las emisiones de un sujeto varón, esto es entre los 0;4 y los 0;6 años, y entre los 0;8 y los 1;6 años. Dicha superposición desapareció considerablemente a los 1;6 años. Se percibió también una expansión del espacio vocálico entre los 0;4 y los 1;6 años de edad que coincidió según los autores con la rápida maduración de los órganos articulatorios hasta

aproximadamente los 1;6 años, tiempo a partir del cual se produjo una maduración más gradual (Vorperian et al. [2005]).

En el estudio longitudinal de Rvachew et al. (2008), en el que se comparó la adquisición de la lengua materna por niños hablantes de inglés canadiense y por niños que hablaban francés de Canadá entre los 0;8 y los 1;6 años de vida (que habían estado expuestos a una sola de las lenguas y cuyos padres eran hablantes nativos de una sola de ellas), los datos confirmaron que una expansión del espacio vocálico tenía lugar conforme transcurría el tiempo y que estaba asociada según los autores al incremento de las vocales de los extremos [i] e [u] producidas por los niños de mayor edad en comparación con los sujetos más jóvenes.

Un trabajo que hemos optado por dejar para el final (al igual que hicimos en el capítulo del estado de la cuestión) por no haber podido averiguar la fecha de publicación, es el de Chi Lin que estudió el desarrollo vocálico de un niño que aprendía taiwanés mediante el análisis de los sonidos vocálicos comprendidos en sílabas aisladas y no nasales que contuvieran una de las seis vocales del taiwanés entre los 1;5 y 1;6 años (etapa primera) y a los 3;0 años de edad (etapa segunda). Chi Lin observó que una superposición en el espacio vocálico entre las diferentes categorías fue más notoria a los 1;6 que a los 3;0 años de edad estando aun así perfectamente separadas las vocales en el espacio vocálico ya a la edad de 1;6 años.

Llegamos en este punto a la quinta hipótesis de partida que planteaba el hecho de que a pesar de los cambios dramáticos que sufren las estructuras del tracto vocal durante el primer año de vida, los dos primeros formantes de los niños no decrecen de manera significativa durante este periodo de tiempo. Dicha hipótesis queda también corroborada en nuestros resultados ya que los ANOVA realizados, en los que se compararon los valores formánticos del primer y segundo formante entre los 0;4 y los 0;8 años con los obtenidos entre los 0;8 y el 1;0 año en las cinco vocales del español, arrojaron un resultado negativo evidenciando así la falta de un descenso significativo en los valores de los dos primeros formantes durante el primer año de vida a pesar de los considerables cambios que sufrió el tracto vocal durante este periodo de tiempo (cambios documentados por múltiples autores y expuestos en el punto 2.1.1 del estado de la cuestión).

Para Gilbert et al. (1997) y Robb et al. (1997) las razones principales de que no se cumpla la teoría acústica expuesta por Fant (1960), según la cual existe una la relación inversa entre el tamaño del tracto vocal y los valores de los formantes de forma que cuanto más crece el tracto vocal más bajos son los valores de estos, son por un lado el hecho de que haya que considerar la reestructuración anatómica como algo más que como un simple alargamiento del tracto vocal, y por otro la necesidad de caracterizar el crecimiento del tracto vocal como un



crecimiento no uniforme en cuanto a longitud, anchura y consecuentemente volumen. Respecto a la primera cuestión, Vorperian et al. (2008:1667) expusieron que ya era evidente en los escritos de Fant la importancia de considerar las dimensiones del tubo (aparte de la longitud del mismo) aunque él usase simples modelos de tubos para hacer diferentes interpretaciones:

“[...] However, it is evident in Fant’s writings (Fant, 1975) that although he used simple tube models to make physiologic-acoustic interpretations, he specified the importance of dimensions other than tube length- specifically laryngeal cavity- and indicated the need for more detailed anatomical studies and calculations. Thus, ultimately it is necessary to have a thorough multi-dimensional understanding of the anatomic development of the acoustic resonator, or the VT, between males and females to help establish anatomic-acoustic correlate during the course of development”.

También Lieberman (2001) nos habló de que la relación entre los formantes y la cavidad no debía limitarse a la longitud de dicha cavidad sino que también había que dar cuenta del ancho de esta.

En cuanto al segundo punto, recordamos que entre las conclusiones de los trabajos de Vorperian et al. (1999), Vorperian et al. (2005) y Vorperian et al. (2008), todas ellas desarrolladas con extensión en el primer apartado del capítulo segundo, destacaba la del crecimiento no lineal de las diferentes estructuras del tracto vocal tras ser estas observadas y medidas con modernas técnicas de resonancia magnética.

Así, para Vorperian et al. (2008) las características acústicas de las emisiones no dependían únicamente de la longitud del tracto vocal sino del diferente crecimiento de todas sus partes. Este crecimiento no uniforme quedó muy bien reflejado en la división sistemática que Vorperian et al. (2008) hicieron del tracto vocal separando las estructuras situadas en la zona anterior, estructuras orales u horizontales, de las ubicadas en la parte posterior del mismo, faríngeas o verticales. Estos autores concretaron que las estructuras anteriores presentaron un crecimiento neural que se caracterizaba por ser un crecimiento muy rápido alcanzando al final de la infancia temprana dos tercios de la estructura del adulto, mientras que las posteriores presentaron un crecimiento somático, es decir, crecimiento muy rápido en una primera fase que iba desde el nacimiento hasta la temprana infancia, pero al final de esta fase de rápido crecimiento la estructura era solo un cuarto de la del adulto. Respecto a las estructuras que están situadas en ambos planos (el vertical y el horizontal) como la longitud del tracto vocal y la longitud de la lengua, estas contaron con una curva de crecimiento combinado, esto es neuronal y somático, que originó que dichas estructuras fueran las que mayor crecimiento mostraron a cualquier edad.

Estos hallazgos coincidieron con los encontrados por investigadores como Buhr (1980), Kent y Murray (1982), Kuhl y Meltzoff (1996), Robb et al. (1997) y Gilbert et al. (1997) (todos citados al final del planteamiento de esta hipótesis en el capítulo tercero) si bien las diferencias surgieron en torno a la edad en la que la máxima de Fant comenzaba a cumplirse. A continuación recuperamos los resultados hallados tras el análisis de varianza o ANOVA, que fueron presentados en el capítulo anterior, y los comparamos con los obtenidos en otros estudios.

En un primer vistazo observamos que un descenso significativo del F2 se produjo para todas las vocales analizadas en ambos niños y en periodos de tiempo muy cercanos, excepto en los sonidos vocálicos anteriores. No obstante, debemos comentar la salvedad hallada dentro de las vocales anteriores puesto que los análisis revelaron que una reducción significativa de los valores del segundo formante tenía lugar entre los 0;4-1;0 año y los 3;0 años de vida (no al comparar el primer año con el segundo ni el segundo con el tercero) en la vocal anterior semialta analizada en el niño (descenso mínimo por otra parte), y entre cuyas causas apuntábamos la de la elevada diferencia en el número de señales analizadas en cada uno de estos periodos.

En cuanto a la evolución del segundo formante en el resto de vocales, estos fueron los resultados obtenidos en la niña: para la vocal [a] el análisis de varianza reveló que los valores descendieron significativamente entre los 1;5 y 1;6 años; unos meses más tarde, esto es entre los 1;8 y los 1;9 años, tuvo lugar el descenso significativo del F2 en la vocal posterior semialta [o] (los valores de este formante no cambiaron de manera significativa en ninguna de estas vocales hasta el final del periodo de estudio); para la vocal posterior alta [u] y a pesar de la imposibilidad de hacer el ANOVA debido a que tan solo contábamos con un dato en el segundo año de vida, la diferencia de 630 Hz entre el primer y el segundo año hace pensar que un descenso significativo se produjo entre el primer año de vida y el segundo, concretamente entre los 1;9 y 1;10 años (el segundo formante de esta vocal no solo no se redujo entre los 1;10 años y el final del periodo de estudio sino que aumentó ligeramente).

En el niño, los análisis de varianza concluyeron lo siguiente: en la vocal central baja [a] el segundo formante experimentó un descenso entre los 1;7 y los 1;8 años. En la vocal posterior semialta [o] los valores se redujeron de manera significativa entre los 1;9 y los 1;10 años (manteniéndose sin variaciones importantes hasta el final del estudio), y en la vocal posterior alta [u], a pesar de no haber podido realizar los análisis de varianza entre el primer y el segundo año de vida por la presencia de una sola señal en el primer año analizado, la diferencia de 615 Hz entre el primer y el segundo año parece indicarnos que sí que hubo un descenso significativo

del F2 entre el primer y el segundo año de vida (dicho descenso se produjo concretamente entre los 1;7 y los 1;8 años). Como podemos observar, los valores del segundo formante descendieron de manera significativa en la vocal central y las vocales posteriores de ambos niños y lo hicieron además alrededor del mismo periodo, esto es, entre los 1;5 y los 1;10 años.

En cuanto a la evolución de los valores del primer formante, también esta fue similar en ambos niños puesto que no hubo descenso significativo en las vocales [a, e, i] en ninguno de los sujetos (salvo el F1 de la [e] de la niña que sí descendió entre el primer y el tercer año pero al ser este descenso tan solo de 92 Hz este podría deberse tal y como apuntábamos al diferente número de señales) y sí lo hubo en la vocal [u] en ambos niños.

Estos fueron los resultados obtenidos en la niña: ninguna reducción significativa se produjo en el F1 de las vocales [a, e, i] en ninguno de los periodos comparados (salvo el ya señalado descenso mínimo del F1 de la [e] entre el primer y el tercer año de vida), pero sí que se halló un descenso significativo de los valores del F1 en la vocal posterior semialta [o] al comparar los datos obtenidos en el primer año de vida con los del segundo (concretamente la reducción se produjo entre los 1;9 y 1;10 años). El primer formante de la vocal posterior alta [u] pareció experimentar también un descenso significativo entre el primer y el segundo año de vida puesto que a pesar de que no pudo realizarse el ANOVA por la falta de señales, la diferencia de 275 Hz entre estos dos periodos (concretamente entre los 0;4 y 1;0 año y los 1;9 y 1;10 años) parece indicar que sí hubo un descenso significativo. El mismo problema encontramos al comparar el segundo año de vida con el tercero, limitado número de señales, pero en este caso la escasa diferencia de 29 Hz entre ambos periodos nos lleva a pensar que no hubo diferencia significativa.

Respecto a la evolución del F1 en el niño, ningún descenso significativo tuvo lugar en las vocales [a, e, i, o] entre el primer y el segundo año de vida salvo el descenso que se produjo en el F1 de la [e] entre los cuatro primeros y los cuatro últimos meses del primer año de vida, reducción que como ya apuntamos no pudo deberse a causas anatómicas debido a que en el segundo año de vida el promedio de este formante fue superior al del primero. En cuanto al primer formante de la vocal posterior alta [u], y a pesar de no poder realizarse el ANOVA por la escasez de señales halladas en ambos periodos (tal y como ocurría en el caso de la niña), la diferencia de 345 Hz entre ambos periodos parece indicar que un descenso significativo tuvo lugar entre el primer y el segundo año de vida (concretamente entre los 1;7 y los 1;8 años). El F1 de esta vocal no experimentó ningún cambio significativo desde los 1;8 años hasta el final del periodo.

A continuación, nos referiremos brevemente a los trabajos que arrojan datos sobre la evolución de los dos primeros formantes a lo largo del tiempo pasando por alto el orden temporal que viene caracterizando esta tesis para comenzar de manera justificada con dos artículos determinados por comprender estos un periodo de estudio semejante al de este trabajo y por ser sus resultados similares a los nuestros y prácticamente idénticos entre ellos (retomaremos después la sucesión temporal predominante en esta tesis). Como en los estudios anteriores, se procederá a equiparar los resultados hallados en cada trabajo con los encontrados en nuestra investigación (siempre que el tiempo de estudio sea semejante y permita dicha comparación) señalando en cada caso si las conclusiones obtenidas sobre la evolución de los formantes se obtuvieron teniendo en cuenta todas las vocales emitidas por el niño en un periodo de tiempo determinado sin que se hiciera diferenciación de timbre, o si esta tuvo lugar.

Se trata en primer lugar del trabajo de Gilbert et al. (1997) en el que se analizaron los sonidos vocálicos, bien aislados bien en contextos consonánticos, de cuatro niños que adquirirían el inglés americano a las edades de 1;3, 1;6, 1;9, 2;0 y 3;0 años. La ausencia de un test de percepción exhaustivo les llevó a clasificar las vocales a partir de sus valores formánticos en amplias categorías como fueron la altura de la lengua en el eje vertical, etiquetándose las vocales en vocal alta, media o baja, y el avance de la lengua en el horizontal, clasificándose las mismas en vocal anterior, central o posterior. El ANOVA determinó que no hubo descenso significativo en la media del F1 y F2 entre los 1;3 y los 1;9 años pero sí entre los 1;3 y los 3;0 años de edad.

En los resultados obtenidos en nuestro estudio el descenso significativo de los formantes se produjo alrededor de la misma edad que en el estudio de Gilbert et al. (1997) ya que en el caso del primer formante de las vocales posteriores de la niña, de la alta y de la semialta, este descenso ocurrió entre los 1;9 y 1;10 años. En el caso del niño, la reducción significativa del primer formante, que tan solo se produjo en la vocal posterior alta, tuvo lugar entre los 1;7 y 1;8 años. En cuanto al F2, los descensos significativos se produjeron en el caso de la niña entre los 1;5 y 1;6 años en la vocal central baja [a], entre los 1;8 y 1;9 años en la posterior semialta [o] y entre los 1;9 y 1;10 años en la posterior alta [u]. En el niño, el descenso del segundo formante tuvo lugar entre los 1;7 y 1;8 años en la [a] y en la [u], y entre los 1;9 y 1;10 años en la [o].

En el mismo año Robb et al. (1997) llevaron a cabo una segunda investigación en la que estudiaron las producciones de veinte niños que adquirirían la lengua inglesa entre los 0;4 y los 2;1 años de edad tomando muestras a los 0;4, 0;8, 1;3, 1;6, y 2;1 años. En esta ocasión los

valores del F1 y F2 no registraron bajadas significativas entre los 0;4 años y los 1;6 años asemejándose estos resultados una vez más a los hallados por nosotros.

Presentamos a continuación otros trabajos en los que se siguió la evolución de los dos primeros formantes en periodos de tiempo que no coincidieron exactamente con los periodos temporales que comprendió nuestro estudio. Comenzamos por el estudio de Eguchi y Hirsh (1969) en el que tras analizar los dos primeros formantes de seis vocales incluidas en dos oraciones que debían pronunciar ochenta y cuatro sujetos agrupados en las edades de 3;0, 4;0, 5;0, 6;0, 7;0, 8;0, 9;0, 10;0, 11;0, 12;0 y 13;0 años, se evidenció un claro descenso de estos a lo largo de todo el periodo de estudio, descenso que se hizo más acusado entre los 3;0 y los 5;0 años. Por otro lado, el descenso fue mayor para el segundo que para el primer formante en todas las vocales excepto en la [a].

De los cuatro sujetos que adquirirían la lengua inglesa y que fueron seguidos por Lieberman (1980) para recoger sus emisiones espontáneas entre los 0;4 y los 5;0 años, tan solo la evolución de los dos primeros formantes de las vocales de uno de estos sujetos se explicita en el artículo. Se trata del sujeto L. S. cuyos dos primeros formantes se mantuvieron estables entre los 0;4 y los 1;4 años de edad debido según Lieberman al escaso crecimiento del tracto vocal supralaríngeo en este periodo. Estos datos coinciden con los nuestros pues ni el F1 ni el F2 de las vocales de los sujetos de nuestro estudio descendieron de manera significativa entre los doce y los dieciséis meses de edad.

El mismo periodo de tiempo comprendió el estudio llevado a cabo por Buhr (1980) en el que se recogió el habla espontánea de un niño entre los 0;4 y los 1;4 años de edad y se analizó la evolución de los tres primeros formantes. Los resultados para las vocales compartidas con el español fueron los siguientes: la vocal [e] apenas presentó cambios a lo largo del tiempo estudiado como sucediera con los resultados obtenidos para los dos sujetos que adquirirían el español; no se especificó la evolución de los dos primeros formantes en las vocales [a] e [i] aludiendo simplemente al hecho de que sus frecuencias registraron fuertes subidas y bajadas en el periodo de tiempo señalado (más en el caso de la [i] que en el de la [a]). El caso contrario fue el de la [u] que no presentó cambios significativos a lo largo del periodo de estudio (resultado idéntico al hallado en las resonancias de esta vocal en los dos sujetos de nuestro estudio). Finalmente, Buhr (1980) añadió que debido al reducido número de muestras no pudieron extraerse conclusiones sobre los cambios formánticos de la vocal [o].

Más reducido fue el periodo de tiempo que comprendió el trabajo de Kuhl y Metlzoﬀ (1996) en el que se estudió la evolución de los dos primeros formantes entre los 0;3 y los 0;5 años. No obstante, los resultados fueron los mismos que los hallados en el estudio anterior y en

el presente trabajo, esto es, ningún cambio significativo tuvo lugar en los dos primeros formantes de las tres vocales de los extremos, es decir [i, a, u], entre los 0;3 y los 0;5 años de vida.

Únicamente en los formantes de la vocal [ɪ] se centró el trabajo de Smith y Kenney (1998) quienes recogieron la pronunciación de la palabra *sissy* en siete sujetos femeninos que fueron grabados en intervalos de año y medio a las siguientes edades: en el tiempo 1 las niñas tenían entre 8;0 y 9;0 años de edad (media de 8;7 años); en el tiempo 2, tenían entre 9;0 y 10;0 años de edad (media 10;1 años); y en el tiempo 3 las edades eran de 11;0 y 12;0 años (media de 11;5 años). Tras realizar un test de varianza los resultados, que se muestran imposibles de comparar con los nuestros debido a la diferencia de edad entre estos sujetos y los nuestros, indicaron que los cambios en el F2 y en el F3 fueron significativos a lo largo del tiempo mientras que el descenso del F1 no registró significatividad.

Del trabajo de Lyakso et al. (2005), investigación que como comentábamos en el estado de la cuestión se apartó en principio de nuestro estudio al intentar averiguar si los adultos eran capaces de reconocer las palabras con significado y las oraciones de niños que adquirían la lengua rusa a lo largo de los 3;0 años de edad, nos quedamos con una reflexión final según la cual los formantes de la vocal [o] alcanzaron valores adultos al final del tercer año de vida. Si bien los autores no concretaron la evolución de los dos primeros formantes de dicha vocal a lo largo del tiempo, deducimos que un descenso significativo hubo de producirse, a pesar de no conocer el momento temporal exacto, para que esta vocal presentara valores adultos al final del tercer año de vida. Por otro lado, este resultado no coincide con los resultados obtenidos para esta vocal en el tercer año de vida de nuestros sujetos puesto que en nuestro estudio los valores de la vocal [o] superaron a los dados por Martínez Celdrán y Fernández Planas (2007) para la vocal posterior semialta del adulto femenino y varón, si bien es cierto que a esta edad (los 3;0 años) los valores formánticos de ambos niños ya habían descendido considerablemente asemejándose más a los de los adultos.

Los resultados arrojados en el estudio de Rvachew et al. (2006) tras analizar los formantes de las vocales en el contexto de CV en cuarenta y tres niños de entre 0;10 y 1;6 años mostraron una vez más la influencia de la lengua ambiente en las producciones de los niños. De esta forma, un descenso producido con el tiempo tuvo lugar en los valores del F1 de las vocales de los niños que adquirían el francés y no en las de los niños de habla inglesa, mientras que en el caso del F2 el descenso solo ocurrió en las vocales de los sujetos que hablaban inglés y no en las de los franceses. Recordamos no obstante, los inconvenientes de realizar los análisis

estadísticos sobre el total de señales obtenidas en el mes (y no sobre las señales agrupadas por timbre vocálico) ya que estos pueden oscurecer la evolución real de los valores formánticos.

Las autoras se refieren en este trabajo al de Rvachew et al. (1996) en el que tras describir las vocales producidas por nueve sujetos que adquirían la lengua inglesa entre los 0;6 y los 1;6 años de edad, se descubrió que la media del F1 fue estable a lo largo de este periodo de tiempo mientras que los valores del F2 aumentaron. Una vez más, los análisis de varianza se realizaron sobre el conjunto total de vocales sin que la diferenciación de timbre se estableciera trayendo consigo esta circunstancia la problemática ya comentada.

En la lengua japonesa, Ishizuka et al. (2007) estudiaron las producciones vocálicas de un sujeto varón entre los 0;4 y 0;6 años, entre los 0;8 y los 1;8 años, a los 1;10, a los 2;0, a los 2;1, a los 2;6, a los 2;10, a los 3;4, a los 3;8 y a los 4;4 años de edad, y las de un sujeto de sexo femenino entre los 0;4 y los 1;10 años, y a los 2;0, 2;1, 2;6, 2;11, 3;4, 3;9, 4;2, 4;7 y 5;0 años de edad. Los resultados hallados fueron los siguientes: los valores del F1 para la [a] en ambos niños decrecieron hasta la edad de entre 1;8 y 2;6 años, permaneciendo sin cambios entre los 2;1 y los 3;4. En nuestros sujetos, sin embargo, el primer formante de la [a] se mantuvo sin cambios a lo largo del periodo mencionado. Para la vocal [e], el F1 descendió en ambos niños hasta los 2;6 años, edad a partir de la cual los valores del F1 de la niña continuaron decreciendo mientras que los del niño apenas variaron. Nuestros resultados mostraron que ningún descenso significativo se producía en el primer formante de esta vocal en el caso de la niña (salvo entre el primer y el tercer año con una pequeña diferencia de 92 Hz), ni en el del niño (excepto el descenso entre los cuatro primeros y los cuatro últimos meses no debido a causas anatómicas). El primer formante de la vocal [i] en ambos sujetos decreció hasta la edad de los 1;8 años edad a partir de la cual los valores del niño se mantuvieron sin cambios mientras que los de la niña continuaron cayendo. Este formante por el contrario se mantuvo sin cambios en las vocales anteriores altas de nuestros dos sujetos entre el primer y el tercer año de vida. En la vocal [o], los valores del primer formante se mantuvieron sin cambios durante todo el periodo para el infante varón mientras que en la pequeña los valores del F1 fueron constantes hasta los 1;8 años y después descendieron. La misma evolución presentó el primer formante de esta vocal en nuestros dos sujetos, esto es, se mantuvo estable en el niño pero descendió en la niña entre los 1;9 y los 1;10 años. Y en cuanto al F1 de la última vocal, el primer formante de la [u] del sujeto varón descendió hasta los 1;8 años y después se mantuvo sin cambios mientras que en el caso de la niña la evolución fue la contraria permaneciendo el F1 de esta vocal sin cambios hasta los 1;8 años y descendiendo después. En nuestros resultados, el primer formante de la [u] descendió de manera significativa en ambos sujetos apareciendo esta reducción en la niña a los 1;10 años y en el niño a los 1;8 años.

Respecto al F2, los valores de este formante para la vocal [a] de los niños japoneses descendieron hasta los 1;8 años y después se mantuvieron sin cambios. En los sujetos de habla española el patrón fue el contrario ya que este formante se situó alrededor de los mismos valores hasta los 1;5-1;6 años en la niña y los 1;7-1;8 años en el niño, edad en la que descendieron significativamente. En la vocal anterior semialta [e] del estudio japonés, el segundo formante analizado en las emisiones del niño permaneció sin cambiar hasta alrededor de los 1;8 años, decreciendo entre esa edad y los 2;6 años y manteniéndose estable a partir de esta última. En la niña japonesa, el segundo formante de la [e] apenas registró cambios entre los 0;4 y los 5;0 años de edad. En cuanto al segundo formante de la vocal anterior semialta de nuestro estudio, este se mantuvo estable a lo largo del tiempo en ambos sujetos con la excepción del cambio significativo que apareció entre el primer y tercer año de vida en el caso del niño por las razones ya explicadas. Coincidentes también son nuestros resultados con los hallados por los autores para el sujeto varón en la vocal anterior alta [i] que no registró ningún cambio entre el primer y el tercer año de vida estudiados. Sin embargo, en el caso del sujeto femenino que aprendía japonés, el segundo formante de esta vocal aumentó con la edad mientras que este se mantuvo sin cambios en los resultados hallados en la niña de habla española. Opuesta fue la evolución del segundo formante de la vocal posterior semialta [o] cuyos valores decrecieron en los dos niños japoneses hasta la edad de los 1;6 años permaneciendo después sin cambiar, mientras que en el caso de los niños españoles este se mantuvo sin cambios hasta los 1;8 y 1;9 años en la niña y los 1;9 y 1;10 años en el niño, momento en el que ambos descendieron de manera significativa. El mismo patrón registrado por la vocal [o] presentó la vocal [u] de los niños japoneses ya que el F2 de ambos niños decreció hasta los 1;6 años manteniéndose después sin cambios notorios. De nuevo la evolución del segundo formante de esta vocal fue la contraria en el caso de los sujetos españoles cuyo segundo formante para esta vocal se mantuvo sin cambios hasta los 1;9 y 1;10 años en la niña y los 1;7 y 1;8 años en el niño, mostrando en estas edades un descenso significativo.

Los resultados obtenidos tras la recopilación de trabajos llevada a cabo por Vorperian y Kent (2007) apoyaron la predicción de la teoría acústica de que los formantes decrecen con la edad según aumenta el tamaño del tracto vocal puesto que los gráficos que elaboraron a partir de los datos obtenidos en los quince trabajos que recogieron en su estudio reflejaron un descenso del área vocálica entre los 0;7 y los 11;0 años de vida. Esto llevó a los autores a considerar los trabajos de Gilbert et al. (1997) y de Robb et al. (1997) como una excepción a dicha teoría acústica puesto que en ellos no se producía ningún descenso significativo de los formantes en los dos primeros años de vida. No obstante y tal y como apuntábamos en el estado de la cuestión, el cálculo del área de los cuadriláteros vocálicos pasó en el estudio de Vorperian y



Kent (2007) de los 0;7 y 1;0 años a los 3;0 años quedando un vacío entre el 1;0 año y los 2;0 años de edad, periodo en el que quizás no se produjo ningún descenso de los valores formánticos. El trabajo de Vorperian y Kent (2007) dejó también un hueco entre los 2;0 años y los 3;0 años de vida tiempo a partir del cual quizás se produjo la caída formántica, como sucedió en el estudio de Gilbert et al. (1997), si es que esta no se había producido anteriormente, es decir a partir de los 1;7-1;10 años de vida, como ocurrió con el segundo formante de las vocales del presente trabajo.

El último trabajo al que nos referiremos es, como en el capítulo segundo, el artículo de Chi Lin que hemos vuelto a dejar para el final debido a la imposibilidad de identificar la fecha en la que se publicó. Este recogía las emisiones de un sujeto que aprendía taiwanés en dos etapas, entre los 1;5 y 1;6 años (etapa I) y a los 3;0 años de edad (etapa II). Después de medir los dos primeros formantes en ambas etapas los resultados fueron los siguientes: la media del primer formante de la [i] fue más baja en la segunda etapa que en la primera mientras que la del segundo formante aumentó en la etapa segunda. Si comparamos estos datos con nuestros resultados observamos que en nuestros sujetos ninguno de los dos formantes de la vocal [i] aumentó o se redujo de manera significativa en el tiempo indicado. Respecto a la [e], ambos formantes fueron más altos en la segunda etapa que en la primera oponiéndose estos resultados a los hallados por nosotros y en los que los dos formantes de esta vocal se mantuvieron sin cambios significativos entre el primer y el tercer año de vida (habiendo sufrido el F1 de la [e] en la niña y el F1 y F2 de esta vocal en el niño un descenso por las causas comentadas). Por su parte, la media del F1 de la [a] fue la misma en ambas etapas aunque en el caso del F2, esta fue significativamente más alta en la etapa II que en la etapa I. En nuestros resultados, tanto en el niño como en la niña la media del primer formante de la [a] no sufrió variaciones a lo largo del tiempo experimentando la del F2 un descenso significativo entre los 1;5 y 1;6 años en la niña y entre los 1;7 y 1;8 años en el niño. Para la [u] ni el F1 ni el F2 variaron de manera significativa con el paso del tiempo a diferencia de lo que ocurrió con estos dos formantes en los sujetos que adquirirían el español y que fue lo siguiente: el F1 de la niña descendió de manera significativa entre los 1;9 y 1;10 años y en el niño entre los 1;7 y 1;8 años. Respecto al F2, este se redujo significativamente a la misma edad que lo hizo el F1 en cada uno de los niños. Finalmente y al no existir datos del taiwanés para la vocal posterior semialta [o], hemos optado por tomar los que este trabajo ofrece para la posterior semibaja [ɔ] teniendo siempre presentes las diferencias existentes entre estas dos vocales, y según los cuales la media del primer y del segundo formante de esta vocal fue mayor en la segunda etapa que en la primera. En nuestro trabajo, el primer formante de esta vocal se mantuvo estable en el caso del niño mientras que en el de la niña este se redujo de manera significativa entre los 1;9 y 1;10 años, esto es entre el primer y el

segundo año de vida. Respecto al F2, hubo un descenso significativo con el paso del tiempo en los valores de los dos sujetos españoles que ocurrió entre los 1;8 y 1;9 años en la niña y entre los 1;9 y 1;10 años en el niño. Según los autores, de estos datos se desprende que únicamente para el primer formante de la [i] se produjo un descenso significativo entre los 1;8 y los 3;0 años de edad mientras que en general en el resto de las vocales los dos primeros formantes aumentaron ligeramente con la edad (resultados muy diferentes a los nuestros y a los hallados en otros trabajos ya revisados como los dados por Gilbert et al. [1997]).

Una vez revisada la literatura en la que se analizó como en nuestro estudio la evolución de los dos primeros formantes, podemos concluir con la existencia de una falta de acuerdo en los artículos publicados hasta el momento en torno a la edad en la que los dos primeros formantes sufren un descenso significativo. Lo que todos ellos comparten, por otra parte, es la ausencia de una reducción notoria de los formantes a lo largo del primer año de vida que es también el resultado que arrojó el ANOVA en nuestro trabajo.

Antes de dejar de lado la evolución formántica de las vocales y de ver si se ha cumplido la última hipótesis planteada, nos gustaría señalar brevemente la diferencia en la evolución de las áreas vocálicas hallada entre este estudio y otros trabajos. Dicha disimilitud radicó en la distinta evolución formántica a partir del primer año de vida puesto que la reducción de los valores del F2 en las vocales [a] y [u] de los sujetos de nuestro estudio y la ausencia de dicho descenso en el segundo formante de la [i] dio lugar a un estiramiento del triángulo hacia la parte central y posterior de la carta de formantes que conllevó un aumento del área vocálica. En el trabajo de Vorperian y Kent (2007) por el contrario, el área del cuadrilátero se redujo entre los 0;7 y los 11;0 años de vida debido al descenso significativo de los dos primeros formantes en ese periodo (a pesar de existir un vacío ya comentado entre los 1;0-2;0 años y los 2;0-3;0 años de edad).

No obstante, el desplazamiento del área vocálica de ambos trabajos fue muy similar ya que frente al estudio de Chi Lin en el que el área del cuadrilátero se desplazó hacia la parte anterior y baja de la carta de formantes, en el trabajo de Vorperian y Kent (2007) y en la presente investigación el área vocálica se desplazó hacia la parte alta y posterior de dicha carta de formantes debido al descenso significativo del F2 en las vocales [u] y [a], aunque al no producirse en nuestro estudio esta disminución en el segundo formante de la [i], el triángulo siguió estando anclado a la parte anterior del gráfico sin desplazarse completamente hacia la parte posterior de este.

Finalmente y antes de abandonar definitivamente el tratamiento acústico de las vocales, creemos conveniente comentar de manera breve la evolución de la desviación estándar en

relación a los resultados hallados en otros trabajos pero solo en aquellos en los que este análisis se calculó sobre vocales concretas y no sobre el conjunto total de las mismas debido a que en este último caso el aumento o descenso de esta podría deberse al predominio o escasez de determinadas vocales.

Tal y como mostraban las tablas del capítulo siete y como comentamos en la comparación de los resultados obtenidos tras los análisis acústicos realizados en las emisiones de niña y en las del niño (apartado 7.1.3 de dicho capítulo), la desviación estándar hallada en los valores de cada una de las vocales fue muy similar en ambos sujetos. Además, aunque esta tendió a descender en la niña con el paso del tiempo (excepto la del F1 de la [a]) y a aumentar en el niño (excepto en los dos formantes de la [u], en los dos de la [e] y en el segundo de la [i]) las diferencias no fueron significativas. Asimismo, en la mayoría de las vocales de ambos niños la desviación estándar del F2 fue mayor que la del F1 si bien la diferencia no fue muy acusada.

Estos hallazgos coinciden con los resultados obtenidos en los trabajos de De Boysson-Bardies et al. (1989) y Gilbert et al. (1997). En el primero, el análisis de la desviación estándar de los dos primeros formantes de las vocales emitidas por veinte niños de 0;10 años que adquirirían francés, inglés, argelino y cantonés reflejó que la diferencia en la desviación estándar del primer formante entre los niños hablantes de una misma lengua era menor que la encontrada en el F2. Esta circunstancia se dio también en nuestros resultados en los que los valores del segundo formante distaron más unos de otros en ambos niños que los del primero, si bien la diferencia entre unos y otros no fue elevada.

A pesar de que resulta imposible la comparación de nuestro trabajo con el de Casal et al. (2002) puesto que en este último no se calculó la desviación estándar de los formantes a lo largo del tiempo sino que se mostró en un solo momento temporal (media de 2;3 años), los valores obtenidos por Casal et al. (2002) a esta edad son mucho más elevados que los nuestros. De esta forma, Casal et al. (2002) hallan una desviación estándar de 452 Hz en el F2 de la [i] de los niños con desarrollo normal que fueron comparados con los que tenían el labio partido, mientras que el valor más alto encontrado en nuestros sujetos fue de 201 Hz. Algo más notoria fue la diferencia entre la desviación estándar del segundo formante de la [e] de este mismo grupo de niños, de 640 Hz, y del grupo de niños con desarrollo normal que se comparó con aquellos que tenían el paladar partido, de 625 Hz, y la encontrada en nuestros sujetos que en ningún caso superó los 189 Hz. Otro dato que supera con creces la desviación estándar de nuestros sujetos es la hallada en el segundo formante de la vocal [i] en el grupo de niños con desarrollo normal que fue equiparado con el de los sujetos que presentaban el paladar partido: 981 Hz frente a los 201 Hz que como ya hemos comentado presentaron nuestros sujetos. En el resto de vocales y de

grupos de niños, la desviación estándar fue similar a la obtenida en nuestra tesis si bien los valores del estudio de Casal et al. (2002) siempre fueron algo más altos. En cuanto a las razones, quizá esta diferencia se deba principalmente al número de sujetos sobre cuyas emisiones se realizó dicho cálculo y que fue de seis por grupo en el trabajo de Casal et al. (2002) y de dos en el nuestro.

Las características propias de la publicación de Vorperian y Kent (2007) al tratarse esta de una recopilación de quince trabajos en los que se llevó a cabo el estudio de las vocales que aparecían en distintos contextos, emitidas por bebés de diferentes edades y que adquirirían distintos dialectos del inglés americano, dificultaron en cierto sentido que una comparación con nuestra investigación pudiera ser hecha. No obstante, los resultados hallados por los autores en cuanto a la desviación estándar fueron los siguientes: hubo una variabilidad mínima en el F1 para las vocales altas, pero esta aumentó en el F2 particularmente para las vocales altas posteriores como [u] a lo largo del desarrollo (no hay datos del resto de las vocales). En nuestro trabajo sin embargo, la desviación estándar de las vocales altas y de la posterior alta fue por el contrario muy parecida a la del resto de vocales y en ningún caso fue la más baja. Sin embargo, insistimos en la dificultad de comparación del trabajo de Vorperian y Kent (2007) con cualquier otro estudio al presentar dicho trabajo datos procedentes de estudios tan heterogéneos.

Diferentes son los resultados sobre la variabilidad formántica obtenidos en el trabajo de Chi Lin que recordemos estudió las vocalizaciones de un niño que adquiriría taiwanés en dos etapas diferentes, entre los 1;5 y 1;6 años (etapa I) y a los 3;0 años de edad (etapa II). La desviación estándar para cada una de las vocales fue la siguiente: el rango formántico de los dos primeros formantes de las vocales [i] y [a] fue más estrecho en la segunda etapa que en la primera coincidiendo estos resultados con los obtenidos en nuestro estudio en el caso de la niña (excepto en la desviación estándar del F1 de la [a] que aumentó ligeramente), pero no en el del niño donde la desviación estándar de estas vocales aumentó ligeramente en el transcurso de una etapa a otra (excepto el F2 de la [i] que descendió ligeramente). Para la [ɔ] y la [u], la diferencia en los valores formánticos tanto del primer como del segundo formante aumentó en la segunda etapa con respecto a la primera, algo que ocurrió solo en los dos primeros formantes de la [o] del sujeto varón de nuestro estudio ya que la desviación estándar de estos formantes para las vocales [o, u] de la niña y [u] del niño descendió con el paso del tiempo. Finalmente para la [e], en el estudio de Chi Lin el rango de valores aumentó en la etapa segunda con respecto a la primera para el primer formante mientras que para el segundo este se redujo. Sin embargo, en los resultados arrojados por nuestro estudio la desviación estándar de los dos primeros formantes de esta vocal disminuyó con el paso del tiempo en ambos sujetos.

Más allá de estas semejanzas y diferencias cabe preguntarse a qué razones se debe la variabilidad que presentan los valores de los formantes de los bebés o niños de corta edad. Parte de estas causas fueron expuestas en el capítulo segundo de esta investigación por lo que creemos conveniente recuperar a continuación los trabajos en los que se planteaban algunos de los motivos de esta varianza formántica.

En el artículo de Eguchi y Hirsh (1969:44) la variabilidad de los formantes descendió según aumentaba la edad del niño, entre los 3;0 y los 11;0 años de edad, de modo que para estos autores la desviación estándar ha de reducirse según se desarrolla el sujeto. Estos atribuyeron la mayor variación hallada a los 3;0 años de edad a la mayor inexactitud de los sujetos en las posiciones articulatorias en comparación con los niños mayores y los adultos (también Kent [1976, 1992] apuntaba a la inmadurez en el control motor como causa de la mayor varianza de los pequeños). Según estos autores:

“[...] The implication is clear that 3-year-old children do not move their tongues to exactly the same position for a particular vowel as it occurs in repetitions of the same sentence, at least not so exactly the same as older children and adults”.

La variabilidad formántica puede ser por tanto explicada por la falta de exactitud o la mayor variabilidad de movimiento, motivo que recogía de manera muy aclaradora el artículo de Sharkey y Folkins (1985) desarrollado en el estado de la cuestión. En un breve recordatorio de los puntos de vista que en este aparecían encontramos que para autores como Bruner (1973) la variabilidad de movimiento (y por tanto la formántica) se reducía una vez que el niño aprendía a realizar una tarea articulatoria; para otros como Bernstein (1967) la variabilidad no era un error sino un fenómeno natural que se daba al repetir una misma tarea y a la imposibilidad de hacerlo con el mismo gesto articulatorio (o al menos con uno exacto). Para este autor la variación en el movimiento podría explicarse con el aumento de la habilidad articulatoria del niño al incrementarse las posibilidades de realizar una misma tarea con diferentes gestos funcionalmente equivalentes (según Bernstein, las repeticiones de una tarea son raramente repetidas con los mismos parámetros de movimiento y en vez de eso las tareas motoras emplean conjuntos de estructuras coordinativas las cuales producen muchos patrones de movimientos funcionalmente equivalentes de modo que según aumenta la habilidad articulatoria, el niño puede aprender nuevas formas sobre cómo organizar la estructura coordinativa –explotar su organización- para realizar la misma tarea pudiendo aumentarse la variabilidad muestra a muestra).

Esta es también la opinión de Kelso y Norman (1978), Kerr y Booth (1978) y Moxley (1979), para quienes la variabilidad se explica como un juego exploratorio que facilita al bebé el

aprendizaje de la lengua. Los propios Sharkey y Folkins (1985) se plantearon cuáles eran las verdaderas razones de este descenso en la variabilidad de movimiento al hallar en su trabajo que la variación o inestabilidad de parte de las variables medidas en su estudio (como el desplazamiento de la mandíbula) no descendía con el paso del tiempo (entre los 4;0 y 10;0 años) mientras que autores como Eguchi y Hirsh (1969), DeSimoni (1974a), (1974b), (1974c), Tingley y Allen (1975), Kent y Forner (1980) y Flege (1982) afirmaron que dicha variabilidad era una cuestión de madurez y por tanto descendía con la edad.

Callan et al. (2000), añadieron un segundo factor como posible determinante de la variedad formántica. En su estudio los autores hablaron de como la configuración del tracto vocal cambiaba durante el desarrollo del bebé y de que por tanto el lugar donde se producían las resonancias y las propiedades de estas se movía. De esta forma, la varianza o desviación estándar que presentan los valores de los formantes del bebé puede deberse según los investigadores al constante cambio al que está sometido el tracto vocal en esas edades.

Para Vorperian et al. (2005) y Vorperian et al. (2008) el aumento del rango de valores de los formantes puede ser debido a la mayor variabilidad del movimiento de la lengua producida por el aumento del tracto vocal que tiene lugar en las primeras etapas del desarrollo del bebé. De este modo tras la reestructuración anatómica, la lengua dispondría de más espacio para moverse dentro de la cavidad bucal y esto aumentaría la variabilidad de su movimiento provocando también un incremento en el rango formántico. Vorperian et al. (2005) calificaron esta variabilidad formántica de fenómeno de dispersión de los valores. Por otra parte, además de la mayor variabilidad de movimiento, los autores creen que los numerosos cambios en el tamaño y en la forma que sufre el tracto vocal del bebé podrían estar detrás de esta varianza (variación que según los autores no conlleva diferencias significativas en los formantes).

También Vorperian y Kent (2007) achacaron la variabilidad formántica característica de los bebés y niños pequeños a estos dos factores: la variabilidad articulatoria en general así como al crecimiento no uniforme del tracto vocal.

Llegamos en este punto a la última hipótesis de partida según la cual a pesar de que la mayor parte de los sonidos vocálicos articulados por niños pertenecientes a una misma comunidad lingüística poseen el mismo timbre, existen ya en las primeras etapas del desarrollo diferencias inter-individuales. Para averiguar si nuestros resultados confirman esta hipótesis volvemos al apartado 6.2 en el que se ofrecía la transcripción de las vocales identificadas en cada uno de los meses de los tres periodos analizados, y observamos que los resultados obtenidos tras el test de percepción la corroboran. Veamos lo que ocurre año por año.

Si nos fijamos en las señales modales del primer periodo analizado, la vocal más producida por ambos niños fue la anterior semialta [e] seguida en el caso de la niña de la vocal central baja [a] y de la anterior alta [i], y en el caso del niño de la anterior alta [i] y de la central baja [a] (si bien la diferencia entre ambas fue pequeña). En penúltimo lugar encontramos en ambos niños la vocal posterior semialta [o] y en el último puesto la posterior alta [u]. En las señales nasales, el orden se invirtió en el caso de la niña ya que la vocal central baja [ã] superó a la anterior semialta [ẽ] mientras que en el niño esta última conservó su posición anterior. Respecto a la [ĩ], esta vocal ocupó el tercer puesto en el niño mientras que no fue identificada en ninguna ocasión en las señales nasales de la niña. En cuanto a las vocales nasales posteriores, tan solo se halló la vocal semialta [õ], vocal que ocupó el último lugar en ambos sujetos.

En cuanto a las vocales modales del segundo año de vida, la vocal anterior alta [i] incrementó su presencia en ambos niños apareciendo en el niño casi el doble de veces que la anterior semialta [e], y quedando solo una muestra por detrás de esta vocal en el caso de la niña. Una diferencia con el periodo anterior es que en ambos niños la vocal central [a] pasó a ocupar las últimas posiciones: la penúltima en el caso de la niña, solo por delante de la [u], y la última en el del niño. Por primera vez las vocales posteriores sobrepasaron a la central baja [a]. Respecto a los sonidos vocálicos nasales, la [ẽ] ocupó el primer puesto y la [ã] el segundo en las emisiones de la niña mientras que en las del niño el orden se invirtió. En ambos sujetos la vocal posterior semialta [õ] superó a la anterior alta [ĩ] quedando en último lugar la vocal posterior alta [ũ] que aumentó notablemente en las producciones del niño.

En el último año de vida observado y en relación a las vocales modales, la vocal posterior semialta [o] aumentó de manera considerable de forma que aunque ocupó la segunda posición en las emisiones de la niña por detrás de [e], la diferencia fue de tan solo una muestra (misma diferencia con respecto a la [i] que quedó una muestra por detrás de la [o]). La [u] ocupó el cuarto lugar superando en número de muestras a la [a] que fue la vocal que menos veces apareció. En el caso del niño, la vocal posterior semialta [o] ocupó el primer lugar quedando en segundo lugar la posterior alta [u] y superando así las vocales posteriores a las anteriores que quedaron en tercer y cuarto lugar ([e, i] respectivamente) sin que la vocal central baja fuera identificada en este periodo. Respecto a las vocales nasales, la frecuencia de aparición fue la misma en ambos niños: [ẽ, ã, õ, ï, ù].

Como conclusión podemos exponer que si bien las vocales anteriores y la central baja fueron las predominantes en las emisiones modales y nasales de ambos niños durante los dos primeros años de vida, el puesto que ocuparon según la frecuencia de aparición cambió con el sujeto (alternándose en el primer puesto la [a] y la [e], y a veces la [i], según el niño del que se

tratara). En el tercer año de vida analizado las preferencias individuales quizás afloraron de manera más clara en las señales modales puesto que la vocal posterior semialta [o] ocupó por primera vez la primera posición en el caso del niño, seguida de la otra vocal posterior [u], sin que esta circunstancia se diera en las emisiones de la niña.

Una vez indicadas las diferencias halladas en el repertorio vocálico de cada uno de los infantes que adquirirían la lengua española, es necesario señalar que tal y como expusimos en el capítulo 5.2.2 también existieron disimilitudes en el número de vocales emitidas por uno y otro niño. De esta forma se extrajeron 188 vocales orales y 489 vocales nasales en la niña, y 112 vocales modales y 335 vocales nasales en el niño.

Se corroboran así en nuestro trabajo las palabras de Ferguson y Farwell (1975), Ferguson (1979), Lieberman (1980), Kent y Bauer (1985) y Ferguson (1986) que afirmaban que hay diferencias individuales importantes tanto en los repertorios de sonidos como en el cómputo de vocales producidas por los niños de la misma experiencia lingüística. Ofrecemos a continuación otros trabajos que confirmaron con anterioridad esta hipótesis.

En el trabajo de Buhr (1980) en el que se llevó a cabo un estudio longitudinal de las vocales emitidas por un sujeto que adquiría inglés americano entre los 0;4 y los 1;4 años de vida, se estableció una comparación entre el repertorio vocálico de este niño y el repertorio de los niños analizados por Irwin (1948) observándose que en este último estudio el número de muestras de [u] antes de los 0;5 o 0;6 años y el de [æ] antes de los 0;11 o 1;0 año, era menor que el hallado en las vocalizaciones del infante de Buhr (1980) en las que además se detectaron estos sonidos desde los 0;4 años (con la particularidad de que los sujetos de ambos estudios estaban aprendiendo la misma lengua ambiente).

Diferencias en el número de vocalizaciones fueron encontradas en el estudio de Camp et al. (1987) en el que después de analizar las vocales de ciento cuarenta bebés entre los 0;4 y el 1;0 año de vida, se observó que las diferencias individuales en cuanto a número de vocalizaciones eran más significativas a los 0;4, 0;5 y 0;6 años.

A pesar de que la mayor parte de las vocales extraídas en el contexto CV de las emisiones de tres niños y tres niñas que adquirirían el inglés americano desde el balbuceo canónico hasta los 3;6 años en el trabajo Davis y MacNeilage (1995) fueron vocales anteriores medias o bajas y centrales, las siguientes diferencias inter-individuales se encontraron: para tres sujetos las vocales medias fueron predominantes, para dos de ellos lo fueron las vocales altas mientras que el último niño prefirió las vocales bajas.



Para Sherkane et al. (2002) la exploración de las habilidades sensoriales y motoras del tracto vocal, como mecanismo que interviene en el desarrollo del habla, implica una enorme variabilidad en la cualidad de las producciones vocálicas en los niños individuales.

Los resultados del trabajo de Van der Stelt et al. (2006) en el que se grabaron las vocales de cuatro niños (incluidos en una base de datos de seis niños y de la que seleccionaron para este estudio a los niños 2, 3, 4 y 5) que adquirían la lengua rusa a las edades de 0;6, 1;0, 1;6 y 2;0 años mostraron que a pesar de las tendencias universales (mayor presencia de vocales centrales y anteriores) diferencias individuales podían observarse. Entre ellas encontraron las siguientes: a los 0;6 años el niño 3 mostró una preferencia mayor que la del resto de sujetos por las vocales posteriores; esta tendencia se repitió en el niño 3 a la edad de 1;0 año frente al niño 2 que presentó más vocales anteriores altas y medias. A los 1;6 años de edad, las vocales centrales prevalecieron en todos los niños aunque estos se distinguieron en la altura vocálica optando por las bajas y las medias los niños 2 y 3 y por las altas el niño 4. A la misma edad, el número de vocales anteriores se redujo en el sujeto 2 y aumentó en el 5 presentando el niño 3 más vocales anteriores que posteriores y decantándose por el patrón contrario el niño 4. A la edad de 2;0 años las vocales posteriores predominaron en el niño 5, se igualó el número de vocales anteriores y posteriores en el niño 2, presentó más vocales posteriores el niño 3 y más anteriores que posteriores el niño 4.

También en otras lenguas encontramos ejemplos de diferencias inter-individuales en las producciones vocálicas características de los bebés. Es el caso del mandarín cuyo estudio fue acometido por Chen (2006) que investigó el repertorio vocálico de dos niños que adquirían dicha lengua entre los 0 y los 2;0 años de edad. Los resultados evidenciaron que a pesar de las similitudes halladas en las vocalizaciones de ambos sujetos como que la vocal anterior alta pasara de ser una vocal bastante frecuente a ocupar los últimos puestos entre el primer y el segundo año de vida, que en el mismo periodo la vocal posterior alta [u] dejara de ser tan poco frecuente y se situara en el segundo y tercer puesto o que fuera la [a] la vocal dominante en todo el estudio excepto en el primer año de vida, existían también diferencias entre los sujetos. Entre ellas encontrábamos la distinta frecuencia de aparición de las cinco vocales del mandarín según el sujeto del que se tratara (ver capítulo segundo). De este modo a los 0;4 años y en el niño varón la frecuencia de aparición fue por este orden de [i, a, o, e, u] mientras que para el sujeto de sexo femenino fue de [i, o, a, u, e]; al 1;0 año la preferencia vocálica en el niño fue por este orden [a, i, o, e, u] mientras que para la niña fue de [a, o, i, e, u]. A los 2;0 años el orden se alteró resultando en la combinación de [a, e, u, o, i] para el sujeto masculino y de [a, u, o, e, i] para el sujeto femenino. Al final del estudio, un mes antes de los 3;0 años de edad, las vocales más frecuentes en el niño fueron por este orden [a, u, e, i, o] y en la niña [a, u, o, e, i].

Observamos pues que las vocales ocuparon distintas posiciones a lo largo de todo el periodo de estudio según el sujeto del que se tratara.

Finalmente, la preferencia por un tipo u otro de vocal en niños pertenecientes a la misma comunidad lingüística se reflejó también, si bien de manera indirecta, en los valores de los formantes. Es lo que ocurrió en el estudio de Boysson-Bardies et al. (1989) en el que tras analizar los dos primeros formantes de las vocales emitidas por veinte niños de 0;10 años procedentes de cuatro lenguas ambiente diferentes (inglesa, francesa, argelina y cantonesa), se observó que existían grandes diferencias en los valores de los formantes de los niños que poseían la misma lengua meta. De esta forma el segundo formante de uno de los niños de habla inglesa era de 2505 Hz frente a los 2805 Hz de uno de sus compañeros; lo mismo sucedió en el caso de la lengua argelina en la que el F2 de uno de los niños fue de 2170 Hz frente a los 2599 Hz de otro de ellos; más llamativo resulta aún, como ya comentamos en el estado de la cuestión, la diferencia entre los valores del segundo formante de uno de los niños que adquiría cantonés cuyo F2 fue de 2061 Hz frente al hallado para uno de sus compañeros que fue de 2566 Hz. Estas disimilitudes, que no pueden ser explicadas por la pertenencia a una lengua ambiente diferente, podrían deberse a la preferencia de los niños pertenecientes a una misma comunidad lingüística por un timbre vocálico determinado que provocaría el aumento o la disminución del promedio de los formantes dependiendo del predominio de una vocal u otra.

Lo mismo puede implicarse del estudio de Gilbert et al. (1997) quienes tras analizar los dos primeros formantes de los sonidos vocálicos de cuatro sujetos que adquirían el inglés americano a los 1;3, 1;6, 1;9, 2;0 y 3;0 años, descubrieron que a los 1;3 años de edad la media del F1 del niño 2 era de 910 Hz y la del niño 4 sin embargo de 1321 Hz. Lo mismo sucedió en el segundo formante y en el mismo periodo temporal en el que la media del F2 del sujeto 1 fue de 2733 Hz y la del sujeto 4 de 2201 Hz. Una vez más la diferencia de valores en los promedios formánticos de niños de la misma edad y que adquirían la misma lengua solo puede explicarse si tenemos en cuenta la inclinación de los niños por un timbre vocálico determinado.

En relación a esta última hipótesis y antes de acabar el capítulo, nos gustaría aclarar que según diferentes estudios estas diferencias individuales, tanto las halladas en el repertorio vocálico como las que tienen que ver con los formantes, no son debidas al diferente sexo de los infantes<sup>9</sup>. Esta es al menos la hipótesis defendida por la mayoría de los autores que tras analizar

---

<sup>9</sup> En nuestro estudio encontramos que los valores de los formantes fueron muy similares en ambos niños para cada una de las vocales resultando la media para algunas vocales mayor incluso en las emisiones del niño, como en caso de la [a], que en las de la niña.

las vocalizaciones de los niños desde el punto de vista acústico y estudiar desde una perspectiva anatómica a los propios sujetos, han concluido que las diferencias de sexo no aparecen sino después del tercer o cuarto año de vida, al menos. Exponemos a continuación, si bien de manera muy breve por no formar parte de nuestro objeto de estudio, algunos de los trabajos en los que se ha investigado a qué edad puede aparecer el dimorfismo sexual.

De acuerdo con Crelin (1973) las diferencias de sexo se reflejan en el tamaño de la laringe a partir de los 3;0 años mientras que en opinión de Eckel et al. (1999) las distinciones relativas al sexo no están presentes en la laringe durante la infancia temprana. Respecto a las cuerdas vocales, el dimorfismo sexual no aparece hasta la edad de 6;0 o 7;0 años según Kazarian et al. (1978) aunque estas diferencias no parecen contribuir a la diferenciación significativa en la  $f_0$  entre hombres y mujeres, distinción que no tiene lugar hasta la pubertad cuando el tamaño laríngeo aumenta el triple en los hombres, junto con aumentos en la longitud de las cuerdas vocales y la diferenciación en su composición (Vorperian y Kent [2007]).

En el trabajo de Bennett (1981), en el que se llevó a cabo un estudio de los cuatro primeros formantes en niños y niñas a la edad de 7;0 y 8;0 años, se encontró una tendencia en los niños hacia la producción de formantes con valores más bajos que los de niñas debido a la mayor longitud faríngea en los sujetos varones.

El estudio de Huber et al. (1999), en el que se analizaron los formantes de la vocal [a] producida bajo la condición de fonación sostenida (2/3 s) en diez sujetos de sexo femenino y diez sujetos de sexo masculino a las edades de 4;0, 6;0, 8;0, 10;0, 12;0, 14;0, 16;0, 18;0 y de entre 20;0 y 30;0 años, concluyó que fuertes diferencias sexuales se daban a partir de los 12;0 años, edad en la que el crecimiento del tracto vocal en los sujetos femeninos parecía estabilizarse haciéndose los formantes de estos sujetos más similares a los de las mujeres adultas, mientras que a esa edad los formantes de los varones descendieron significativamente debido a una gran aceleración del crecimiento del tracto vocal entre los 12;0 y 14;0 años, y continuaron descendiendo, aunque de manera más escalonada, indicando una estabilización en el crecimiento del tracto vocal más tardía que en el caso de los hablantes femeninos.

Para Vorperian y Kent (2007:1515) el dimorfismo sexual aparece a partir de los 4;0 años tal y como exponen en su trabajo en el que recordemos recopilaron quince estudios en los que se analizaron los formantes de niños de diferentes edades que hablaban distintos dialectos del inglés americano:

“[...] A sex difference in the acoustic space begins to emerge even in the data for 4-year-olds, especially for the low vowels in which the F1 values are about 150–200 Hz

lower for males than for females. This difference becomes more pronounced with age, such that progressively less overlap is noted in the vowel quadrilaterals for the two sexes. By the age of 16 years, the quadrilaterals do not overlap. An additional potentially interesting feature is that there is a sex difference in F1 frequency for low vowels across all age groups, with males having lower F1 values”.

Otros autores como Merow y Broadbent (1990) creen que no hay diferencias sexuales hasta la pubertad.

De esta breve revisión de trabajos que indagaron en el reflejo de las distinciones de sexo en las vocalizaciones de los bebés pueden desprenderse dos conclusiones: que ninguno de los autores manifiesta la aparición de las distinciones sexuales antes de los 3;0 años de edad, y la disparidad de conclusiones acerca de la edad en la que este dimorfismo sexual hace su aparición y que es debida según Vorperian y Kent (2007) a los diferentes aspectos de la materia en los que se han centrado los autores.

## 9. CONCLUSIONES.

Una vez presentados los resultados y expuesta la discusión, nos referimos a continuación a las conclusiones que hemos podido extraer a partir del desarrollo del presente estudio:

1. Los valores de los dos primeros formantes de las vocales analizadas en los bebés y niños pequeños superan con creces a los de los adultos. Esta diferencia fue encontrada tras comparar los valores formánticos obtenidos en nuestros dos sujetos con los valores medios establecidos para los hablantes adultos (además ha sido hallada también en lenguas diferentes al idioma español). La razón principal de esta disimilitud reside en la configuración anatómica particular del tracto vocal del bebé y en las características de sus distintas estructuras que causan que el aparato fonador del bebé y del niño pequeño sean fisiológica y funcionalmente muy diferentes al del adulto.
2. La mayoría de los sonidos vocálicos producidos por un bebé en los primeros meses de vida son articulados en la parte central y anterior de la cavidad bucal y poseen, al menos al principio, una altura media y baja (estableciéndose de nuevo una coincidencia con los resultados obtenidos en otras lenguas). Esta conclusión se desprende de la ubicación de las vocales en las cartas de formantes del apartado 6.1, así como de los resultados arrojados por del test de percepción en el punto 6.2 donde se observó una mayor frecuencia de estas vocales en ambos niños durante al menos el primer año de vida. Aunque entre las causas volvemos a encontrar condicionantes de tipo anatómico (como la reducida movilidad de la lengua debido a su gran tamaño), a estos debe sumársele el escaso control articulatorio de los diferentes órganos del tracto vocal siendo este control necesario para la producción de otros timbres vocálicos. Un ejemplo lo constituye el deficiente control articulatorio que posee el bebé para la articulación de la [o] y la [u], puesto que la producción de estas vocales implica además del movimiento mandibular y el de la lengua, el control de los labios, frente al suficiente control articulatorio que presenta el bebé para la articulación de las vocales centrales y anteriores medias y bajas.
3. Dentro del ángulo posterior, la vocal [u] es la última en aparecer y no lo hace al menos hasta que el bebé ha cumplido los seis meses de vida según los resultados obtenidos en el test de percepción. Los motivos de la tardanza

residen en que al hecho de tratarse de una vocal posterior para cuya articulación es necesario el movimiento coordinado de distintos órganos articulatorios, que como para la vocal [o] dificultan su producción y favorecen su escasez, a la pronunciación de la [u] se le suma una mayor dificultad del gesto articulatorio que se traduce como hemos visto en: una mayor protusión de los labios (mayor que para la [o]), una constricción perfectamente definida en la cavidad oral y una coordinación de todos los órganos implicados mucho más precisa.

4. Las producciones vocálicas de los bebés están determinadas, al menos en parte, por la lengua ambiente en la que estos están inmersos. Así lo confirman en nuestro estudio la elevada frecuencia de la vocal [e] en ambos niños a lo largo de todo el periodo observado y la exigua presencia de la vocal [u] en las vocalizaciones de ambos sujetos durante todo el tiempo analizado, circunstancias ambas que coinciden con la frecuencia vocálica computada para la lengua española. Estos resultados parecen pues reflejar la intervención de otros componentes, además del de la configuración anatómica y el del nivel de control motor del sujeto, en la adquisición de una lengua materna por parte del niño.
5. El bebé es capaz de producir cualquier vocal del español una vez que se superan los primeros siete u ocho meses de vida. Si observamos las cartas de formantes y comparamos la distribución de las señales con los resultados obtenidos en el test de percepción, podemos concluir que a los siete meses de edad en el caso de la niña y a los ocho en el del niño, el bebé es capaz de articular todas las vocales del español (a estas edades ya han aparecido al menos en una ocasión todas las vocales de la lengua española).
6. Los valores de los dos primeros formantes son muy similares en ambos sujetos. Esta similitud se dio al equiparar los valores obtenidos en los dos primeros formantes de las vocales de cada uno de los sujetos después de que estos se calcularan sobre las vocales de un mismo timbre (evitando así que la predilección por una determinada vocal oscureciera los promedios mensuales y anuales).
7. Se crearon campos de dispersión alrededor de todas las vocales analizadas debido a la articulación de los niños de una misma vocal con pequeñas diferencias en los valores formánticos. Este fenómeno puede observarse en las

cartas de formantes realizadas y puede desprenderse de la presencia de la desviación estándar (si bien esta no fue muy elevada).

8. La variabilidad de los valores formánticos dentro de una vocal puede deberse a que el tracto vocal del bebé no está aún configurado como el del adulto y sufre cambios continuos tanto en su longitud como en su anchura y volumen. Estos cambios, al ser constantes, pueden provocar que una vocal presente unos valores determinados para el F1 y el F2 en una sesión y unos valores ligeramente distintos en una sesión diferente porque en ese periodo de tiempo se han producido ajustes anatómicos. Por otro lado, el aumento del movimiento de algunos órganos articulatorios (como el de la lengua), la falta de exactitud en las posiciones a la hora de articular una vocal determinada así como la dificultad de repetir una tarea con el mismo gesto articulatorio, facilitan en el bebé la variabilidad de los valores formánticos.
9. Los valores de los dos primeros formantes no sufren ninguna alteración significativa durante el primer año de vida a pesar de los cambios dramáticos que experimenta el tracto vocal en este periodo temporal. Este hecho puede inferirse tras contemplar la evolución de los promedios mensuales a partir de los gráficos y de las tablas que recogen las medias mensuales y anuales de los dos primeros formantes en cada uno de los niños, y puede confirmarse a partir de los resultados obtenidos después de haber realizado el cálculo estadístico del análisis de varianza (ANOVA). Entre las razones que se aducen para explicar que los valores formánticos no desciendan a pesar del crecimiento del tracto vocal que tiene lugar en este periodo de tiempo (teoría acústica de Fant, 1960), está la necesidad de considerar otras variables en el crecimiento de dicho tracto vocal como la variable de la longitud, la de la anchura y la del volumen sin limitarse a la longitud del mismo, así como la obligación de contemplar dicho crecimiento como un crecimiento no lineal en el que las distintas estructuras del tracto vocal presentan diferentes ritmos y velocidades de crecimiento (crecimiento no uniforme del tracto vocal).
10. Existe una variedad inter-individual tanto en el número de vocales extraídas a lo largo del periodo de estudio como en su cualidad vocálica. Así lo confirma el recuento final de las vocales analizadas en las emisiones de ambos bebés y en el que los sonidos vocálicos de la niña, tanto modales como nasales, superaron notoriamente a los extraídos en el niño. Por otro lado, y a pesar de la tendencia

de ambos bebés hacia la producción de vocales centrales y anteriores bajas y medias, existieron preferencias individuales hallándose diferencias en la frecuencia de aparición de las distintas vocales en determinados periodos temporales según el sujeto.



## 10. EPÍLOGO.

Antes de finalizar esta tesis, nos gustaría dejar sobre la mesa cuestiones tales como de qué forma dicho trabajo ha podido contribuir a la literatura acerca del estudio acústico de las vocales del bebé que aprende una lengua, en qué medida sus datos son representativos, qué otras variables sería necesario estudiar para tener una visión más completa de la acústica vocálica infantil o cuáles son sus posibles aplicaciones.

Por lo que respecta a la contribución del mismo en el limitado panorama internacional sobre la investigación acústica de los sonidos vocálicos emitidos por bebés y niños pequeños, este trabajo ha ayudado a comprender mejor el inventario vocálico de los pequeños no solo mediante una descripción detallada del mismo sino mediante la indagación en cuestiones tales como el porqué de la presencia de determinadas vocales y no de otras en los meses inmediatos al nacimiento, las causas de la aparición tardía de algunas de ellas, o los motivos por los que los valores formánticos no sufren cambios significativos durante el primer año de vida. Además, al abordar estas cuestiones desde una triple perspectiva, a saber, acústica, anatómica y perceptiva, se consigue un entendimiento más completo del desarrollo vocálico en la infancia más temprana. Por otro lado, el que se hayan corroborado en el idioma español todas las hipótesis de partida, confirmadas con anterioridad en otras lenguas, dota a esta investigación de un carácter de universalidad que se traduce en un paso hacia adelante en la comprensión del desarrollo vocálico en bebés y niños pequeños independientemente de la lengua que estos hablen.

Dentro del reducido número de trabajos que estudian el componente fonético a lo largo del proceso de adquisición de la lengua española, debemos subrayar que se trata del primer trabajo sobre el análisis acústico de los dos primeros formantes de las emisiones vocálicas de bebés grabados y estudiados durante tan extenso periodo de tiempo. No existe otro estudio en el idioma español con el que pueda compararse por lo que su contribución es notoria ya que constituye los cimientos de cualquier otra investigación posterior en el campo del español al mismo tiempo que abre y facilita el camino a futuros investigadores en esta materia.

Finalmente y en cuanto a los marcos más teóricos en materia de adquisición, este estudio es de gran utilidad ya que según Dinnsen (1992) la información respecto a la variación en los inventarios fonéticos de los sistemas fonológicos en desarrollo es un paso necesario para establecer cualquier teoría de adquisición.

En relación a la representatividad de este trabajo, queda patente que el reducido número de sujetos en los que se basó este estudio no favorece la misma ya que los resultados sobre el repertorio fonético así como los del análisis y la evolución de los dos primeros formantes son

solo los presentados por dos individuos. Sin embargo, esta circunstancia sitúa nuestra investigación dentro de la tónica general de los estudios longitudinales en los que por sus propias características (explicadas en el capítulo quinto de esta tesis) resulta realmente complicado la observación y el estudio de un mayor número de sujetos debido a las múltiples dificultades que conlleva el realizar grabaciones periódicas, en nuestro caso semanales, durante un considerable periodo de tiempo, de casi tres años en nuestra investigación (se oponen a estos trabajos los estudios transversales en los que se analizan las emisiones de un mayor número de sujetos aunque en un periodo de tiempo mucho más limitado). Estas circunstancias hacen que la representatividad radique en otros aspectos como son por un lado el elevado número de grabaciones y de muestras para analizar y estudiar en cada uno de los sujetos, y por otro en el hecho de que se trate, como ya hemos comentado, del primer estudio de estas características realizado en el idioma español.

Si hablamos de la utilidad de esta monografía encontramos que las aplicaciones son múltiples y variadas. Por un lado, y quizás entre las más evidentes, está su aprovechamiento en la investigación sobre el desarrollo del habla por la luz que este estudio arroja acerca del componente fonético, y más concretamente acústico y vocálico, de las emisiones de bebés y niños pequeños que están adquiriendo una lengua. Por otro lado, encontraríamos también aplicaciones clínicas ya que al comprender el estudio un periodo de tiempo tan inmediato al nacimiento, este puede servir de gran ayuda para detectar desórdenes del habla. En este sentido, Vorperian y Kent (2007) aluden a la utilidad de este tipo de trabajos a la hora de localizar problemas como la disartria, la sordera y otros desórdenes del habla (como la reducida inteligibilidad), desórdenes que se descubrirían tras comparar por ejemplo los datos del tamaño del espacio acústico de estos niños con el hallado en niños con un desarrollo normal. Por último, el análisis formántico de las vocales puede ser de gran ayuda para el desarrollo, en el campo de la ingeniería del sonido, de sistemas de síntesis y reconocimiento de voz del habla infantil.

A pesar de la considerable cantidad de información que aporta un estudio de tipo acústico en el que se analizan los dos primeros formantes de las vocales de bebés y niños pequeños, creemos que dicha información podría completarse con investigaciones futuras mediante la observación de otras variables. Dentro de un estudio puramente acústico sería interesante el análisis del tercer formante así como el del ancho de banda formántico y el de la frecuencia fundamental. Por otro lado, la integración de la perspectiva anatómica, desde el desarrollo neurológico a la medición de algunos índices de crecimiento, sería de gran ayuda en la comprensión de los resultados acústicos debido a la estrecha relación que guardan estos dos componentes en la producción del habla.

Por último, la escasez de trabajos centrados en el análisis acústico de las emisiones vocálicas de bebés y niños pequeños nos lleva a insistir en la necesidad de estudios futuros cuyas conclusiones contribuyan a un mejor entendimiento del habla infantil y para los que esperamos haber abierto camino y despertado el interés suficiente con esta monografía.

## BIBLIOGRAFÍA CITADA.

- ABBS J. H. y GRACO V. L. (1984): "Control of complex motor gestures: orofacial muscle responses to load perturbations of lip during speech", *J. Neurophysiol* 51, 705-723.
- ABRAMSON, A. S. y LISKER, L. (1967): "Discriminability along the voicing continuum: Cross-language tests", *Status Report on Speech Research* 11, 17-22 (Haskins Laboratories, New York).
- ABRAMSON, A. S. y LISKER, L. (1967): Voice Timing: Cross-language Experiments in Identification and Discrimination, *Journal of the Acoustical Society of America* 44, 377.
- AINSWORTH, W. A. (1975): "Intrinsic and extrinsic factors in vowel judgments". En: FANT, G. y TATHAM, M. (eds.). *Auditory Analysis and Perception of Speech*, (Academic London), 103-113.
- ALARCOS LLORACH, E. (1961): *Fonología española*, Madrid, Gredos, 3ª ed.
- ALARCOS LLORACH, E., BOREL-MAISONNY, S., HECAEN, H., MANDIN, D. y GUIRAUD, P. (1968): *La adquisición del lenguaje por el niño. Desórdenes, funciones secundarias y representaciones gráficas del lenguaje*, vol. 3, Buenos Aires, Nueva Visión.
- AMANO, S., KATO, K. y KONDO, T. (2002): Development of Japanese infant speech database and speaking rate analysis, *Proc. of the 7<sup>th</sup> International Conference on Spoken Language Processing (ICSLP)* 1, 317-320.
- AMAYREH, M. M. y DYSON, A. T. (2000): "Phonetic inventories of young Arabic-speaking children", *Clinical linguistics & phonetics* 14(3), 193-215.
- ASLIN, R. N., PISONI, D. B., HENNESSY, B. L. y PERRY, A. J. (1981): "Discrimination of voice onset time by human infants: New findings and implications for the effect of early experience", *Child development* 52, 1135-1145.
- ASSMAN, P. F. y KATZ, W. F. (2000): "Time-varying spectral change in the vowels of children and adults", *Journal of the Acoustical Society of America* 108(4), 1856-1866.
- ATAL, B. S. (1970): "Determination of the vocal tract shape directly from the speech wave", *Journal of the Acoustical Society of America* 47, 65 (A).

- ATAL, B. S., CHANG, J. J., MATHEWS, M. V. y TUKEY, J. W. (1978): "Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer sorting technique", *Journal of the Acoustical Society of America* 63, 1535-1555.
- BAHRICK, L. E. y PICKENS, J. N. (1988): "Classification of bimodal English and Spanish language passages by infants", *Infant Behavior and Development* 11, 277-296.
- BAILEY, P. y SNOWLING, M. (2002): "Auditory processing and the development of language and literacy", *British Medical Bulletin* 63, 135-146.
- BAKEN, R. J. y ORLIKOFF, R. F. (2000): *Clinical Measurement of Speech & Voice*, Singular Publishing Group.
- BATEMAN, W. G. (1915): "Two children's progress in speech", *J. Educ., Psicol.* 6, 475-493.
- BEDDOR, P. S. y STRANGE, W. (1982): "Cross-language study of the oral-nasal distinction", *Journal of the Acoustical Society of America* 71, 1551-1561.
- BEERS, M. (1995): The phonology of normally developing and language-impaired children, Doctoral Dissertation, University of Amsterdam, Amsterdam.
- BENDERS, T. (2011): Dutch infants discrimination of /a:/ and /ɑ/: Development into the second year of life, The International Child Phonology conference 15-18 June 2011, The University of York.
- BENQUEREL, A. P. y LAFARGUE, A. (1981): "Perception of vowel nasalization in French", *Journal of Phonetic* 9, 309-321.
- BENNER, A., GRENON, I. y ESLING, H. (2007): Infant's phonetic acquisition of voice quality parameters in the first year of life, Department of Linguistics, University of Victoria, Victoria, BC, Canada (pp. 2073-2076).
- BENNET, S. (1981): "Vowel formant frequency characteristics of preadolescent males and females", *Journal of the Acoustical Society of America* 69, 231-238.
- BERG, K. M. y BOSWELL, A. E. (2000): "Noise increment detection in children 1 to 3 years of age", *Percep. Psychophys*, 62(4), 868-873.
- BERNSTEIN, N. A. (1967). *The Coordination and Regulation of Movements*. Oxford: Pergamon Press.

- BERNSTEIN, N. y LUBEROFF A. (1984): “Cues to post-vocalic voicing in mother–child speech”, *Journal of Phonetics* 12, 285–289.
- BEST, C. T. y MCROBERTS, G. W. (1989): Phonological influence on infant’s perception of two nonnative speech contrasts, Paper presented at the meeting of the Society for Research in Child Development, Kansas City, MD.
- BEST, C.T., SINGH, L., BOUCHARD, J., CONNELLY, G., COOK, A. y FABER, A. (1997): “Developmental changes in infants’ discrimination of non-native vowels that assimilate to two native categories”. En: Meeting of the Society of Research in Child Development, Providence, RI, 3–6 April 1997.
- BICKLEY, C.A. (1989): Acoustic evidence for the development of speech (RLE Technical Report 548): Cambridge: MIT.
- BICKLEY, C., LINDBLOM, B. y ROGUÉ, L. (1986): Acoustic measures of rhythm in infant’s babbling. Paper presented at the Proceedings of the Twelfth International Congress on Acoustics, Toronto.
- BIRNHOLTZ, J.C. y BENACERRAF, B.R. (1983): “The development of human fetal hearing”, *Science* 222,516-518.
- BISCHOFF, D. M. (1976): “Secondary acoustic characteristics and vowel identification”, *Journal of the Acoustical Society of America* 51, 648-651.
- BLAKEMORE, S. y CHOUDHURY, S. (2006): “Development of the adolescent brain: Implications for executive function and social cognition”, *Journal of Child Psychology and Psychiatry* 47 (3-4), 296-312.
- BLOOM, L. (1970): *Language Development: Form and Function in Emerging Grammars*, Cambridge, MA: MIT Press.
- BLOOM, K. (1988): “Quality of adult vocalizations affects the quality of infant vocalizations”, *Journal of Child Language* 15, 469–480.
- BLOOM, K. (1998): “The missing link's missing link. Syllabic vocalizations at 3 months of age”, Open Peer Commentary MacNeilage, Evolution of speech, *Behavioral and Brain Sciences* 21, 514-515.
- BLOOM, K., RUSSELL, A. y WASSENBERG, K. (1987): “Turn taking affects the quality of infant vocalizations”, *Journal of Child Language* 14, 211–227.

- BOË, L. J. (1999): Modeling the growth of the vocal tract vowel spaces of newly-born infants and adults. En: *Proceedings of the XIVth international congress of phonetic sciences* (pp. 2501–2504), San Francisco, USA.
- BOË, L. J. y ABRY, C. (1986): “Nomogrammes et systèmes vocaliques. Actes des 15èmes journées d’étude sur la parole”, *Société Française d’Acoustique*, 303–306.
- BOË, L. J., PERRIER, P. y BAILLY, G. (1992): “The geometric vocal tract variables controlled for vowel production: proposal for constraining acoustic-to-articulatory inversion”, *Journal of phonetics* 20, 27-38.
- BOERSMA, H., van der LINDEN, F. P. G. M. y PRAHL-ANDERSEN, B. (1979): Craniofacial development. En: PRAHL-ANDERSEN, B., KOWALSKY, C. J. y HEYDENDAEL, P. (eds.): *Interdisciplinary Study of Growth and Development* (Academic, New York).
- BOHN, W. W. (1914): “First steps in verbal expression”, *Ped. Sem.* 21, 578-595.
- BOHN, O-S, y FLEGE, J. E. (1990): “Interlingual identification and the role of foreign language experience in L2 vowel perception”, *Appl. Psycholinguistics* 11, 303-328.
- BOLIEK, C. A., HIXON, T. J., WATSON, P. J. y MORGAN, W. J. (1996): “Vocalization and breathing during the first year of life”, *Journal of Voice* 10, 1-22.
- BOLIEK, C.A., HIXON, T. J., WATSON, P. J. y MORGAN, W. J. (1997): “Vocalization and breathing during the second and third years of life”, *Journal of Voice* 11, 373–390.
- BOND, Z. S., PETROSINO, L. y DEAN, R. (1982): “The emergence of vowels: 17 to 26 months”, *Journal of Phonetics* 10, 417-422.
- BORZORNE de MANRIQUE, A. M. (1977): “On the recognition of isolated Spanish vowels”. En: HOLLIEN, H. y HOLLIEN, P. (eds.). *Current Issues in the Phonetic Sciences. Proceedings of the IPS-77 Congress*, Amsterdam, John Benjamins.
- BORZORNE de MANRIQUE, A. M. y GURLEKIAN, J. A. (1977): *Study of sounds*, XVIII.
- BOSCH, L. (2004). *Evaluación fonológica del habla infantil*, Madrid, Elsevier.
- BOSCH, L. y SEBASTIÁN-GALLÉS, N. (2003a): “Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life”, *Language and Speech* 46, 217–243.

- BOSCH, L. y SEBASTIÁN-GALLÉS, N. (2009): “Developmental changes in the discrimination of vowel contrast in bilingual infants”. En: COHEN, J., MCALISTER, K., ROLTAD, K. y MACSWAN, J. (eds.) ISB4. *Proceedings of the 4 The International Symposium on Bilingualism*. Somerville, MA: Cascadilla Press (pp. 354-363).
- BOSMA, J. F. (1963): “Maturation of function of the oral and pharyngeal region”, *Journal of Orthodontics* 49(2), 94-104.
- BOSMA, J. F. (1975a): “Anatomic and Physiologic Development of the Speech Apparatus”. En: TOWER, D. B. (ed.). *The Nervous System*, vol. 3: *Human Communication and its disorders*, Raven Press, New York, vol. 3, 469-481.
- BOSMA, J. F. (1975b): *Introduction In Symposium on Development of Upper Respiratory Anatomy and Function: Implications Regarding Sudden and Unexpected Infant Death*, edited by J. F. Bosma and J. Showacre (Government Printing Office, Washington, D. C.), p. 5-44.
- BOSMA, J. F. (1976): Discussion of the paper, “Postnatal development of the basicranium and vocal tract in man”, by J. T. Laitman and E. S. Crelin, in *Symposium on the Development of the Basicranium*, edited by J. Bosma. DHEW Publication No. 76-989 (PHS- NIH, Bethesda, MD).
- BOSMA, J. F. (1985): “Postnatal ontogeny of performance of the pharynx, larynx and mouth”, *Am. Rev. Respir. Dis.* 131, S10-S15.
- BOSMA, J. F. (1986): *Anatomy of the infant head*. Johns Hopkins University Press, Baltimore.
- BOYSSON-BARDIES, B., SAGART, L. y BACRI, N. (1980): “Phonetic analysis of late babbling: a case study of French child”, *Journal of child Language* 8, 511-524.
- BOYSSON-BARDIES, B., SAGART, L. y DURAND, C. (1984): “Discernible differences in the babbling of infants according to target-language”, *Journal of Child language* II, 1-15.
- BOYSSON-BARDIES, B., SAGART, L. y DURAND, C. (1986): “Acoustic investigation of crosslinguistic variability in babbling”. En: LINDBLOM, B. y ZETTERSTROM, R. (eds.). *Precursors of early speech* (pp. 113-126): New York: Stockton Press.



- BOYSSON-BARDIES, B., HALLE, P., SAGART, L. y DURAND, C. (1989): "A cross-linguistic investigation of vowel formants in babbling", *Journal of Child Language* 16, 1-18.
- BOYSSON-BARDIES, B. y VIHMAN, M. M. (1991): "Adaptation to language: Evidence from babbling and first words in four languages", *Language* 67, 297-319.
- BOYSSON-BARDIES, B., VIHMAN, M. M., ROUG-HELLICIUS, L., DURAND, C., LANDBERG, I. y ARAO, F. (1992): "A cross-Linguistic Phonetic Study". En: FERGUSON, C., MENN, L. y STOEL-GAMMON, C. (eds.). *Phonological development. Models, Research, Implications*, Timonium, Maryland, pp. 369-391.
- BOYSSON-BARDIES, B., VIHMAN, M. M., ROGU-HELLICHIUS, L., DURAND, C., LANDBERG, I. y ARAO, F. (1992): Material evidence of infant selection from the target language: a cross-linguistic phonetic study. En: FERGUSON, C., MENN, L. y STOEL-GAMMON, C. (eds.). *Phonological development: Models, research, implications* (pp. 369-391): Timonium, MD: York Press.
- BRUNER, J. S. (1973): "Organization of early skilled action", *Child Development* 44, 1-11.
- BRYAN, A. C. y WOHL, M. E. B. (1986): Respiratory mechanics in children. En: MACKLEM, P. T. y MEAD, J. (eds.). *Handbook of physiology*, vol. 3, The respiratory system (pp. 179-191), Bethesda, MD: American Physiological Society.
- BUCHAILLARD, S., PERRIER, P. y PAYAN, Y. (2009): "A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning", *Journal of the Acoustical Society of America* 126 (4), 2033-2051.
- BÜHLER, K. (1930): *The Mental Development of the Child*, Londres, Kegan Paul.
- BUHR, R. D. (1980): "The emergence of vowels in an infant", *Journal of speech and hearing research* 23, 73-94.
- BURNHAM, D. K. (1986): "Developmental loss of speech perception: Exposure to and experience with a first language", *Applied psycholinguistics* 7, 207-240.
- BUSBY, P. A. y PLANT, G. L. (1995): "Formant frequency values of vowels produced by preadolescent boys and girls", *Journal of the Acoustical Society of America* 97, 2603-2606.

- BUTTERFIELD, E. y CAIRNS, G. (1974): "Discussion summary- Infant perception research". En: SCHIEFELBUSCH, R. y LLOYD, L. (eds.). *Language Perspectives, Acquisition, Retardation, and Intervention*. Baltimore, Md.: University Park Press), pp. 75-102.
- CALLAN, D. E., KENT, R. D., GUENTHER, F. H. y VORPERIAN, H. K. (2000): "An auditory-feedback-based neural network, model of speech production that is robust to developmental changes in the size and shape of the articulatory system", *Journal of Speech, Language and Hearing Research* 43, 721-736.
- CAMP, B. W., BURGESS, D., MORGAN, L. J. y ZERBE, G. (1987): "A longitudinal study of infant vocalization in the first year", *Journal of Pediatric Psychology* 12(3), 321-331.
- CARLSÖÖ, S. y LEIJON, G. (1960): "A radiographic study of the position of the hyo-laryngeal complex in relation to the skull and the cervical column in man", *Trans. R. Sch. Dent. Umea* (Stockh) 5, 13-35.
- CARRÉ, R., LINDBLOM, B. y MACNEILAGE, P. F. (1995): Acoustic factors in the evolution of the vocal tract. (Translation) *C. R. Academie des Sciences Paris*, t 320, Series IIb, 471-476.
- CASAL, C., DOMÍNGUEZ, C., FERNÁNDEZ, A., SARGET, R., MARTÍNEZ-CELDRÁN, E., SENTÍS-VILALTA, J. y GAY-ESCODA, C. (2002). "Spectrographic Measures of the Speech of Young Children with Cleft Lip and Cleft Palate", *Folia Phoniatr, Logop.* 54(5), 247-257.
- CATFORD, I. (1981): "Observations on the recent history of vowel classification", Towards a History of Phonetics. Papers contributed in honour of David Abercrombie, R.E. Asher & E.J.T. Henderson (eds.), Edinburgh: Edinburgh University Press, pp. 19-32.
- CHAMPNEYS, F. H. (1881): "Notes on an infant", *Mind* 6, 104-107.
- CHAN, C. (2001): A Consonant-vowel co-occurrence in early speech development of normal Cantonese-speaking children, Undergraduated thesis, University of Hong Kong.
- CHANG, H.-W. y TREHUB, S. E. ( 1977 ): "Infants' perception of temporal grouping in auditory patterns", *Child Development* 48(4), 1666-1670 .
- CHARPENTIER, F. (1984): "Determination of the vocal tract shape from the formants by analysis of the articulatory-to-acoustic nonlinearities", *Speech Communication* 3, 291-308.

- CHEN, L-M (2006): "Developmental changes in the acoustic vowel space of Mandarin-learning children before 2 years of age", *Journal of the Acoustic Society of America* 119(5), pt. 2, p. 3421.
- CHEN, L-M (2007): Longitudinal Development of Vowels before 4 Years of Age. *Proceedings of the 20th Symposium of the Acoustical society of the Republic of China*.
- CHEN, L-M y KENT, R. D. (2005): "Consonant-vowel co-occurrence patterns Mandarin-learning infants", *Journal of Child Language* 32, 507-534.
- CHEN, L-M y KENT, R. D. (2010): "Segmental production in Mandarin-learning infants", *Journal of Child Language* 37, 341-371.
- CHEN, M. (1997): "Acoustic correlates of English and French nasalized vowels", *Journal of the Acoustical Society of America* 102 (4), 2360-2370.
- CHENG, R. L. (1966): "Mandarin phonological structure", *Journal of Linguistics* 2 (2), 135-262.
- CHENG, C. C. (1973): *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.
- CHENG, H. Y., MURDOCH, B. E., GOOZEE, J. V. y SCOTT, D. (2007): "Physiological development of tongue-jaw coordination from childhood to adulthood", *Journal of Speech, Language, and Hearing Research* 50, 352-60.
- CHI LIN, Y. (in press). *Vowel development of a 3 year old Taiwanese child*, National Chung Cheng University, 1-20.
- CHIBA, T. y KAJIYAMA, M. (1941): *The Vowel: Its Nature and Structure* (Tokyo-Kaiseikan, Tokyo, Japan).
- CHIEDE: Corpus de Habla Infantil Espontánea del Español, Laboratorio de Lingüística Informática de la Universidad Autónoma de Madrid.
- CHILDERS, D. G. y WU, K. (1991): "Gender recognition from speech. Part II: Fine analysis", *Journal of the Acoustical Society of America* 90, 1841-1856.
- CHISTOVICH, L. A. y LUBLINSKAYA, V. V. (1979): "The 'center of gravity' effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli", *Hear. Res.* 1, 185-195.

- CHISTOVICH, L. A., SHEIKIN, R. L. y LUBLINSKAYA, V. V. (1979): "Centres of gravity' and spectral peaks as the determinants of vowel quality". En: LINDBLOM, B. y OHMAN, S. (eds.). *Frontiers of Speech Communication Research*, (Academic, London) pp. 143-157.
- CLARKSON, M. G., CLIFTON, R. K. y PERRIS, E. E. (1988): "Infant timbre perception: Discrimination of spectral envelopes", *Perception & Psychophysics* 43, 15-20.
- CLEMENT, C. J. (2004): Early speech development in normally hearing and hearing impaired infants. En: CLEMENT, C. J. (ed.). *Development of vocalizations in deaf and normally hearing infants*, Academisch Proefschrift, 2004.
- CLEMENT, C. J. y WIJNEN, F. (1994): "Acquisition of vowel contrast in Dutch", *Journal of Speech and Hearing Research* 37(1), 83-89.
- CLOPPER C. G., PISONI D. B. y JONG de, K. (2005): "Acoustic characteristics of the vowel systems of six regional varieties of American English", *Journal of the Acoustical Society of America* 118, 1661-1676.
- CONNAGHAN, K. P., MOORE, C. A., CHRISTOPHER, A. y HIGASHAKAWA, M. (2004): "Respiratory kinematics during vocalization and nonspeech respiration in children from 9 to 48 months", *Journal of Speech, Language, and Hearing Research* 47, 70-84.
- COOPER J. S. y COOPER R. (2000): Slower speaking rate increases attention to infant-directed speech in 1- and 4-month-olds. *Paper Presented at International Society of Infant Studies (ICIS2000)*, Brighton, England, UK.
- COOPER, F. S., DELATTRE, P. C., LIBERMAN, A. M., BORST, J. M. y GERSTMAN, L. J. (1952): "Some experiments on the perception of synthetic speech sounds", *Journal of the Acoustical Society of America* 24, 597-606.
- COWAN, N. y MORSE, P. A. (1986): "The use of auditory and phonetic memory in vowel discrimination", *Journal of the Acoustical Society of America* 79, 500-507.
- CRELIN, E. S. (1973): *Functional Anatomy of the Newborn*. New Haven: Yale University Press.
- CRELIN, E. S. (1976): "Development of the upper respiratory system", *Clin. Symp.* 28, 1-30.
- CRELIN E. S. (1987): *The human vocal tract: anatomy, function, development, and evolution*. New York: Vantage Press.

- CRUTTENDEN, A. (1970): "A phonetic study of babbling", *British Journal of Disorders of communications* 5, 110-117.
- CUTTING, J. E., ROSNER, B. S. y FOARD, C. F. (1975): Rise time in nonlinguistic sounds and models of speech perception. Haskins Laboratories Status Report on Speech Research SR-41, 71-93.
- DARWIN, C. (1877): "A biographical sketch of an infant", *Mind* 2(7), 285-294.
- DAVIS, B. L. y MACNEILAGE, P. F. (1990): "Acquisition of correct vowel production: a quantitative case of study", *Journal of Speech and Hearing Research* 33, 16-27.
- DAVIS, B. L. y MACNEILAGE, P. F. (1994): "Organization of babbling: a case study", *Language and Speech* 37, 341-355.
- DAVIS, B. L. y MACNEILAGE, P. F. (1995): "The articulatory basis of babbling", *Journal of Speech and Hearing Research* 38, 1199-1211.
- DAVIS, B. L., MACNEILAGE, P. F. y MATYEAR, C. L. (2002): "Organization of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production", *Phonetica* 59, 75-107.
- DAVIS, E. R. y DAVIS, B. O. (2002): "Early Sound Patterns in the Speech of two Brazilian Portuguese speakers", *Language and Speech* 45(2), 179-204.
- DAWN WARNER-CZYZ, A. (2005): The development of accuracy in early speech acquisition: relative contributions of production and auditory perceptivo factors, University of Texas, Austin.
- DEARBORN, G. V. N. (1910): *Moto-sensory development: Observations on the first three years of a child*, Baltimore: Warwick and Cork, 215.
- DECASPER, A. J. y FIFER, W. P. (1980): "Of human bonding: Newborns prefer their mothers' voices", *Science* 208, 1174 –1176.
- DECASPER, A. J. y SPENCE, M. J. (1986): "Prenatal maternal speech influences newborns' perception of speech sounds", *Infant Behaviour and Development* 9, 133-150.
- DECASPER, A. J., LECANUET, J. P., BUSNEL, M-C, GRANIER-DEFERRE, C. y MAUGEAIS, R. (1994): "Fetal reactions to recurrent maternal speech", *Infant Behavior and Development* 17, 159 – 164.

- DEHAENE-LAMBERTZ, G. y HOUSTON, D. (1998): "Language discrimination response latencies in two-month-old infants", *Language and Speech* 41, 1, 21-43.
- DEJARDINS, R. N., TRAINOR, L. J. (1998): "Fundamental frequency influences vowel discrimination in 6-month-old infants", *Can.Acoust.Proc.* 26, 96-97.
- DELATTRE, P. (1948): "Un triangle acoustique des voyelles orales du français", *French Review* (FR) 21, 477-484.
- DELATTRE, P. (1954): "Les attributs acoustiques de la nasalité vocalique et consonantique", *Studia Linguist.* 8, 103-109.
- DELATTRE, P. (1965): Comparing the phonetic features of English, German, Spanish and French, Heidelberg, Julius Groos Verlag.
- DELATTRE, P. (1968): "Divergences entre la nasalité vocalique et consonantique en français", *Word* 24, 64-72.
- DELATTRE, P., LIBERMAN, A. M. y COOPER, F. S. (1951): "Voyelles synthétiques à deux formantes et voyelles cardinales", *Le Maître Phonétique* 96, 30-36.
- DELATTRE, P., LIBERMAN, A. M., COOPER F. S. y GERSTMAN, L. J. (1952): "An experimental study of the acoustic determinants of vowel color; Observations on one- and two-formant vowels synthesized from spectrographic patterns", *Word* 8(3), 192-210.
- DELGUTTE, B. (1997): "Auditory neural Processing of speech". En: HARDCASTLE, W. J. y LAVER, J. (eds.). *The Handbook of Phonetic Sciences*, Blackwell: Oxford, p. 507-538.
- DENES, P. y PRINSON, E. (1963): The speech chain. Bell Telephone Laboratories, Baltimore: Waverly.
- DEPAOLIS, R. A., VIHMAN, M. M. y KUNNARI, S. (2008): "Prosody in production at the onset of word use: A cross- linguistic study", *Journal of Phonetics* 36, 406-422.
- DESIMONI, F. G. (1974a): "Influence of vowel environment on the duration of consonants in the speech of three, six, and nine-year old children", *Journal of the Acoustical Society of America* 55, 360-361.

- DESIMONI, F. G. (1974b): "Influence of consonant environment on duration of vowels in the speech of three, six, and nine-year old children", *Journal of the Acoustical Society of America* 55, 362-363.
- DESIMONI, F. G. (1974c): "Influence of utterance length upon bilabial closure for /p/ in three, six, and nine-year old children", *Journal of the Acoustical Society of America* 55, 1353-1354.
- DIEHL, R. L., BUCHWALD McCUSKER, S. y CHAPMAN, L. S. (1981): "Perceiving vowels in isolation and in consonantal context", *Journal of the Acoustical Society of America* 69(1), 239- 247.
- DIETRICH, C., SWINGLEY, D. y WERKER, J. (2007): "Native language governs interpretation of salient speech sound differences at 18 months", *PNAS* 104(41), 16027-16031.
- DIEZ-ITZA, E. (1992): *Adquisición del lenguaje*, Oviedo: Pentalfa.
- DINNSEN D. A. (1992): "Variation in developing and fully developed phonologies". En: FERGUSON, C. A., MENN, L., STOEL-GAMMON, C. (eds.). *Phonological development: Models, research, implications*. Timonium, MD: York Press, pp. 191–210.
- DISNER, S. F. (1980): "Evaluation of vowel normalization procedures", *Journal of the Acoustical Society of America* 67 (1), 253-261.
- DISNER, S. F. (1986): On Describing Vowel Quality. En: OHALA, J. J. y JAEGER, J. J. (eds.): *Experimental Phonology* (pp. 69–79): New York: Academic Press.
- DORMAN, M. F., LOIZOU, P. C., KIRK, K. I., y SVIRSKY, M. (1998): Channels, children and the Multisyllabic Lexical Neighborhood Test (MLNT), *NIH Neural Prosthesis Workshop*, Bethesda, MS; October, 1998.
- DRUMMOND, M. (1916): "Notes on speech development I", *Child Study* 9, 83-86.
- DUBRUL, E. I. (1977): "Origin of the speech apparatus and its reconstruction in fossils", *Brain lang* 4, 365-381.
- DUNN, L. M. (1965): *Peabody Picture Vocabulary Test*, American Guidance Service, Minneapolis, Minn.

- DYSON, A. T. (1988): "Phonetic inventories of 2- and 3-year-old children", *Journal of Speech and Hearing Disorders* 53, 89- 93.
- ECKEL, H. E., KOEBKE, J., SITTEL, C., SPRINZL, G. M., POTOTSCHNIG, C. y STENNERT, E. (1999): "Morphology of the human larynx during the first five years of life studied on whole organ serial sections", *Annals of Otolaryngology, Rhinology, & Laryngology* 108(3), 232-238.
- ECKEL, H. E., SPRINZL, G. M., SITTEL, C., KOEBKE, J., DAMM, M. y STENNERT, W. (2000): "Anatomy of the vocal folds and subglottic airway in children" [German], *HNO* 48, 501-507.
- EGUCHI, S. y HIRSCH, I. J. (1969): "Development of speech sounds in children", *Acta Otolaryngologica Suppl.* 257, 5-51.
- EILERS, R. E. y MINFIE, F. D. (1975): "Fricative discrimination in early infancy", *Journal of Speech and Hearing Research* 18, 158-167.
- EILERS, R. E., OLLER, D. K. y GAVIN, W. J. (1978): A cross-linguistic study of infant speech perception. Paper presented at the southeastern Conference on Human Development. Atlanta, Georgia.
- EIMAS, P. D. (1978): "Developmental aspects of speech perception". En: HELD, R., LEIBOWITZ, H. W. y TEUBER, H. L. (eds.) *Handbook of sensory Physiology* (vol. 8), Berlin: Springer-Verlag.
- EIMAS, P. D., SIQUELAND, E. R., JUSCZYK, P. y VIGORITO, J. (1971): "Speech perception in infants", *Science* 171, 303 – 306.
- EISENBERG, L. S., SHANNON, R. V., MARTÍNEZ, A. S., WYGONSKY, J. y BOOTHROYD, A. (2000): "Speech recognition with reduced spectral cues as a function of age", *Journal of the Acoustical Society of America* 107, 2704-2710.
- ELLIOTT, L. L., LONGINOTTI, C., MEYER, D., RAZ, I. y ZUCKER, K. (1981): "Developmental differences in identifying and discrimination of CV syllables", *Journal of the Acoustical Society of America* 67, 343-344.
- ENGLUND, K. y BEHNE, D. (2006): "Changes in infant directed speech in the first six months", *Infant and Child Development* 15, 139-160.



- ENGLUND, K. y BEHNE, D. (2005): "Infant directed speech in natural interaction - Norwegian vowel quantity and quality", *Journal of Psycholinguistic Research* 34(3), 259-280.
- ENLOW, D. H. y HARRIS, D. B. (1964): "A study of the postnatal growth of the human mandible", *American Journal of Orthodontics* 50, 25-50.
- ESCUDERO, P. y POLKA, L. (2003): "A cross-language study of vowel categorization and vowel acoustics: Canadian English versus Canadian French," *15th International Congress of the Phonetic Sciences*, Barcelona, Spain, Vol. 1, pp. 861-864.
- FAGAN, M. K. e IVERSON, J. M. (2007): "The influence of Mouthing on Infant Vocalization", *Infancy* 11(2), 191-201.
- FAHEY, R. P., DIEHL, R. L. y TRAUNMÜLLER, H. (1996): "Perception of back vowels: Effects of varying F1-F0 Bark distance", *Journal of the Acoustical Society of America* 99(4), 2350-2357.
- FAIRBANKS, G. (1942): "An acoustical study of the pitch of infant wails", *Child Development* 13, 227-232.
- FANT, G. (1956): "On the Predictability of Formant Levels and Spectrum Envelopes from Formant Frequencies". En: HALLE, M- LUNT, MACLEAN, H. (eds.). *For Roman Jakobson*. Mouton: The Hague. pp. 109-120; En: LEHISTE, I. (ed.). *Readings in Acoustic Phonetics*. Cambridge: The MIT Press, 1967, pp. 44-55.
- FANT, G. (1960): *Acoustic Theory of Speech Production. With Calculations based on X-Ray Studies of Russian Articulations*. The Hague: Mouton (Description and Analysis of Contemporary Standard Russian, 11):
- FANT, G. (1965): "Formants and Cavities", en *Proceedings of the 5th International Congress of Phonetic Sciences*. Münster, 1964. Basel: S. Karger. pp. 120-141.
- FANT, G. (1968): "Analysis and synthesis of the speech processes". En: MALMBERG, B. (ed.): *Manual of Phonetics*. Amsterdam: North Holland Publishing Company, pp. 173-277.
- FANT, G. (1973): "Acoustic description and classification of phonetic units". En: *Speech sounds and features*. (pp. 32-83): Cambridge, MA: The MIT Press. (Original work published 1959).

- FANT, G. (1975): A note on vocal tract size factors and non-uniform F pattern scalings. *Speech Transmission Laboratory Quarterly Progress and Status Reports (Royal Institute of Technology, Stockholm)* 4, 22-30.
- FANT, G. (2004): "Acoustical analysis of speech". En *Speech acoustics and phonetics: Selected writings*. (pp. 162-74): Dordrecht: Kluwer. (Original work published 1997).
- FANT, G., CARLSON, R., y GRANSTRÖM, B. (1974): "The [e]-[ɛ] ambiguity", *Speech Communication Seminar*, vol. 3, *Speech Transmission Laboratory*, 117-121.
- FANT, G., ISHIZAKA, K., LINDQVIST, J. y SUNDBERG, J. (1972): "Subglottal formants," KTH Speech Transmission Laboratory Quarterly Progress and Status Report 1, 1-12.
- FELD, V. (1999): Antecedentes en las investigaciones del juego vocal. Su observación en niños prematuros de menos de 1500 gramos, Hospital Materno-Infantil Ramón Sardá, Esteban de Luca 2151 (C.P.:1246) Conferencia en Buenos Aires-Argentina. Sección Neuropediatria, Neuropsicología Infantil.
- FENNELL, C.T., HUDON, T. y SPRING, M. (2010): French-English bilingual infants' discrimination of vowel contrasts, Ottawa, ON, *Development 2010: A Canadian Conference on Developmental Psychology*.
- FENTRESS, J. C. (1984): "The development of coordination", *Journal of Motor Behavior* 16, 99-134.
- FERGUSON, C. A. (1963): Baby talk in six languages, *Am.Anth* 66, 103-114.
- FERGUSON, C. A. (1979): Phonology as an individual access system: some data from language acquisition. En: FILLMORE, C. J., KEMPLER, D. y WANG, W. S-Y (eds.). *Individual Differences in Language Ability and Language Behavior* (Academic, New York), pp. 189-201.
- FERGUSON, C. A. (1986): Discovering sound unit and constructing sound system: it's child play. En: PERKELL, J. S. y KLATT, D. H. (eds.). *Invariance and variability in speech processes*. Hillsdale NJ: Erlbaum.
- FERGUSON, C. A. y FARWELL, C. B. (1975): "Words and sounds in early language acquisition", *Language* 51, 419-439.

- FERNALD, A., TAESHCHER, T., DUNN, J., PAPOUSEK, M., de BOYSSON-BARDIES, B. y FUKUI, I. (1989): "A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants", *Journal of Child Language* 16, 477-501.
- FITCH, H. (2002): Comparative vocal production and the evolution of speech: reinterpreting the descent of the larynx. En: WRAY, A. (ed.). *The Transition to language*, Oxford, Oxford University Press.
- FITCH, T. W. y GIEDD, J. (1999): "Morphology and development of the human vocal tract: A study using magnetic resonance imaging", *Journal of Acoustic Society of America* 106 (3), 1511-1522.
- FLANAGAN, J. L. (1955): "A difference limen for vowel formant frequency", *Journal of the Acoustical Society of America* 27, 613-617.
- FLANAGAN, J. L. (1972): *Speech analysis, synthesis and perception*, 2<sup>nd</sup> ed. Springer, New York, Heidelberg, Berlin.
- FLEGE, J. E. (1982): Timing in the speech of children and adults. Paper presented at the meeting of the Acoustical Society of America, Chicago, IL.
- FLETCHER, S. G. (1973): Maturation of the speech mechanism, *Folia Phoniat.* 25(3), 161-172.
- FLETCHER, N. (1992): *Acoustic Systems in Biology*. New York, Oxford University Press.
- FOK, A. (1979): "The frequency of occurrence of speech sounds and tones in Cantonese". En: LORD, M. R. (eds.). *Hong Kong Language Papers* (Hong Kong: Hong Kong University Press), pp. 150-157.
- FOLKINS, J. W. y ABBS, J. H. (1975): "Lip and jaw motor control during speech: Responses to resistive loading of the jaw", *Journal of Speech and Hearing Research* 18, 207-220.
- FOSNOT, S. (1997): Vowel development of i and u in 15-36 month old children at risk and not at risk to stutter, EuroSpeech '97, 5<sup>th</sup>, European.
- FOULKE, K. y STINCHFIELD, S. (1929): "The speech development of four infants under two years of age", *Pd. Sem. And J. Genet. Psicol.* 36, 140-171.
- FRIED, M. (1983): *The larynx: a multidisciplinary approach* (2nd ed.), St. Louis, Mo: Mosby.
- FRIED, M. P., KELLY, J. H. y STROME, M. (1982): "Comparison of the adult and infant larynx", *J. Fam. Pract.* 15(3), 557-561.

- FRY, D. B. (1979): *The physics of speech*. Cambridge: Cambridge University Press.
- FRY, D. B., ABRAMSON, A. S., EIMAS, P. D. y LIBERMAN, A. M. (1962): “The identification and discrimination of synthetic vowels”, *Language and Speech* 5, 171-189.
- FUJIMURA, O. (1960): Spectra of nasalized vowels, *Res. Lab. Electron, Q. Prog. Rep.* n° 58, MIT, 214-218 (July, 15, 1960).
- FUJIMURA, O. (1961): Analysis of nasalized vowels, *Res. Lab. Electron. A Prog. Rep* n° 62, MIT, 191-192 (July, 15, 1961).
- FUJIMURA, O. y LINDQVIST, J. (1971): “Sweep-tone measurements of vocal-tract characteristics”, *Journal of the Acoustical Society of America* 49, 541-558.
- FUJISAKI, H. y KAWASHIMA, T. (1968): “The roles of pitch and higher formants in the perception of vowels”, *IEEE Trans. Audio Electroacoustic*, AV-16, 73-77.
- FUJISAKI, H. y KAWASHIMA, T. (1969): “On the modes and mechanisms of speech perception”, *Annual Report of the Engineering Research Institute* 28, 67-73.
- GASSER, R. F. (1967): “The development of the facial muscles in man”, *American Journal of Anatomy* 120, 357-376.
- GELFAND, I. M., GURFINKEL, V. S., TSETLIN, M. L. y SHIK, M. L. (1971): “Some problems in the analysis of movements”. En: GELFAND, I., GURFINKEL, V., FOMIN, S. y TSETLIN, M. (eds.). *Models of the Structural-Functional Organization of Certain Biological Systems*. Cambridge, MA: MIT Press.
- GEORGE, S. L. (1978): The relationship between cranial base angle morphology and infant vocalizations. Doctoral Dissertation, University of Connecticut, Storrs.
- GERARD, J. M., WILHELMS-TRICARICO, R., PERRIER, P. y PAYAN, Y. (2003): “A 3D dynamical biomechanical tongue model to study speech motor control. Recent Research”, *Developments in Biomechanics* 1, 49-64.
- GIESBRECHT, N. (2002): “Creaky voice in a Pre-Babbling infant”, *Linguistics* 382, 1-23.
- GIL FERNÁNDEZ, J. (1988): *Los sonidos del lenguaje*, Madrid, Síntesis.

- GILBERT, J. H. (1970): "Formant concentration positions in the speech of children at two levels of linguistic development", *Journal of the Acoustical Society of America* 48(6) part 2, 1404-1406.
- GILBERT, H. R., ROBB, M. P. y CHEN, Y. (1997): "Formant frequency development: 15 to 36 months", *Journal of Voice* 11(3), pp. 260-266.
- GOLDSTEIN, U. G. (1979): Modeling children's vocal tracts, Paper presented at the 97<sup>th</sup> Meeting of the Acoustical Society of America, Cambridge, Mass.
- GOLDSTEIN, U. G. (1980): An articulatory model for the vocal tracts of growing children, Ph.D. thesis, MIT, Cambridge, MA.
- GOLLIN, W. (1981): "Development and plasticity". En: GOLLIN, E. (ed.). *Developmental Plasticity* (Academic, New York), 231-252.
- GONZÁLEZ ÁLVAREZ, J. (1997): "La percepción del lenguaje", en *Psicología del lenguaje. Libro de prácticas*. Universidad Jaume I: Castellón.
- GOTO, H. (1971): "Auditory perception by normal Japanese adults of the sounds 'L' and 'R' ", *Neuropsychologia* 9, 317-323.
- GOTTFRIED, T. L. y STRANGE, W. (1980): "Identification of coarticulated vowels", *Journal of the Acoustical Society of America* 68, 1626-1635.
- GREEN, J. R., MOORE, C. A., RUARK, J. L., RODDA, P. R., MORVÉE, W. T. y Van WITZENBURG, M. J. (1997): "Development of chewing in children from 12 to 48 months: Longitudinal study of EMG patterns", *Journal of Neurophysiology* 77, 2704-2727.
- GREEN, J. R., MOORE, C. A., HIGASHIKAWA, M. y STEEVE, R. W. (2000): "The physiological development of speech motor control: lip and jaw coordination", *Journal of Speech, Language, and Hearing Research* 43, pp. 239-255.
- GREEN, J. R., MOORE, C. A. y REILLY, K. J. (2002): "The sequential development of jaw and lip control for speech", *Journal of Speech, Language, and Hearing Research* 45, 66-79.
- GRUNWELL, P. (1987): *Clinical Phonology*, 2nd Edn (London: Croom-Helm).

- GRUPO DE INVESTIGACIÓN EN INFORMÁTICA APLICADA AL PROCESADO DE SEÑAL E IMAGEN: *Ancalvoz* (2009), Universidad Politécnica de Madrid.
- GRUPO DE INVESTIGACIÓN EN INFORMÁTICA APLICADA AL PROCESADO DE SEÑAL E IMAGEN: *PangurBan* (2011), Universidad Politécnica de Madrid.
- GRUPO DE INVESTIGACIÓN KOINÉ: Corpus Koiné, Universidad de Santiago de Compostela.
- GUENTHER, F. H. (1994): “A neural network model of speech acquisition and motor equivalent speech production, Running title: Speech acquisition and motor equivalence”, *Biological Cybernetics* 72, 43-53. (1-20).
- GUENTHER, F. H. (1995): “Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production”, *Psychological Review* 102, 594–621.
- GUENTHER, F. H., HAMPSON, M. y JOHNSON, D. (1998): “A theoretical investigation of reference frames for the planning of speech movements”, *Psychological Review* 105, 611-633.
- GUENTHER, F. H. y PERKELL, J. (2004): A neural model of speech production and supporting experiments, From sound to sense, MIT, p. 98-106.
- GUIRAO, M. y BORZONE de MANRIQUE, A. M. (1972): “Fonemas, sílabas y palabras del español de Buenos Aires”, *Filología*, XVI, págs. 135-165.
- HAGIWARA, R. E. (1997): “Dialect variation and formant frequency: the American English vowels revisited”, *Journal of the Acoustical Society of America* 102, 655-658.
- HALLE, M. y STEVENS, R. N. (1959): Analysis by synthesis. En: WATHEN-DUNN, W. y WOODS, I. E. (eds.). Proceedings of the seminar on speech compression and processing, AFCRC-TR-59-198, December 1959, vol. II, paper D7.
- HARDCASTLE, W. J., MORGAN BARRY, R. A. y CLARK, C. J. (1987): “An instrumental phonetic study of lingual activity in articulation-disordered children”, *Journal of Speech and Hearing Research* 30, 171-184.
- HARE, G. (1983): “Development at 2 years”. En: IRWIN, J. V. y WONG, S. P. (eds.). *Phonological Development in Children: 18-72 months* (Carbondale: Southern Illinois University Press), Pp. 55-88.

- HARSHMAN, R., LADEFOGED, P. y GOLDSTEIN, L. (1977): "Factor analysis of tongue shapes", *Journal of the Acoustical Society of America* 62, 693-707.
- HATTORI, S., YAMAMOTO, K. y FUJIMURA, O. (1956): "Nasalization of vowels and nasals", *Rep. Kobayashi sci. Inst.* 6, 226-235.
- HATTORI, S., YAMAMOTO, K. y FUJIMURA, O. (1958): "Nasalization of vowels in relation to nasals", *Journal of the Acoustical Society of America* 30, 267-274.
- HAWKINS, S. y STEVENS, K. N. (1983): "A cross-language study of the perception of nasal vowels", *Journal of the Acoustical Study of America* Suppl. 1 73, S54.
- HAWKINS, S. y STEVENS, K. (1985): "Acoustic and perceptual correlates of the non-nasal-nasal distinction for vowels", *Journal of the Acoustical Society of America* 77(4), 1560-1575.
- HEINZ, J. M. (1967): "Perturbation functions for the determination of vocal-tract area function from vocal-tract eigenvalues", *Speech Transmission Laboratory, Quarterly Progress and Status Report* 7, 1-14.
- HEPPER, P., SCOTT, D. y SHAHIDULLAH, S. (1993): "Newborn and fetal response to maternal voice", *Journal of Reproductive & Infant Psychology* 11, 147-153.
- HERSHENSON, M. B. (1992): The respiratory muscles and chest wall. En: BECKERMAN, R. C., BROUILLETTE, R. T. y HUNT, C. E. (eds.). *Respiratory control disorders in infants and children* (pp. 28-44): Baltimore: Williams & Wilkins.
- HILLEBRAND, J. M., GETTY, L. A., CLARK, M. J. y WHEELER, K. (1995): "Acoustic characteristics of American English vowels", *Journal of the Acoustical Society of America* 97, 3099-111.
- HILLS, E. C. (1914): "The speech of a child two years of age", *Dialect Notes* 4, 84-100.
- HIRANO, M., KURITA, S. y NAKASHIMA, T. (1981): "The structure of the vocal folds". En: STEVENS, K. S. e HIRANO, M. (eds.). *Vocal fold Physiology*, University of Tokyo Press.
- HIRANO, M., KURITA, S. y NAKASHIMA, T. (1983): "Growth development and aging of the vocal folds". En: BLESS, D. M. y ABBS, J. H. (eds.). *Vocal fold physiology: Contemporary research and clinical issues* (pp. 22-43), San Diego, CA: College-Hill Press.

- HODGE, M. M. (1989): A comparison of spectral-temporal measures across speaker age: Implications for an acoustic characterization of speech maturation. Unpublished doctoral dissertation, University of Wisconsin-Madison.
- HOEMEKE, K. A. y DIEHL, R. L. (1994): "Perception of vowel height: the role of F1-F0 distance", *Journal of the Acoustical Society of America* 96(2), 661-673.
- HOLMBERG, T. L., MORGAN, K. V. y KUHL, P. A. (1977): "Speech perception in early infancy: Discrimination of fricative consonants", *Journal of the Acoustical Society of America* 62, Supplement 1,599.
- HOLMGREN, K., LINDBLÖM, B., AURELIUS, G., JALLING B. y ZETTERSTRÖM, R. (1986): "On the phonetics of infant vocalization". En: LINDBLÖM, B. y ZETTERSTRÖM, R. (eds.). *Precursors of Early speech* (Stockton, New York).
- HONDA, K. (1992): "Physiological background of speech production (tutorial, in Japanese)", *Journal of the Acoustical Society of Japan* 48, 9-14.
- HONDA, K. (1996): "The organization of tongue articulation for vowels", *Journal of Phonetics* 24, 1, 39-52.
- HONDA, K., KUSAKAWA, N. y KAKITA, Y. (1992): "An EMG analysis of sequential control cycles of articulatory activity during /əVp/ utterances", *Journal of Phonetic* 20, 53-63.
- HONDA, K., KURITA, T. y KAKITA, Y. (1995): "Physiology of the lips and modeling of lip gestures", *Journal of Phonetics* 23, 243-254.
- HOPKINS, B. (2005): *The Cambridge Encyclopedia of Child Development*, Cambridge University Press.
- HOUSE, A. S. y STEVENS, K. N. (1956): "Analog studies of the nasalization of vowels", *Journal of Speech and Hearing Disorders* 21, 218-232.
- HOWATT, W. F. y DEMUTH, G. R. (1965): "The growth of lung function: II. Configuration of the chest", *Pediatrics* 35, 177-184.
- HOWELL, P. (1981): "Identification of vowels in and out of context", *Journal of the Acoustical Society of America* 70(5), 1256-1260.



- HUBER, J. E., STATHOPOULOS, E. T., CURIONE, G. M., ASH, T. A y JOHNSON, K. (1999): "Formants of children, women and men. The effects of vocal intensity variation", *Journal of the Acoustical Society of America* 106 (3), 1532-1542.
- HUMPHREY, T. (1964): "Some correlations between the appearance of human fetal reflexes and the development of the nervous system", *Progress in Brain Research* 4, 93-135.
- HUMPHREYS, M. W. (1880): "A contribution to infantile linguistics", *Trans. Amer. Philol. Assoc* 11, 5-17.
- HUNTER, J. (1771): *The Natural History of the Human Teeth: Explaining Their Structure, Use, Formation, Growth and Diseases*, London.
- IRIBAR, A. (2012). Caracterización fonética experimental del vocalismo vasco-románico. Tesis doctoral.
- IRWIN, O. C. (1948): "Infant speech: Development of vowel sounds", *Journal of Speech and Hearing Disorders* 13, 31-34.
- IRWIN, O. C. y CHEN, A. P. (1946): "Infant speech: vowel and consonant frequency", *Journal of Speech and Hearing Disorders* 11, 123-124.
- IRWIN, J. V. y WONG, S. P. (1983): *Phonological Development in Children: 18 to 72 Months*, Carbondale, IL: Southern Illinois University Press.
- ISHIZUKA, K., MUGITANI, R., KATO, H. y AMANO S. (2007): "Longitudinal developmental changes in spectral peaks of vowels produced by Japanese infants", *Journal of the Acoustical Society of America* 121(4), 2272-2282.
- IVERSON, P. y KUHL, P. K. (1995): "Mapping the perceptive magnet effect for speech using signal detection theory and multidimensional scaling", *Journal of the Acoustical Society of America* 97, 553-562.
- IVERSON, J. M. y FAGAN, M. K. (2004): "Infant vocal-motor coordination: Precursor to the gesture-speech system?", *Child Development* 75, 1053-1066.
- JACKSON MENALDI, M. C. (1992): *La voz normal*, Buenos Aires, Editorial Médica Panamericana.
- JAKOBSON, R. (1941): *Lenguaje infantil, afasia y leyes fonéticas generales*, Upsala. Traducción al español: *Lenguaje infantil y afasia*, 1974, ed. Ayuso.

- JOHNSON, K., FLEMMING, E. y WRIGHT, R. (1993): "The hyperspace effect: Phonetic targets are hyperarticulated", *Language* 69, 505-528.
- JOOS, M. (1948): "Acoustic phonetics", *Language Monogr.* 23 (suppl. to 24, 2).
- JUNG, T. (2009): Acoustic articulatory evidence for quantal vowel categories: the features [low] and [back], Submitted to the Harvard-MIT Division of Health Sciences and Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Speech and Hearing Bioscience and Technology.
- JUSCZYK, P. W. (1997): *The Discovery of Spoken Language*, MIT Press, Cambridge, Massachusetts.
- JUSCZYK, P. W., FRIEDERICI, A. D., WESSELS, J. M. I., SVENKERUD, V. Y. y JUSCZYK, A. M. (1993): "Infants' sensitivity to the sound patterns of native language words", *Journal of Memory and Language* 32, 402-420.
- KADOKAWA, Y. y SUZUKI, J. (1969): "A simple method for the vocal-tract configuration from the first three formant frequencies", *Electronics and Communication in Japan* 52-C (9), 123-130.
- KAROUSOU, A. (2003): Análisis de las vocalizaciones tempranas. Su patrón evolutivo y su función determinante en la emergencia de la palabra. Tesis doctoral, Universidad Complutense de Madrid.
- KAROUSOU, A. y LÓPEZ ORNAT, S. (2013): "Prespeech Vocalizations and the Emergence of Speech: A Study of 1005 Spanish Children", *The Spanish Journal of Psychology* 16, e32, 1-21.
- KAZARIAN, A. G., SARKISSIAN, L. S. e ISAAKIAN, D. G. (1978): "Length of the human vocal cords by age", *Zhurnal Eksperimentalnoi I Klinicheskoi Meditsiny* 18, 105-109.
- KELSO, J. A. S. y NORMAN, P. E. (1978): "Motor schema formation in children", *Developmental Psychology* 14, 153-156.
- KELSO, J. A. S., TULLER B., VATIKIOTIS-BATESON, E. y FOWLER, C. A. (1984): "Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures", *Journal of Experimental Psychology: Human Perception and Performance* 10, 812-832.

- KENT, R. D. (1976): "An anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies", *Journal of Speech and Hearing Research* 19, 421-447.
- KENT, R. D. (1981): "Articulatory-acoustic perspectives on speech development". En: STARK, R. (ed.). *Language Behavior in Infancy and Early Childhood*, North-Holland, Elsevier, Amsterdam, pp. 105-239.
- KENT, R. D. (1981): "Sensorimotor aspects of speech development". En: ASLIN, R. N., ALBERTS, J. R. y PETERSON, M. R. (eds.). *Development of perception, Audition, Somatic Perception and the chemical senses*, vol. 1.
- KENT, R. D. (1984): "Psychobiology of speech development: co-emergence of language and a movement system", *American Journal of Physiology* 246(6), R888-R894.
- KENT, R. D. (1992): "The biology of phonological development". En: FERGUSON, C., MENN, L. y STOEL-GAMMON, C. (eds.). *Phonological development. Models, Research, Implications*, Timonium, Maryland, pp. 65-90.
- KENT, R. D. (1999): "Motor control: Neurophysiology and functional development". En: CARUSO, A. y STRAND, E. (eds.). *Clinical Management of Motor Speech Disorders in Children*, Thieme Medical Publishers, New York.
- KENT, R. D. (2004): Development, Pathology and Remediation of Speech, From Sound to Sense: June 11-June 13, 2004 at MIT.
- KENT, R. D. y FORNER, L. L. (1979): "Development study of vowel formant frequencies in an imitation task", *Journal of the Acoustical Society of America* 65, 208-217.
- KENT, R. D. y FORNER, L. L. (1980): "Speech segment durations in sentence recitations by children and adults", *Journal of Phonetics* 8, 157-168.
- KENT, R. D. y MURRAY, A. D. (1982): "Acoustic features of infant vocalic utterances at 3, 6, and 9 months", *Journal Acoustical Society of America* 72 (2), 353-365.
- KENT, R. D. y BAUER, H. R (1985): "Vocalizations of one-year-olds", *Journal of Child Language* 12, 491-526.

- KENT, R. D., OSBERG, M. J., NETSELL, R. y HUSTEDDE, C. (1987): "Phonetic development in identical twins differing in auditory function", *Journal of Speech and Hearing Disorders* 52, 64-75.
- KENT, R. D., MITCHELL, P. R. y SANCIER, M. (1991): "Evidence and role of rhythmic organization in early vocal development in human infants". En: FAGARD, J. y WOLFF, P. H. (eds.). *The development of timing control and temporal organization in coordinated actions* (pp. 135-139), Elsevier: New York.
- KENT, R. D. y READ, C. (1992): *The acoustic analysis of speech*, San Diego, CA: Singular Publishing Group.
- KENT, R. D. y MIOLO, G. (1995): "Phonetic Abilities in the First Year of Life". En: FLETCHER, P. y MCWHINNEY, B. (eds.). *The handbook of child language*, 303-304, Oxford, Blackwell.
- KENT, R. D. y VORPERIAN, H. K. (1995): "Anatomic development of the craniofacial-oral-laryngeal systems: A review", *Journal of Medical Speech-Language Pathology* 3, 145-190.
- KENT, R. D. y VORPERIN, H. K. (2006): "In the mouths of Babes: Anatomic, Motor and Sensory foundations of speech Development in Children". En: RHEA, P. (ed.). *Language disorders from a developmental perspective: Essays in honour of. Robin S. Chapman*. London: Lawrence Erlbaum.
- KENYÈRES, E. (1927): "Les premiers mots de l'enfant et l'apparition des espèces de mots dans le langage", *Arch. De Psychol.* 21, 191-218.
- KERN, S. y DAVIS, B. L. (2009): "Emergent complexity in early vocal acquisition: Cross linguistic comparisons of canonical babbling". En: CHITORAN, I., COUPÉ, C., MARSICO, E. y PELLEGRINO, F. (eds.). *Approaches to Phonological Complexity*, Phonology and Phonetics Series, 353-376. Berlin: Mouton de Gruyter.
- KERN, S., DAVIS, B., MACNEILGAE, P., KOÇBAS, D., KUNTAY, A. y ZINK, I. (in press): Crosslinguistic similarities and differences in babbling: Phylogenetic implications. En: HOMBERT, J. M. (ed.). *Towards the Origins of Language and Languages*, OMLL volume.
- KERR, B. y BOOTH, B. (1978): "Specific and varied practice of motor skill", *Perceptual and Motor Skills* 46, 395-401.

- KEWLEY-PORT, D. y WATSON, C. (1994): “Formant frequency discrimination for isolated English vowels”, *Journal of the Acoustical Society of America* 95(1), 485-496.
- KING, E. W. (1952): “A roetgenographic study of pharyngeal growth”, *Angle Orthod.* 22, 23–37.
- KISAKAWA, N., HONDA, K. y KAKITA, Y. (1993): Construction of articulatory trajectories in the space of tongue muscle contraction force (in Japanese): *ATR Technical Repoprt*, TR-A-0717 (ATR, Kyoto, Japan, 1993).
- KISILEVSKY, B. S., HAINS, S. M. J., BROWN, C. A., LEE, C. T., COWPERTHWAITTE, B., STUTZMAN, S. S., SWANSBURG, M. L., LEE, K., XIE, X., HUANG, H., YE, H. H., ZHANG, K. y WANG, Z. (2009): “Fetal sensitivity to properties of maternal speech and language”, *Infant behavior and development* 32, 59-71.
- KITAMURA C. y BURNHAM D. (2003): “Pitch and communicative intent in mother’s speech: adjustments for age and sex in the first year”, *Infancy* 4(1), 85–110.
- KOOPMANS-van BEIMUN, F. J., CLEMENT, C. J. y van den DIKKENBERG-Pot, I. (2001): “Babbling and the lack of auditory speech perception: a matter of coordination?”, *Dev. Sci.* 4, 61-70.
- KOOPMANS-van BEINUM, F. J. y van der STELT, J. M. (1986): “Early stages in the development of speech movements”. En: LINDBLOM, B. y ZETTERSTROM, R. (eds.). *Precursors of Early Speech*, Stockton, New York.
- KORNFELD, J. (1971): Spectrographic analysis of liquids in young children. Quarterly Progress Report of the Research Laboratory of Electronics, No. 101, MIT. In Lees, R. B. *Grammar of English nominalizations*. Bloomington, Indiana.
- KRAKOW, R. y BEDDOR, P. (1991): Coarticulation and the perception of nasality, *Proceeding of the XII International Congress of Phonetic Sciences*, 4, 38-41.
- KRATZENSTEIN, C. G. (1780): “Sur la naissance de la formation des voyelles”, *Journal of Physiology* 21 (1782), 358-81 (translated from Acta Academic Petrograd).
- KUHL, P. K. y MELTZOFF, A. N. (1982): “The bimodal development of speech in infancy”, *Science* 218, 1138–1141.

- KUHL, P. K. (1983): "Perception of auditory equivalence classes for speech in early infancy", *Infant Behavior and Development* 6, 263-285.
- KUHL, P. K. (1991): "Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not", *Perception and Psychophysics* 50, 93-107.
- KUHL, P. K. y MILLER J. D. (1982): "Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants", *Perception and Psychophysics* 341, 279-92.
- KUHL, P. K., WILLIAMS, K. A., LACERDA, F., STEVENS, K. N. y LINDBLOM, B. (1992): "Linguistic experience alters phonetic perception", *Science* 255, 606-608.
- KUHL, P. K. y MELTZOFF, A. N. (1996): "Infant vocalizations in response to speech: vocal imitation and development change", *Journal of Acoustic Society of America* 100 (4), 2425-2438.
- KUHL, P. K., ANDRUSKI, J. E., CHISTOVICH, I. A., CHISTOVICH, L. A., KOZHEVNIKOVA, E. V., RYSKINA, V. L., STOLYAROVA, E. I., SUNDBERG, U. y LACERDA, F. (1997): "Cross-language analysis of phonetic units in language addressed to infants", *Science* 277, 684-686.
- KUSAKAWA, N., HONDA, K. y KAKITA, Y. (1993): Construction of articulatory trajectories in the space of tongue muscle contraction force (in Japanese): *ATR Technical Report*, TR-A-0717 (ATR, Kyoto, Japan, 1993).
- LACHARITE, D. y PARADIS, C. (1997): "Category preservation and proximity versus phonetic approximation in loanword adaptation", *Linguistic Inquiry* 36, 223-258.
- LADEFOGED, P. (1982): *A course in phonetics*. Orlando, FL Harcourt Brace Jovanovich.
- LADEFOGED, P. (2001): *Vowels and Consonants: An introduction to the sounds of languages*. Oxford: Blackwell Publishers.
- LADEFOGED, P. y BROADBENT, D. (1957): "Information conveyed by vowels", *Journal of the Acoustical Society of America* 29, 98-104.

- LADEFOGED, P., De CLERK, J., LINDAU, M. y PAPCUN, G. (1972): "An auditory-motor theory of speech production", *UCLA Working Papers in Phonetics* 22, 48-76.
- LADEFOGED, P., HARSHMAN, R., G., L. y RICE, L. (1978): "Generating Vocal Tract Shapes from Formant Frequencies", *Journal of Acoustical Society of America* 64(4), 1027-1035.
- LAITMAN, J. T. y CRELIN, E. S. (1976): Postnatal development of the basicranium and vocal tract region in man, in *Symposium on the Development of the Basicranium*, edited by J. Bosma, DHEW Publication No 76-989 (PHS-NIH, Bethesda, MD), pp. 206-220.
- LAITMAN, J. T., CRELIN, E. S. y CONLOGUE, G. J. (1977): "The function of the epiglottis in monkey and man", *Yale Journal of Biology and Medicine* 50, 43-48.
- LANDBERG, L. R. y LUNDBERG, L. J. (1989): "Phonetic development in early infancy: a study of four Swedish children during the first eighteen months of life", *Journal of Child Language* 16, 19-40.
- LANGLOIS, A., BAKEN, R. J. y WILDER, D. N. (1980): Pre-speech behavior during the first year of life. En: MURRAY, T. y MURRAY, J. (eds.): *Infant communication, cry and early speech*. Houston, Texas, College Hill Press.
- Le HUCHE, F. y ALLALI, A. (2004): *La voz*, Barcelona, Ed. Masson.
- LECANUET, J. P., GRANIER-DEFERRE, C., DECASPER, A. J., MAUGEAIS, R., ANDRIEU, A. J. y BUSNEL, M. C. (1987): "Perception et discrimination foetales de stimuli langagiers, mise en évidence à partir de la réactivité cardiaque, résultats préliminaires", *C R Acad Sci Paris (III)* 305, 161-164.
- LECANUET, J. P., GRANIER-DEFERRE, C. y BUSNEL, M. C. (1995): "Human fetal auditory perception". En: LECANUET, J. P., FIFER, W. P., KRASNEGOR, N. A. y SMOTHERMAN, W. P. (eds.). *Fetal Development: A Psychobiological Perspective*, pp. 239-62. Hillsdale, NJ: Lawrence Erlbaum Associates.
- LECANUET, J. P., GRANIER-DEFERRE, C., JACQUET, A. Y. y DECASTER, A. J. (2000): "Fetal discrimination of low pitched musical notes", *Developmental Psychobiology* 36, 29-39.

- LEE, S., POTAMIANOS, A. y NARAYANAN, S. (1999): “Acoustic of children’s speech: Developmental changes of temporal and spectral parameters”, *Journal of the Acoustical Society of America* 105(3), 1455-1468.
- LEE, S., DAVIS, B. y MACNEILAGE, P. (2009): “Universal production patterns and ambient language influences in babbling: A cross-linguistic study of Korean and English-learning infants”, *Journal of Child Language* 37, 293-318.
- LEEPER, H. A., TISSINGTON, M. L. y MUNHALL, K. G. (1988): “Temporal aspects of velopharyngeal function in children”, *Cleft Palate-Craniofacial Journal* 35, 215-221.
- LENNEBERG, E. H. (1967): *Biological foundations of language*, New York: John Wiley and Sons, Inc.
- LEOPOLD, W. (1939): *Speech development of a bilingual child: A linguist's record*. Vol I. *Vocabulary growth in the first two years*. Evanston, IL: Northwestern University Press.
- LEOPOLD, W. (1947): *Speech development of a bilingual child: A linguist's record*. Vol II. *Sound learning in the first two years*. Evanston, IL: Northwestern University Press.
- LEOPOLD, W. (1949a): *Speech development of a bilingual child: A linguist's record*. Vol III. *Grammar and general problems*. Evanston, IL: Northwestern University Press.
- LEOPOLD, W. (1949b): *Speech development of a bilingual child: A linguist's record*. Vol IV. *Diary from age two*. Evanston, IL: Northwestern University Press.
- LEVELT, C. C. (1996): Consonant-vowel interactions in child language. En: BERNHARDT, B., GILBERT, J., e INGRAM, D. (eds.). *Proceeding of the UBC international conference on phonological acquisition*. Somerville: Cascadilla Press.
- LEVITT, A. G. y WANG, Q. (1991): “Evidence for language specific rhythmic influences in the reduplicative babbling of French-and-English learning infants”, *Language and Speech* 34, 235-249.
- LEVITT, A. G. y UTMAN, J. G. A. (1992): “From babbling towards the sound systems of English and French – a longitudinal - case study”, *Journal of Child Language* 19, 19-49.
- LIBERMAN, A. M., HARRIS, K. S., HOFFMAN, H. S. y GRIFFITH, B. C. (1957): “The discrimination of speech sounds within and across phoneme boundaries”, *Journal of Experimental Psychology* 54, 358 –368.



- LIBERMAN, A. M. y MATTINGLY, I. (1985): "The motor theory of speech perception revised", *Cognition* 21, 1-36.
- LIEBERMAN, P. (1968): "Primate vocalizations and human linguistic ability", *Journal of the Acoustical Society of America* 44, 1574-1584.
- LIEBERMAN, P. (1975): *On the Origins of Language*, New York: MacMillan.
- LIEBERMAN, P. (1977): *Speech Physiology and Acoustic Phonetics*, New York: MacMillan.
- LIEBERMAN, P. (1980): "On the development of vowel production in young children". En: YENI-KOMSHIAN, G. H., KAVANAGH, J. F. y FERGUSON, C. A. (eds.). *Child phonology I: Production*. New York, NY: Academic Press.
- LIEBERMAN, P. (1984): *The Biology and Evolution of Language*. Harvard University Press, Cambridge, MA.
- LIEBERMAN, P. (1986): "The acquisition of intonation by infants: Physiology and neural control". En: JOHNS-LEWIS, C. (eds.): *Intonation in Discourse* (pp. 239-257): San Diego, CA: Colege-Hill Press.
- LIEBERMAN, P., HARRIS, K. S., WOLFF, P. y RUSSELL, L. H. (1971): "Newborn infant cry and nonhuman primate vocalization", *J. Speech Hear. Res.* 14, 718-727.
- LIEBERMAN, P., CRELIN, E. S. y KLATT, D. H. (1972): "Phonetic ability and related anatomy of the newborn and adult human, Neanderthal man, and the chimpanzee", *Am. Anthropol.* 74, 287-307.
- LIEBERMAN, P. y BLUMSTEIN, S. (1988): *Speech physiology, speech perception and acoustic phonetics*, Cambridge University Press, New York.
- LIEBERMAN, D. E. y MCCARTHY, R. C. (1999): "The ontogeny of cranial base angulation in humans versus chimpanzees and its implications for reconstructing pharyngeal dimensions", *J. Hum. Evol.* 36, 487-517.
- LIEBERMAN, D. E., MCCARTHY, R. C., HIIEMAE, K. M. y PALMER, J. B. (2001): "Ontogeny of postnatal hyoid and larynx descent in humans", *Archives of Oral Biology* 46, 117-128.
- LIN, T. H. (1989): Autosegmental treatment of segmental processes in Chinese phonology. PhD dissertation, University of Texas, Austin.

- LIN, Q. (1990): Speech production theory and articulatory speech synthesis. Thesis, R.I.T. Stockholm.
- LINDBLOM, B. (1961): Accuracy and limitations of Sona-Graph measurements. En: *Proceedings of the 4th International Congress of Phonetic Sciences*. Helsinki: Mouton Press.
- LINDBLOM, B. (1984): "Can the models of evolutionary biology be applied to phonetics problems?" En: van der BROECKE, M. P. y COHEN, A. (eds.). *Proceedings of the Tenth International Congress of Phonetic Sciences*, Dordrecht: Foris.
- LINDBLOM, B. (1990): "Explaining phonetic variation: a sketch of the H&H theory". En: HARDCASTLE, W. y MARCHAL, A. (eds.). *Speech Production and Speech Modeling*, Kluwer: Dordrecht; 403-439.
- LINDBLOM, B. y SUNDBERG, J. (1969): A quantitative method of vowel production and the distinctive features of Swedish vowels, *Quarterly Progress Status Report*, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, 1, 1432.
- LINDBLOM, B. y SUNDBERG, J. (1971): "Acoustical consequences of lip, tongue, jaw, and larynx movement", *Journal of the Acoustical Society of America* 50, 1166-1180.
- LINDQVIST, J. y SUNDBERG, J. (1976): "Acoustic properties of the nasal tract", *Phonetica* 33, 161-168.
- LISKER, L. y ABRAMSON, A. S. (1970): The voicing dimension: Some experiments in comparative phonetics. En: *Proceedings of the Sixth International Congress of Phonetic Sciences* Prague Academia.
- LIU, H. M., KUHL P. K. y TSAO, F. M. (2003): "An association between mothers' speech clarity and infants' speech discrimination skills", *Developmental Science* 6(3), 1-10.
- LLEÓ, C. (1997): *La adquisición de la fonología de la primera lengua y de las lenguas extranjeras: modelos teóricos y métodos lingüísticos de análisis*, Madrid, Visor.
- LLOYD, R. J. (1890a): *Some Researches into the Nature of Vowel-Sound*, Turner and Dunnet, Liverpool, England.
- LLOYD, R. J. (1891): "Speech sounds: their nature and causation (II-IV)", *Phonetische Studien* 4, 37-67, 183-214, 275-306.

- LLOYD, R. J. (1892): "Speech sounds: their nature and causation (V-VII)", *Phonetische Studien* 5, 1-32, 129-141, 263-271.
- LLOYD, P. M. y SCHNITZER, R. D. (1967): "A statistical study of the structure of the Spanish syllable", *Linguistics* 37, 58-72.
- LOCKE, J. L. (1993): *The child's path to spoken language*. Cambridge, MA: Harvard University Press.
- LOCKE, J. L. y PEARSON, D. M. (1992): "Vocal learning and the emergence of phonological capacity. A neurobiological approach". En: FERGUSON, C., MENN, L. y STOELGAMMON, C. (eds.). *Phonological development. Models, Research, Implications*, Timonium, Maryland, pp. 91-129.
- LOGAN, J. S., LIVELY, S. E. y PISONI, D. B. (1991): Training Japanese listeners to identify English [r] and [l]: A first report, *Journal of the Acoustical Society of America* 89, 874-886.
- LONCHAMP, F. (1978): Recherches sur les indices perceptifs des voyelles orales et nasales: application à la structure du système vocalique français et de diverses autres langues, Thèse 3<sup>ème</sup> cycle, l'Université de Nancy.
- LÓPEZ ORNAT, S. (1994b): "La adquisición del lenguaje: talón de Aquiles y poción mágica de la teoría cognitiva", *Cognitiva* 6 (2), 213-239.
- LÓPEZ ORNAT, S. y KAROUSOU, A. (2005): Las vocalizaciones tempranas (8-30 meses) y su relación con el vocabulario y la gramática. Su medida en el "CDI español": resultados preliminares. En: MAYOR CINCA, M. A., ZUBIAUZ DE PEDRO, B. y DÍEZ-VILLORIA, E. (eds.): *Estudios sobre la adquisición del lenguaje*, Aquilafuente Eds. Universidad de Salamanca, 401-420.
- LÓPEZ ORNAT, S. y KAROUSOU, A. (2013): Vocalización de los bebés y desarrollo lingüístico. Suplemento de Ciencia del Notiweb de Madrid.
- LÓPEZ RAMÓN y CAJAL., C. (1996): "Description of human fetal laryngeal functions: phonation", *Early Human Development* 45, 63-72.
- LUND, J. P. y ENOMOTO, S. (1988): The generation of mastication by the central nervous system. En: COHEN, A. ROSSIGNOL, S. y GRILLNER, S. (eds.). *Neural control of rhythmic movements*, Wiley.

- LYAKSO, E. E., GROMOVA, A. D., FROLOVA, O. V. y ROMANOVA, O. D. (2005): “Acoustic aspects of the formation of speech in children in the third year of life”, *Neuroscience and Behavioral Physiology* 35(6), 573-583.
- LYAKSO, E.E., van der STELT, J. M., WEMPE, T. G. y ZAJDÓ, K. (2006): Vowel data of early speech development in several languages, *Proc. Multiling 2006*, Stellenbosch.
- MACKAIN, K. S., BEST, C. Y. y STRANGE, W. (1980): “Native language effects on the perception of liquids”, *Journal of the Acoustical Society of America* 27, 527.
- MACKAIN, I. S., BEST, C. T. y STRANGE, W. (1981): “Categorical perception of English /r/ and /l/ by Japanese bilinguals”, *Applied Psycholinguistics* 2, 369-390.
- MACMILLAN, A. S. y KELEMEN, G. (1952): “Radiography of the supraglottic speech organs”, *Arch. Otolaryngol.* 55 (6), 671-688.
- MACNEILAGE, P. F. (1998): “The Frame/Content Theory of Evolution of Speech Production”, *Behavioral and Brain Sciences* 21 (4), 499–546.
- MACNEILAGE, P. F. y DAVIS, B. L. (1990): Acquisition of speech production: Frames then content. En: JEANNEROD, M. (ed.). *Attention and Performance XII: Motor Representation and Control* (pp. 453–76): Hillsdale, NJ: Lawrence Erlbaum.
- MACNEILAGE, P. F., DAVIS, B. L., KINNEY, A. y MATYEAR, C. L. (2000): “The motor core of speech: A comparison of serial organization patterns in infants and languages”, *Child Development* 71, 153-163.
- MACWHINNEY, B. (2000). *The CHILDES Project: Tools for Analyzing Talk*. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- MADDIESON, I. y LADEFOGED, P. (1995): *Sounds of the World's Languages*. Oxford: Blackwell.
- MAEDA, S. (1982a): The role of the sinus cavities in the production of vowels, *Proc. ICASSP-82*, Paris, 911-914.
- MAEDA, S. y HONDA, K. (1994): “From EMG to Formant Patterns of Vowels: The implication of vowel space”, *Phonetica* 51, 17-29.
- MAGRIPLES, U. y LAITMAN, J. T. (1987): “Developmental change in the position of the fetal human larynx”, *Am. J. Phys. Anthropol.* 72, 463-472.

- MAJOR, D. R. (1906): *First steps in mental growth: A series of studies in the psychology of infancy*, New York: Macmillan, pp 360 (pp. 317-333):
- MAREAN, G., C., LYNNE, A., WERNER, L. A. y KUHL, P. (1992): "Vowel categorization by very young infants", *Developmental Psychology* 28(3), 396-405.
- MARTIN, J. (1981): *Voice, Speech and Language in Child Development and Disorders*, New York: Springer-Verlag.
- MARTIN, P. (2002): "Le système vocalique du français du Québec. De l'acoustique à la phonologie," *La Linguistique* 38(2), 71-88.
- MARTÍNEZ CELDRÁN, E. (1984): *Fonética*, Barcelona, Teide (4ª ed., 1993).
- MARTÍNEZ CELDRÁN, E. (1995b): "En torno a las vocales del español: análisis y reconocimiento", *Estudios de Fonética Experimental*, VII, Universitat de Barcelona, Laboratori de Fonètica, pp. 195-218.
- MARTÍNEZ CELDRÁN, E. (1998): *Análisis espectrográfico de los sonidos del habla*, Barcelona, Ariel.
- MARTÍNEZ CELDRÁN, E. y FERNÁNDEZ PLANAS, A. M. (2007): *Manual de fonética española*, Barcelona, Ariel.
- MATTOCK, K., RVACHEW, S., POLKA, L. y TURNER, S. (2005): "A comparison of vowel formant frequencies in the babbling of infants exposed to Canadian English and Canadian French", *Journal of the Acoustical Society of America* 117 (4), 2402-2402.
- MATTOCK, K. y BURNHAM, D. (2006): "Chinese and English infants' tone perception: Evidence for perceptual reorganization", *Infancy* 10, 241 - 65.
- MATTOCK, K., AMITAY, S. y MOORE, D. R. (2010): "Auditory development and learning". En: PLACK, C. (ed.): *Oxford Handbook of Auditory Science: Auditory Perception*. Oxford: Oxford University Press.
- MATYEAR, C. L., MACNEILAGE, P. F. y DAVIS, B. L. (1998): "Nasalization of vowels in nasal environments in babbling: Evidence for Frame dominance", *Phonetica* 55, 1-17.
- MAXWELL, E. M. (1979): "Competing Analyses of a Deviant Phonology", *Glossa* 13, 181-213.

- MAYE, J., WERKER, J. F. y GERKEN, L. A. (2002): "Infant sensitivity to distributional information can affect phonetic discrimination", *Cognition* 82(3), B101-B11.
- MCALLISTER, T. K. (2009): *The articulatory basis of positional asymmetries in phonological acquisition*, Massachusetts Institute of Technology.
- MCCAFFREY, A. (1971): Speech perception in infancy. Unpublished doctoral dissertation, Cornell University.
- MCCARTHY, J. J. y KIRK, S. A. (1961): Illinois Test of psycholinguistic Abilities: Experimental Edition, Univ. of Illinois Press, Urbana.
- MCGOWAN, R. S., NITTROUER, S. y CHENAUSKY, K. (2008): "Speech production in 12 month old children with and without hearing loss", *J. Speech Lang. Hear. Res.* 51, 879-888.
- MCNEILL, D. (1970): *The acquisition of language: The study of developmental psycholinguistics*, New York: Harper.
- MEHLER, J., BERTONCINI, J. y BARRIÈRE, M. (1978): "Infant recognition of mother's voice", *Perception* 7, 491-497
- MEHLER, J., JUSCZYK, P., LAMBERTZ, G., HALSTED, N., BERTONCINI, J. y AMIELTISON, C. (1988): "A precursor of language acquisition in young infants", *Cognition* 29, 143-178.
- MELTZOFF, A. N. y MOORE, M. K. (1989): "Imitation in Newborn Infants: Exploring the Range of Gestures Imitated and the Underlying Mechanisms", *Developmental Psychology* 25, (6), 954-962.
- MELTZOFF, A. N. y MOORE, M. K. (1977): "Imitation of facial and manual gestures by human neonates", *Science* 198, 75-7.
- MÉNARD, L. (2003): "Acoustic variability and adaptive articulatory strategies during vocal tract growth revealed by the rounding contrast in French", *15<sup>th</sup> International Congress of Phonetic Sciences* 3, 3169-3172.
- MÉNARD, L. (2005): "Production-perception relationship development", *Journal of the Acoustical Society of America* 117(4), pt. 2, 2398.

- MÉNARD, L. y BOË, L-J. (2004): “L'emergence du système phonologique chez l'enfant: l'apport de la modélisation articulatoire”, *Canadian Journal of Linguistics/Revue canadienne de linguistique* 49(2), 155-174.
- MÉNARD, L., SCHWARTZ, J-L y BOË, L-J (2004): “Role of vocal tract morphology in speech development: perceptual targets and sensorimotor maps for synthesized French vowels from birth to adulthood”, *Journal of Speech, Language and Hearing Research* 47, 1059-1079.
- MERMELSTEIN, P. (1967): “Determination of the vocal-tract shape from measured formant frequencies”, *Journal of the Acoustical Society of America* 41, 1283-1294.
- MERMELSTEIN, P. (1978): “Difference limens for formant frequencies of steady-state and consonant-bound formants”, *Journal of the Acoustical Society of America* 63, 572-580.
- MEROW, W. W. y BROADBENT, B. H. (1990): Cephalometrics, in *Facial Growth* 3<sup>rd</sup> ed. Edited by d. H. Enlow (Saunders, Philadelphia).
- MICKENS, C. W. (1897-1898): “Practical results of child-study”, *Child Study Monthly* 3, 196-205.
- MILLER, J. D. (1953): “Auditory tests with synthetic vowels”, *Journal of the Acoustical Society of America* 18, 114-121.
- MILLER, J. D. (1987): “Auditory-perceptual interpretation of the vowel”, *Journal of the Acoustical Society of America* 85(5), 2114-2134.
- MILLER, J. D., LEE, S., UCHANSKI, R. M., HEIDBREder, A. H., RICHMAN, B. B. y TADLOCK, J. (1996): Creation of two children's speech databases, *Proceedings of the ICASSP* (Atlanta, GA), pp. 849–852.
- MIYAWAKI, K., STRANGE, W., VERBRUGGE, R. R., LIBERMAN, A. M., JENKINS, J. J. y FUJIMURA, O. (1975): “An effect of linguistic experience: the discrimination of (r) and (l) by native speakers of Japanese and English”, *Perception and Psychophysics* 18, 331-340.
- MOFFITT, A. R. (1971): “Consonant cue perception by twenty- to twenty-four week old infants”, *Child Development* 42, 717 -731.
- MOL, H. (1963): *Fundamentals of Phonetics*, Jallua Linguorum, N 26, The Hague.

- MONSEN, R. B. y ENGEBRETSON, A. M. (1983): "The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear prediction", *Journal of Speech and Hearing Research* 26, 89-97.
- MOON, C. (2009): "Newborn infant perception of vowels is affected by ambient language", *Journal of the Acoustical Society of America* 125(4), 2778.
- MOON, C., COOPER, R. P. y FIFER, W. P. (1993): "Two-day-olds prefer their native language", *Infant Behavior and Development* 16, 495-500.
- MOORE, K. C. (1895-1897): "The mental development of a child", *Psicol. Monog.* 1(3), pp. 150.
- MOORE, D. R. (2002): "Auditory development and the role experience", *British Medical Bulletin* 63, 171-181.
- MOORE, C. A. y RUARK, J. L. (1996): "Does speech emerge from earlier appearing oral motor behaviors?", *Journal of Speech and Hearing Research* 39, 1034-1047.
- MOORE, C. A., CAULFIELD, T. J. y GREEN, J. R. (2001): "Relative kinematics of the rib cage and abdomen during speech and nonspeech behaviors of 15-month-old children", *Journal of Speech, Language and Hearing Research* 44, 80-94.
- MOROSAN, D. E. y JAMIESON, D. G. (1989): "Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task", *Journal of Speech and Hearing Research* 32, 501-511.
- MORRONGIELLO, B. A., ROBSON, R., C., BEST, C. T. y CLIFTON, R. K. (1984): "Trading relations in the perception of speech by 5-year-old children", *Journal of Exp. Child Psych.* 37, 231-250.
- MORRONGIELLO, B. A. y TREHUB, S. E. (1987): "Age-related changes in auditory temporal perception", *Journal of Experimental Child Psychology* 44, 413-426.
- MORSE, P. A. (1978): "Infant speech perception: Origins, processes and alpha centauri". En: MINIFIE, F. D. y LLOYD, L. L. (eds.). *Communicative and cognitive abilities. Early behavioral assessment*, Baltimore, MD: University Park Press.
- MOSKOWITZ, A. I. (1970): "The two-year-old stage in the acquisition of English, *Lg.* 46, 426-441.



- MOWRER, D. E. (1980): "Phonological development during the first year of life". En: LASS, N. (ed.). *Speech and Language. Advances in Basic Research and Practice*, Academic Press.
- MOXLEY, S. E. (1979): "Schema: The variability of practice hypothesis", *Journal of Motor Behavior* 11, 65-70.
- MRAYATI, M. (1976): Contribution aux etudes sur la production de la parole, Thèse de doctorat, Grenoble, France.
- MRAYATI, M., CARRÉ, R. y GUÉRIN, B. (1988): "Distinctive regions and modes: a new theory of speech production", *Speech Communication* 7, 257-286.
- MUNHALL, K. G. y JONES, J. A. (1998): "Articulatory evidence for syllabic structure", *Behavioral and Brain Sciences* 21, 524-525.
- NAKAZIMA, S. (1962): "A comparative study of the speech development of Japanese and American English in childhood", *Stud. Phonol.* 2, 27-39.
- NAKAJIMA, T., OMURA, H. e ISHIZAKI, S. (1975): Estimation of vocal tract area functions by adaptive inverse filtering methods and identification of articulatory model. En: FANT, G. (ed.): *Speech communication seminar*, New York: Wiley.
- NAVARRO TOMÁS, T. (1918): *Manual de pronunciación española*, Madrid, CSIC, 27ª ed. (1999)
- NAVARRO TOMÁS, T. (1946): "Escala de frecuencia de los fonemas españoles" en *Estudios de fonología española*, págs. 15-30, Syracuse, 1ª ed. Syracuse, New York, Syracuse University 1946, 2ª ed.
- NEAREY, T. M. (1989): "Static, dynamic and relational properties in vowel perception", *Journal of the Acoustical Society of America* 85, 2088-2113.
- NEREAY, T. (1976): Phonetic features for vowels, unpublished Doctoral Dissertation, University of Connecticut, Storrs.
- NIP, I. S. B. y GREEN, J. R. (2006): The development of speaking rate: A kinematic perspective, *Paper presented at the Conference on Motor Speech Austin, TX*.

- NISHIMURA, T. (2003): “Comparative morphology of the hyo-laryngeal complex in anthropoids: two steps in the evolution of the descent of the larynx”, *Primates* 44, 41-49.
- NISHIMURA, T., MIKAMI, A., SUZUKI, J. y MATSUZAWA, T. (2003): Descent of the larynx in chimpanzee infants. *Proceedings of the National Academy of Sciences of the USA*. 100, 6930-6933.
- NISHITANI, N., SCHÜRMAN, M., AMUNTS, K. y HARI, R. (2005): “Broca's region: From Action to Language”, *Psychology* 20, 69.
- NITTROUER, S. (1992): “Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries”, *J. Phonetics* 20, 351–382.
- NITTROUER, S. (1993): “The emergence of mature gestural patterns is not uniform: Evidence from an acoustic study”, *Journal of Speech and Hearing Research* 36, 959-972.
- NITTROUER, S. (1995): “Children learn separate aspects of speech production”, *Journal of the Acoustical Society of America* 97(1), 520-530.
- NITTROUER, S. (2005): “Age-related differences in weighting and masking of two cues to word-final stop voicing in noise”, *Journal of the Acoustical Society of America* 118, 1072-1088.
- NITTROUER, S. y STUDDERT-KENNEDY, M. (1987): “The role of coarticulatory effects in the perception of fricatives by children and adults”, *Journal of Speech and Hearing Research* 32, 120-132.
- NITTROUER, S. y MILLER, M. (1997): “Developmental weighting shifts for noise components of fricative-vowel syllables”, *Journal of the Acoustical Society of America* 102, 572-580.
- NORDSTRÖM, P. E. (1975): “Attempts to simulate female and infant vocal tracts from male are functions”, *STL-QPSR* 2-3, 20-30.
- NORDSTRÖM, P. E. (1977): “Female and infant vocal tracts simulated from male area functions”, *Journal of Phonetics* 5, 81-92.
- O'SHAUGHNESSY, D. (1996): “Critique: Speech perception: Acoustic or articulatory?” *Journal of the Acoustical Society of America* 99(3), 1726-1729.

- OHALA, J. (1996): "Speech perception is hearing sounds, not tongues", *The Journal of the Acoustical Society of America* 99(3), 1718-1725.
- OHDE, R. N., HALEY, K. L. y MCMAHON, C. W. (1996): "A developmental study of vowel perception from brief synthetic consonant-vowel syllables", *Journal of the Acoustical Society of America* 100, 3813-3824.
- OHDE, R. N. y HALEY, K. L. (1997): "Stop-consonant and vowel perception in 3 and 4 year-old children", *Journal of the Acoustical Society of America* 102, 3711-3722.
- OLLER, D. K. (1978): "Infant vocalization and the development of speech", *Allied Health Behav Sci.* 1, 523-549.
- OLLER, D. K. (1980): "The emergence of the sounds of speech in infancy". En: YENIKOMSHIAN, G., KAVANAG, C. y FERGUSON, C. (eds.). *Child Phonology: perception and production*. London: Academic Press.
- OLLER, D. K. (1986): "Metaphonology and Infant vocalizations". En: LINDBLOM, B. y ZETTERSTROM, R. (eds). *Precursor of Early Speech*, Pp. 21-35, Stockton Press.
- OLLER, D. K. (2000): *The Emergence of the Speech Capacity*, Lawrence Erlbaum Associates, Mahwah, New Jersey.
- OLLER, D. K., WIEMAN, L. A., DOYLE, W. J. y ROSS, C. (1976): "Infant babbling and speech", *Journal of Child Language* 3, 145-156.
- OLLER, D. K. y EILERS, R. E. (1981): "Similarity of babbling in Spanish and English-learning babies", *Journal of Child Language* 9, 565-77.
- OLLER, D. K. y MACNEILAGE, F. (1983): "Development of speech production: perspectives from natural and perturbed speech". En: MACNEILAGE, P. (ed.). *The production of speech*, pp. 99-108 (Springer, New York).
- OLLER, D. K., EILERS, R. E., BULL D. H. y CARNEY, A. E. (1985): "Prespeech vocalizations of a deaf infant: a comparison with normal metaphonological development", *Journal of Speech and Hearing Research* 28(1), 47-63.
- OLLER, D. K. y EILERS, R. E. (1988): "The role of audition in infant babbling", *Child Dev.* 59(2), 441-9.

- OILER, D. K. y STEFFANS, M. L. (1993): "Syllables and segments in infant vocalizations and young child speech". En: YAVAS, M. (ed.). *First and second language phonology* (p. 45-62): San Diego, CA: Singular Publishing Co.
- OLMSTEAD, D. (1971): *Out of the mouth of babes*, The Hague: Mouton.
- OLNEY, R. L. y SCHOLNICK, E. K. (1975): "Adult judgments of age and linguistic differences in infant vocalization", *Journal of Child Language* 3, 145-156.
- OTOMO, K. y STOEL-GAMMON, C. (1992): "The Acquisition of unrounded vowels in English", *Journal of Speech and Hearing Research* 35, 604-616.
- PALETHORPE, S., WALES, R., CLARK, J. y SENSERRICK, T. (1996): "Vowel-classification in children", *Journal of the Acoustical Society of America* 100(6), 3843-3851.
- PANNETON, R., KITAMURA, C., MATTOCK, K. y BURNHAM, D. (2006): "Slow Speech Enhances Younger but not Older Infants' Perception of Vocal Emotion", *Human Development* 3(1), 7-19.
- PARNELL, M. M. y AMERMAN, J. D. (1978): "Maturational influences on perception of coarticulatory effects", *Journal of speech and Hearing Research* 21, 682-701.
- PASCHALL, L. (1983): "Development at 18 months". En: IRWIN, J. V. y WONG, S. P. (eds.). *Phonological Development in Children: 18 to 72 Months* (Carbondale, IL: Southern Illinois University Press), pp. 27-54.
- PENTZ, A. y GILBERT, H. (1983): Comparison of formants in preadolescent children's vowel productions. Paper presented in a poster session at 1983 Annual Convention of the American Speech-Language-Hearing Association. En: KENT, R. D. (1994). *Reference manual for communicative sciences and disorders: Speech and language*. San Antonio, TX: Pro-Ed. P. 73.
- PÉREZ, B. (1885): *The first three years of childhood*. Trans by Alice M. Christie, Chicago: A. N. Marquis, pp. X, 292 (pp. 234-262).
- PERKELL, J. (1969): *Physiology of speech Production: Results and Implications of a Quantitative Cineradiographic Study*, Cambridge, Mass: M.I.T. Press.

- PERKELL, J. (1980): "Phonetic features and the physiology of speech production". En: BUTTERWORTH, B. (ed.). *Language Production*. Vol.1. London: Academic Press; 1980. pp. 337-372.
- PERRIER, P. (2005): "Control and representations in speech production", *ZAS Papers in Linguistics* 40, 109-132.
- PERRIER, P., PAYAN, Y. y NAZARI, M. (2011): Biomechanical models of speech articulators to study speech motor control, "The 6th International Conference on Speech Motor Control, Groningen: Netherlands (2011)".
- PERRY, T. L., OHDE, R. N. y ASHMEAD, D. H. (2001): "The acoustic bases for gender identification from children's voices", *Journal of the Acoustical Society of America* 109 (6), 2988-2998.
- PETERSON, G. E. (1961): "Parameters of vowel quality", *Journal of Speech and Hearing Research* 4, 10-29.
- PETERSON, G. E. y BARNEY, H. L. (1952): "Control methods used in a study of the vowels", *Journal of the Acoustical Society of America* 24, 175-184.
- PETITTO, L. A. (1993): "On the ontogenetic requirements for early language acquisition". En: de BOYSSON-BARDIES, B., de SCHONEN, S., JUSCYK, P., MACNEILAGE, P. y MORTON, J. (eds.). *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 365-383): Dordrecht, Holland: Kluwer Academic Publishers.
- PIAGET, J. (1955): *The Language and Thought of the Child*, Translated by M. Gabain, Cleveland, Meridian Books.
- PICKETT, J. M. (1980): *Perspectives in audiology series. The sounds of speech communication: A primer of acoustic phonetics and speech perception*, Baltimore: University Park Press.
- PIERREHUMBERT, J. (2011): "Phonetic diversity, statistical learning and acquisition of phonology", *Language and Speech* 46 (2-3), 115-154.
- PISONI, D. B. (1971): On the nature of categorical perception of speech sounds. *Status report on speech research* (SR-27): New Haven: Haskins Laboratories, p. 101.
- PISONI, D. B. (1973): "Auditory and phonetic memory codes in the discrimination of consonants and vowels", *Percept. Psychophys.* 13, 253-260.

- PISONI, D. B. (1975): "Auditory short-term memory and vowel perception", *Memory and Cognition*, 3(1), 7-18.
- PISONI, D. B., ASLIN, R. N., PEREY, A. J. y HENNESSY, B. L. (1982): "Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants", *Journal of Experimental Psychology: Human perception and performance* 8, 297-314.
- PLOMP, R. (1975): "Auditory analysis and timbre perception". En: FANT, G. y TATHAM, M. A. A. (eds.): *Auditory Analysis and Perception of Speech*, (Academic London), pp. 7-22.
- PLOOG, D. (1979): "Phonation, emotion, cognition, with reference to the brain mechanisms involved", *CIBA Foundation Symposium* 69, 79-98.
- POLGAR, G. y WENG, T. (1979): "The functional development of the respiratory system: From the period of gestation to adulthood", *Am. Rev. Respir. Dis.* 120, 625-695.
- POLGAR, G. y WENG, T. (1979): *Pulmonary function testing in children: Techniques and standards*. Philadelphia: W.B. Saunders.
- POLKA, L. y WERKER, J. F. (1994): "Developmental changes in perception of nonnative vowel contrast", *Journal of Experimental Psychology: Human perception and Performance* 20(2), 421-435.
- POLKA, L. y BOHN, O-S. (1996): "A cross-language comparison of vowel perception in English-learning and German-learning infants", *Journal of the Acoustical Society of America* 100(1), 577-592.
- POLKA, L. y BOHN, O-S. (2003): "Asymmetries in vowel perception", *Speech and Communication* 41, 221-231.
- POLKA, L., RVACHEW, S. y MATTOCK, K. (2007): "Experiential influences on speech perception and production in infancy". En: HOFF, E. y SHATZ, M. (eds.). *Handbook of Language Development*, Blackwell Publishing.
- POLKA, L. y BOHN, O-S (2011): "Natural referent vowel (NRV) framework: an emerging view of early phonetic development", *Journal of Phonetics* 39, 467-478.
- POLLOCK, F. (1878): "An infant's progress in language", *Mind* 3, 392-401.

- POLLOCK, K. E. y KEISER, N. J. (1990): "An examination of vowel errors in phonologically disordered children", *Clinical Linguistic and Phonetics* 4, 161-178.
- POLLOCK, K. E. y HALL, P. K. (1991): "An analysis of the vowel misarticulations of the children with development apraxia of speech", *Clinical Linguistics and Phonetics* 5, 207-224.
- POLS, L. C. W., LYAKSO, E., van der STELT, J. M., WEMPE, T. G. y ZAJDÓ, K. (2006): Vowel data of early speech development in several languages, Proc. Multiling 2006, Stellenbosch.
- PRESTON, M., YENI-KOMSHIAN, G. y STARK, R. (1967): Voicing in initial stop consonants produced by children in the prelinguistic period from different language communities. Johns Hopkins University School of Medicine, *Annual Report of Neurocommunications Laboratory* 2, 305-323.
- PYE, C., INGRAM, D. y LIST, H. (1987): "A comparison of initial consonant acquisition in English and Quiche". En: NELSON, K. E. y VAN KLEEK, A. (eds). *Children's language: Vol. 6*, Hillsdale, NJ: Lawrence Erlbaum Associates.
- QUERLEU, D., RENARD, S. y VERSYP, F. (1981): "Les perceptions auditives du foetus humain" *Medecine et Hygiene* 39, 2101-2110.
- QUERLEU, D., BOUTTEVILLE, C., RENARD, X. y CREPIN, G. (1985): "Sound stimulation test and fetal well-being", *American Journal of Obstetrics and Gynecology* 151 (6), 829-30.
- QUILIS, A. y ESGUEVA, M. (1980): "Frecuencia de los fonemas en el español hablado", *Lingüística Española Actual*, II, 1-25.
- QUILIS, A. y ESGUEVA, M. (1983): "Realización de los fonemas vocálicos españoles en posición fonética normal". En: ESGUEVA, M. y CANTARERO, M. (eds.). *Estudios de Fonética I*, Madrid, CSIC.
- REMEZ, R. E. (1995): "Critique: auditory form and gestural topology in the perception of speech", *Journal of the Acoustical Society of America* 99(3), 1695-1698.
- REPP, B. H., HEALEY, A. F. y CROWDER, R. G. (1979): "Categories and context in the perception of isolated steady-state vowels", *J. Exp. Psychol. Human Percept. Perform* 5, 129-145.

- REPP, B. H. y CROWDER, T. G. (1990): "Stimulus order effects in vowel discrimination", *Journal of the Acoustical Society of America* 88, 2080-2090.
- RIZZOLATTI, G., FADIGA, L., GALLESE, V. y FOGASSI, L. (1996a): "Premotor cortex and the recognition of motor actions", *Brain Res Cogn Brain Res* 3, 131-141.
- RIZZOLATTI, G., FADIGA, L., MATELLI, M., BETTINARDI, V., PAULESU, E., PERANI, D. y FAZIO, F. (1996b): "Localization of grasp representations in humans by PET: 1. Observation versus execution", *Experimental Brain Research* 111, 246-252.
- RIZZOLATTI, G., FOGASSI, L. y GALLESE, V. (1997): "Parietal cortex: from sight to action", *Current Opinion in Neurobiology* 7, 562-567.
- ROBB, M. y CACACE, A. (1995): "Estimates of formant frequencies in infant cry", *Int. J. Pediatr. Otorhinolaryngol.* 32, 57-67.
- ROBB, M., CHEN, Y. y GILBERT H. R. (1997): "Developmental aspects of formant frequency and bandwidth: 4-25 months", *Folia phoniatrica et logopaedica: official organ of International Association of Logopedics* 49, 88-95.
- ROCHAT, P. (1989): "Object manipulation and exploration in 2- to 5-month-old infants", *Developmental Psychology* 25, 871-884.
- ROCHE, A. F. y BARKLA, D. H. (1965): "The level of the larynx during childhood", *Ann. Otol. Rhinol. Laryngol.* 74, 645-654.
- RUARK, J. L. y MOORE, C. A. (1997): "Coordination of lip muscle activity by 2-year-old children during speech and nonspeech task", *Journal of Speech, Language, and Hearing Research* 40, 1373-1385.
- RUCH, F. L. y ZIMBARDO, P. G. (1971): *Psychology and life* (8<sup>th</sup> ed.) Glenview, III.: Scott, Foresman.
- RUFF, H. A., SALTARELLI, L. M., CAPOZZOLI, M. y DUBINER, K. (1992): "The differentiation of activity in infants' exploration of objects", *Developmental Psychology* 28, 851-86.
- RUSSELL, G. O. (1928): *The vowel*, Ohio State University Press, Columbus, Ohio.



- RUZZA, B., ROCCA, F., BOERO, D. L. y LENTI, C. (2003): "Investigating the musical quality of early infant sounds", *Annals of the New York Academy of Sciences* 999, 527-529.
- RVACHEW, S. y SLAWINSKI, E. (1995): "The acoustic characteristics of babble produced by infants with and without early onset otitis media with effusion", *Journal of the Acoustical Society of America* 97(5), 3363.
- RVACHEW S., SLAWINSKI E. B., WILLIAMS M. y GREEN C. L. (1996): "Formant frequencies of vowels produced by infants with and without early onset otitis media", *Canadian Acoustics* 24, 19-28.
- RVACHEW, S., MATTOCK, K., POLKA L. y MÉNARD, L. (2006): "Developmental and cross-linguistic variation in the infant vowel space: The case of Canadian English and Canadian French", *Journal of Acoustical Society of America* 120(4), 2250-2259.
- RVACHEW, S., ALHAIDARY, A., MATTOCK, K. y POLKA, L. (2008): "Emergence of the corner vowels in the babble produced by infants exposed to Canadian English or Canadian French", *Journal of Phonetics* 36, 564-77.
- SAPIENZA, C. M., RUDDY, B. H. y BAKER, S. (2004): "Laryngeal structure and function in the pediatric larynx: Clinical application", *Language Speech and Hearing Services in Schools* 35, 299-307.
- SASAKI C. T., LEVINE, P. A., LAITMAN, J. T. y CRELIN, E. S. (1960): "Postnatal descent of the epiglottis in a man", *American Medical Association*, 1960-[c1985], Chicago, 1977, Mar, 103(3), 169-171.
- SCAMMON, R. E. (1930): The measurement of the body in childhood. En: HARRIS, J. A. JACKSON, C. M., PATTERSON, D. G. y SCAMMON, R. E. (eds.). *The measurement of man* (pp. 173-215), Minneapolis: University of Minnesota Press.
- SCHENEIDER, B., TREHUB, S., MORRONGIELLO, B. y THORPE, L. (1986): "Auditory sensitivity in preschool children", *Journal of the Acoustical Society of America* 79, 447-452.
- SCHNEIDER, B. y TREHUB, S. (1992): "Sources of Developmental Change in Auditory Sensitivity". En: WERNER, L. y RUBEL, E. (eds.). *Developmental Psychoacoustics*, American Psych. Assoc., Washington, D.C.

- SCHROEDER, M. R. (1967): "Determination of the geometry of the human vocal tract by acoustic measurements", *Journal of the Acoustical Society of America* 41, 1002-1010.
- SCHROEDER, M. R. (1975): Models of hearing, *Proc. IEEE* 63, 1332-1450.
- SCHROEDER, M. R., ATAL, B. S. y HALL, J. L. (1979): "Objective measure of certain speech signal degradations based on masking properties of human auditory perception". En: LINDBLOM, B. y OHMAN, S. (eds.). *Frontiers of Speech Communication Research*, Academic, London, pp. 217-229.
- SCHWARTZ, J. L. y ESCUDIER, P. (1989): "A strong evidence for the existence of a large scale integrated spectral representation in vowel perception", *Speech Comm.* 8, 235-259.
- SCHWARTZ, J. L., ABRY, C., BOË, L. J., MÉNARD, L. y VALLÉE, N. (2005): "Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory", *Speech Communication* 45, 425-434.
- SCOTT, B. L. (1976): "Temporal factors in vowel perception", *Journal of the Acoustical Society of America* 60, 1354-1365.
- SELBY, J. C., ROBB, M. P. y GILBERT, H. R. (2000): "Normal vowel articulations between 15 and 36 months of age", *Clinical Linguistics and Phonetics* 14(4), 255-265.
- SENECAIL, B. (1979): L'Os Hyhode: Introduction Anatomique a l'Etude de Certains Mechanisms de la Phonation, *Mémoires du Laboratoire d'Anatomie et de la Faculté de Médecine de Paris*, Paris.
- SERKHANE, J., SCHWARTZ, J. L., BOË, L. J., DAVIS, B. L. y MATYEAR, C. L. (2002): Motor specifications of a baby robot via the analysis of infant's vocalizations. ICSLP'2002, Denver, Colorado, 45-48.
- SERKHANE, J. E., SCHWARTZ, J. L., BOË, L. J., DAVIS, B. L. y MATYEAR, C. L. (2007): "Infant's vocalizations analyzed with an articulatory model: A preliminary report", *Journal of Phonetics* 35, 321-340.
- SERRA, M., SERRAT, E., SOLÉ, R., BEL, AURORA y APARICI, M. (2000): *La adquisición del lenguaje*, Barcelona, Ariel.
- SHANKWEILER, D., STRANGE, W. y VERBRUGGE, R. R. (1977): "Speech and the problem of perceptual constancy". En: SHAW, R. y BRANSFORD, J. (eds.).

*Perceiving, Acting, and Knowing: Toward an Ecological Psychology*, (Lawrence Erlbaum, Hillsdale, NJ).

- SHARKEY, S. G. y FOLKINS, J. W. (1985): “Variability of lip and jaw movements in children and adults: Implications for the development of speech motor control”, *Journal of Speech and Hearing Research* 28, 8-15.
- SHARP, J. T., DRUZ, W. S., BALAGOT, R. C., BANDELIN, V. R. y DANON, J. (1970): “Total respiratory compliance in infants and children”, *Journal of Applied Physiology* 29, 775-779.
- SHERRINGTON, C. S. (1906): *The Integrative Action of the Nervous System*, New Haven: Yale University Press.
- SINCLAIR, D. (1985): *Human Growth after Birth*, ed. 4, Oxford, Oxford Press.
- SINGH, S. y BLACK, J. W. (1966): “Study of twenty-six intervocal consonants as spoken and recognized by four language groups”, *Journal of the Acoustical Society of America* 39, 371-387.
- SINNOTT, J. M. y ASLIN, R. N. (1985): “Frequency and intensity discrimination in human infants and adults”, *Journal of the Acoustical Society of America* 78, 1986–1992.
- SLOAN, R. F. (1967): “Neuronal histogenesis, maturation and organization related to speech development”, *J. Comm. Disord.* 1, 1-15.
- SMITH, A. (2010): “Development of neural control of orofacial movements for speech”. En: HARDCASTLE, W., LAVER, J. y GIBBON, F. (eds.). *Handbook of Phonetic Sciences* (2ªed.), Oxford: Blackwell.
- SMITH, B. L. y GARTEBERG, T. E. (1984): “Initial observations concerning developmental characteristics of labio-mandibular kinematics”, *Journal of the Acoustical Society of America* 75, 1599-1605.
- SMITH, B. L. y KENNEY, M. K. (1994): “Assessing duration and variability in speech and non-speech tasks performed by adults”, *Journal of the Acoustical Society of America* 96, 699-705.
- SMITH, B. L. y KENNEY, M. K. (1998): “An assessment of several acoustic parameters in children's speech production development: longitudinal data”, *Journal of Phonetics* 26, 95-108.

- SMITH, A., y GOFFMAN, L. (2004): "Interaction of motor and language factors in the development of speech". En: MAASSEN, B., KENT, R., PETERS, H., van LIESHOUT, P. y HULSTIJN, W. (eds.). *Speech Motor Control in Normal and Disordered Speech* (pp. 227-252): Oxford, England: Oxford University Press.
- SMITH, A. y ZELAZNIK, H. N. (2004): "Development of functional synergies for speech motor coordination in childhood and adolescence", *Developmental Psychobiology* 45, 22-33.
- SOMMERS, M. S., MOODY, D. B., PROSEN, C. A. y STEBBINS, W. C. (1992): "Formant frequency discrimination by Japanese macaques (*Macaca fasciata*)", *Journal of the Acoustical Society of America* 91(6), 3499-3510.
- SONDHI, M. M. (1979): "Estimation of vocal-tract areas: The need for acoustical measurements", *IEEE Transactions on Acoustic, Speech and Signal Processing* 27, 268-273.
- SONDHI, M. M. y GOPINATH, B. (1971): "Determination of vocal-tract shape from impulse response of the lips", *Journal of the Acoustical Society of America* 49, 1867-1873.
- STARK, R. E. (1980): "Stages of speech development in the first year of life". En: YENIKOMSHIAN, G. H., KAVANAGH, J. F. y FERGUSON, C. A. (eds.). *Child Phonology: Vol. 1, Production*, (Academic New York) vol., 1, pp. 73-92.
- STARK, R. E., ROSE, S. N. y MCLAGEN, M. (1975): "Features of infant sounds: The first eight weeks of life", *Journal of Child Language* 2, 205-21.
- STARK, R. E., ROSE, S. N. y BENSON, P. J. (1978): "Classification of infant vocalization", *British Journal of Communication Disorders* 13, 41-7.
- STATHOPOULOS, E. T. (1995): "Variability revisited: An acoustic, aerodynamic, and respiratory kinematic comparison of children and adults during speech", *Journal of Phonetics* 23, 67-80.
- STATHOPOULOS, E. T. y SAPIENZA, C. (1993): "Respiratory and laryngeal measures of children during vocal intensity variation", *Journal of the Acoustical Society of America* 94(5), 2531-2543.

- STEEVE, R. W., MOORE, C. A., GREEN, J. R., REILLY, K. J. y RUARK, J. (2008): “Babbling, chewing, and sucking: Oromandibular coordination at 9 months”, *Journal of Speech, Language and Hearing Research* 51, 1390-1404.
- STELT, Van der, J. M., ZAJDÓ, K. y WEMPE, T. (2005): “Exploring the acoustic vowel space in two-year-old children: Results for Dutch and Hungarian”, *Speech Communication* 47, 143-159.
- STELT, Van der J. M., LYAKSO, E., POLS, L. C. W. y WEMPE, A. G. (2006): Developing a Russian vowel space in infancy: the first two years, 6-page paper presented at the SpeCom 2006 Conference in St. Petersburg, Russia, 28th of June 2006.
- STELT, Van der J. M., WEMPE, A. G. y POLS, L. C. W. (2008): “Comparing Deaf and Hearing Dutch Infants: changes in the vowel space in the first two years”, *J. Clinical Linguistics & Phonetics* 5, 1-10.
- STERN D. N., SPIEKER, R. K., BARNETT R. K. y MACKAIN K. (1983): “The prosody of maternal speech: infant age and context related changes”, *Child Language* 10(1), 1-15.
- STEVENS, K. (1972): “The quantal nature of speech: Evidence from articulatory-acoustic data”. En: DAVIS, E. E. y DENES, P. B. (eds.). *Human Communication: A Unified View*. McGraw-Hill, New York, pp. 51–66.
- STEVENS, K. (1983): “Design Features of Speech Sound Systems”. En: MACNEILAGE, P. F. (ed.). *The Production of Speech*. New York: Springer Verlag, pp. 247-261
- STEVENS, K. (1989): “On the quantal nature of speech”, *Journal of Phonetics* 17, 3-45.
- STEVENS, K. (1997): “Articulatory-acoustic-auditory relationships”. En: HARDCASTLE, W. J. y LAVER, J. (eds.): *The Handbook of Phonetic Sciences*, Blackwell: Oxford, 462-506.
- STEVENS, K. (1998): *Acoustic Phonetics*, MIT Press, Cambridge, Massachusetts.
- STEVENS, K. y HOUSE, A. S. (1961): “An acoustical theory of vowel production and some of its implications”, *Journal of Speech and Hearing Research* 4, 303-320.
- STEVENS, K., LIBERMAN, A. M., STUDDERT-KENNEDY, M. y OHMAN, S. E. G. (1969): “Cross-language study of vowel perception”, *Language and Speech* 12, 1-23.

- STEVENS, K. y HOUSE, A. (1995): "Development of a Quantitative Description of Vowel Articulation", *Journal of the Acoustical Society of America*, 27(3), 484-493.
- STOEL-GAMMON, C. y COOPER, J. A. (1984): "Patterns of early lexical and phonological development", *Journal of Child Language* 11, 247-271.
- STOEL-GAMMON, C. y HERRINGTON, P. B. (1990): "Vowel systems of normally developing and phonologically disordered children", *Clinical Linguistics and Phonetics*, 4, 145-160.
- STOKES, S. F. y WONG, I. M. (2002): "Vowel and diphthong development in Cantonese-speaking children", *Clinical Linguistics and Phonetics* 16(8), 597-617.
- STRANGE, W., VERBRUGGE, R. y SHANKWEILER, D. (1974): "Consonantal environment specifies vowel identity", *Journal of the Acoustical Society of America* 55, (S1):S54.
- STRANGE, W., EDMAN, T. R. y JENKINS, J. J. (1979): "Acoustic and phonological factors in vowel identification", *J. Exp. Psychol: Human Percept. Perf.* 5, 643-656.
- STUDDERT-KENNEDY, M. (1986): "Sources of variability in early speech development". En: PERKELL, J. S. y KLATT, D. H. (eds.). *Invariance and variability in speech processes*. Hillsdale NJ: Erlbaum.
- SUNDBERG, U. (1998): Mother tongue-phonetic aspects of infant-directed speech. Doctoral Dissertation, PERILUS, Stockholm.
- SUSSMAN, J. (2001): "Vowel perception by adults and children with normal language and specific language impairment: based on steady states or transitions?", *Journal of the Acoustical Society of America* 109(3), 1173-1180.
- SUSSMAN, J. E. y CARNEY, A. E. (1989): "Effects of transition length on the perception of stop consonants by children and adults", *Journal of Speech and Hearing Research* 32, 151-160.
- SUSSMAN, J. E. y LAUCKNER-MORANO, V. J. (1995): "Further tests of the perceptual magnet effect in the perception of [i]: Identification and change/no change discrimination", *Journal of the Acoustical Society of America* 97, 539-552.
- SWEENEY, S. (1973): "The importance of imitation in the early stages of speech acquisition: a case report", *J. Speech Hearing Dis.* 38, 490-494.

- SWOBODA, P. J., MORSE, P. A. y LEAVITT, L. A. (1976): "Continuous vowel discrimination in normal and at risk infants", *Child Development* 47, 459-465.
- SYRDAL, A. y GOPAL, H. (1986): "A perceptual model of vowel recognition based on the auditory representation of American English vowels", *Journal of the Acoustical Society of America* 79(4), 1086-1091.
- SZE C. P. (2000): The Development of Vowel Production in Cantonese-speaking Infants. Bachelor Thesis. University of Hong Kong.
- TAINE, M. (1877): "The acquisition of language by childrens", *Mind* 2, 252-259.
- TAKEUCHI, S., KASUYA, H. y KIDO, K. (1974): "An active model for extraction of nasality and its perceptual evaluation". En: FANT, G. (ed.). *Speech Communication: vol. 3. Speech Perception and Automatic Recognition* (Almqvist and Wiksell, Stockholm), pp. 141-148.
- TAKEUCHI, S., KASUYA, H. y KIDO, K. (1975): "On the acoustic correlate of nasality", *Journal of the Acoustical Society of America*, Jpn. 31, 298-309.
- TARQUINIO, N., ZELAZO, P. R. y WEISS, M. J. ( 1990 ): "Recovery of neonatal headturning to decreased sound pressure level", *Developmental Psychology* 26, 752-758.
- TEES, R. C. y WERKER, J. F. (1984): "Perceptual flexibility: Maintenance or recovery of the ability to discriminate nonnative speech sounds", *Canadian Journal of Psychology* 38, 579-590.
- TEMPLIN, M. (1957): Certain language skills in children: their development and interrelationships. *Institute of Child Welfare Monographs*, 26, (Minneapolis: University of Minnesota Press).
- TEMPLIN, M. y DARLEY, F. (1969): *Templin-Darley Test of Articulation*. Iowa City, IA: Bureau of Educational Research and Service University of Iowa.
- THEVENIN, D., EILERS, R. E., OLLER, D. K. y LAVOIE, L. (1985): "Where's the drift in babbling drift? A cross-linguistic study", *Applied Psycholinguistics* 6, 3-15.
- THOM, S., HOIT, J., HIXON, T. y SMITH, A. (2005): "Velopharyngeal function during vocalization in infants", *Cleft Palate-Craniofacial Journal* 43, 539-546.

- THOMPSON, A. E. y HIXON, T. J. (1979): "Nasal air flow during normal speech production" *Cleft Palate Journal* 16, 412-420.
- TINGLEY, B. y ALLEN, G. (1975): "Development of speech timing control in children", *Child Development* 46, 186-194.
- TITZE, I. (1989): "Physiologic and acoustic differences between male and female voices", *Journal of the Acoustical Society of America* 85, 1699-1707.
- TITZE, I. (1994): *Principles of Voice Production* (Prentice-Hall, EnglewoodCliffs, NJ).
- TRACY, F. (1895): *The psychology of childhood*, 3d ed, Boston: D. C. Heath, pp. XIII, 170 (pp. 114-151).
- TRAUNMÜLLER, H. (1981): "Perceptual dimension of openness in vowels", *Journal of the Acoustical Society of America* 69(5), 1465-1475.
- TREHUB, S. E. (1973): "Infant's sensitivity to vowel and tonal contrasts", *Developmental psychology* 9(1), 91-96.
- TREHUB, S. E. (1976): "The discrimination of foreign speech contrasts by infants and adults", *Child Development* 47, 466-472.
- TREHUB, S. y RABINOVITCH, M. (1972): "Auditory-Linguistic Sensitivity in Early Infancy", *Developmental Psychology* 6(1), 74-77.
- TREHUB, S. E., THORPE, L. A., y MORRONGIELLO, B. A. (1987): "Organizational processes in infants' perception of auditory patterns", *Child Development* 58, 741-749.
- TREHUB, S., CHENIDER, B., MORRONGIELLO, B. y THORPE, L. (1988): "Auditory sensitivity in school-age children", *J. of Exp. Child Psych.* 46, 273-285.
- TREHUB, S. E., ENDMAN, M. W. y THORPE, L. A. (1990): "Infants perception of timbre: Classification of complex tones by spectral structure", *Journal of experimental child psychology* 49, 300-313.
- TSE, A. C. (1991): *The Acquisition Process of Cantonese Phonology: a case study*. Unpublished doctoral dissertation, University of Hong Kong.
- TYLER, A. A. y LANGSDALE, T. E. (1996): "Consonant-vowel interactions in early phonological development", *First Language* 16, 159-191.



- VERBRUGGE, R., STRANGE, W. y SHANKWEILER, D. (1974): What information enables a listener to map a talker's vowel space? *Haskins Status Report*, SR-37/38, 199-208.
- VIHMAN, M. (1992): "Early syllables and the construction of phonology". En: FERGUSON, C. A., MENN, L. y STOEL- GAMMON, C. (eds.). *Phonological development: Models, Research, Implications* (p.393-422): Maryland: York Press.
- VIHMAN, M. (1996): *Phonological development. The origins of language in the child*. Blackwell Publ.
- VIHMAN, M. M. y MILLER, R. (1988): "Words and babble at the threshold of language acquisition". En: SMITH, M. D. y LOCKE, J. L. (eds.). *The Emergent Lexicon: The Child's Development of Linguistic Vocabulary*, (Academic, New York), pp. 151-183.
- VIHMAN, M. y MCCUNE, L. (1994): "When is a word a word", *Journal of Child Language* 21, 517-542.
- VORPERIAN, H. K. (2000): Anatomic Development of the Vocal Tract Structures as Visualized by MRI. Ph. D. dissertation, University of Wisconsin-Madison.
- VORPERIAN, H. K., KENT, R. D., LINDELL, R. G. y YANDELL, B. S. (1999): "Magnetic resonance imaging procedures to study the concurrent anatomic development of vocal tract structures preliminary results", *International Journal of Pediatric Otorhinolaryngology* 49, 197-206.
- VORPERIAN, H. K., KENT, R. D., LINDSTROM, M. J., KALINA, C. M., GENTRY, L. R. y YANDEL, B. S. (2005): "Development of vocal tract length during early childhood: a magnetic resonance imaging study", *Journal of the Acoustical Society of America* 117(1), 338-350.
- VORPERIAN, H. K. y KENT, R. D. (2007): "Vowel acoustic space development in children. A Synthesis of Acoustic and Anatomic data", *Journal of Speech Language and Hearing Research*, Scholar one, 1-83.
- VORPERIAN, H. K., WANG, S., CHUNG, M. K., SCHIMEK, E. M., REID, B., DURTSCHI, R. B. y KENT, R. D. (2008): "Anatomic development of the oral and pharyngeal portions of the vocal tract: an imaging study", *Journal of the Acoustic Society of America* 125(3), 1666-1678.

- WAKITA, H. (1973): "Direct estimation of the vocal-tract shape by inverse filtering of acoustic speech waveform", *IEEE Transactions on Audio and Electroacoustics*, 21(5), 417-427.
- WAKITA, H. (1979): "Estimation of vocal-tract shapes from acoustical analysis of the speech wave: The state of the art", *IEEE Transactions on Acoustics, Speech and Signal Processing*, 27(3).
- WAKITA, H. y GRAY, A. H. (1975): "Numerical determination of the lip impedance and vocal tract area functions", *IEEE Transactions on Acoustics, Speech and Signal Processing* 23(6), 574-580.
- WALKER, J. (1994): "Craniofacial changes". En: PINKHAM, J. R., CASAMASSIMO, P. S, TIGUE, D. J., FIELDS, H. W. y NOWAK, A.(eds.). *Pediatric Dentistry: Infancy Through Adolescence*, (Saunders, Philadelphia).
- WALLEY, A. C., PISONI, D. B. y ASLIN, R. N. (1981): The role of early experience in the development of speech perception. En ASLIN, R. N., ALBERTS, J. R. y PETERSEN, M. R. (eds.). *Development of Perception: vol. 1* (Academic, New York), pp. 219-256.
- WALSH, B. y SMITH, A. (2002): "Articulatory movements in adolescents: Evidence for protracted development of speech motor control processes", *Journal of Speech, Language, and Hearing Research* 45, 1119–1133.
- WALTON, G. E., SHOUP-PECENKA, A. G. y BOWER, T. G. (1991): "Speech categorization in infants: Newborns can detect phonetic invariance across talkers", *Journal of the Acoustical Society of America* 90, 2296.
- WATERSON, N. (1971): "Child phonology: A prosodic view", *Journal of Linguistics* 7, 179.
- WATKINS, K. y FROMM, D. (1984): "Labial coordination in children: Preliminary considerations", *Journal of the Acoustical Society of America* 75, 629-632.
- WATSON, J. B. (1925): *Behaviorism*, New Cork: Peoples Institute Publishing, Co. Pp. 251 (pp. 182-183).
- WEIJER, J. van de (2001): "Vowels in infant and adult directed speech", Lund University, Dept. of Linguistics, *Working Papers* 49, 172-175.
- WEIR, R. (1966): "Some questions on the child's learning of phonology". En: SMITH, F. y MILLER, G. (eds.). *The genesis of language*, Cambridge, Mass: M. I. T.

- WELLMAN, B. L., CASE, I. M., MENGERT, I. B. y BRADBURY, D. E. (1931): "Speech sounds of young children", *University of Iowa Studies in Child Welfare* 2, pp. 5.
- WERKER, J. F. (1982): The development of cross language speech perception: the effect of age, experience and context on perceptual organization. Unpublished doctoral dissertation, University of British Columbia, Vancouver, B. C.
- WERKER, J. F., GILBERT, J. H. V., HUMPHREY, K. y TEES, R. (1981): "Developmental aspects of cross-language speech perception", *Child Development* 52, 349-355.
- WERKER, J. F. y TEES, R. C. (1983): "Developmental changes across childhood in the perception of non-native speech sounds", *Canadian Journal of Psychology* 37, 278-286.
- WERKER, J. F. y TEES, R. C. (1984a): "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life", *Infant Behavior and Development* 7, 49-63.
- WERKER, J. F. y TEES, R. C. (1984b): "Phonemic and phonetic factors in adult cross-language speech perception", *Journal of the Acoustical Society of America* 75(6), 1866-1878.
- WERKER, J. F. y LOGAN, J. (1985): "Cross-language evidence for three factors in speech perception", *Perception and Psycholinguistics* 37, 35-44.
- WERKER, J. F. y LALONDE, C. E. (1988): "Cross-language speech perception: Initial capabilities and developmental change", *Developmental psychology* 24, 672-683.
- WERKER, J. F. y PEGG, J. E. (1992): "Infant speech perception and phonological acquisition". En: FERGUSON, C., MENN, L. y STOEL-GAMMON, C. (eds.). *Phonological development: Models, Research, Implications*, Timonium, Maryland, pp. 285-311.
- WERKER, J. F. y POLKA, L. (1993): "Developmental changes in speech perception: new challenges and new directions", *Journal of Phonetics* 21, 83-101.
- WESTHORPE, R. N. (1987): "The position of the larynx in children and its relationship to the ease of intubation", *Anaesth. Intensive Care* 15(3), 122-132.
- WHALEN, D. H., LEVITT, A. G. y WANG, Q. (1991): "Intonational differences between the reduplicative babbling of French- and English learning infants", *Journal of Child Language* 18, 501-16.

- WHITAKER, H. A. (1973): "Comments on the innateness of language". En: SHUY, R. W. y WASHINGTON, D. C. (eds.). *Some New Directions in Linguistics*, Georgetown University Press, 95-120.
- WHITAKER, H. A. (1976): "Neurobiology of language". En: CARTERETTE, E. C. y FRIEDMAN, M. P. (eds.). *Handbook of Perception: Language and Speech*, New York: Academic, vol. 7.
- WHITESIDE, S. P. (2001): "Sex-specific fundamental and formant frequency patterns in a cross-sectional study", *Journal of the Acoustical Society of America* 110, 464-478.
- WHITESIDE, S. P. y HODGSON, C. (2000): "Some acoustic characteristics in the voices of 6- to 10-year-old children and adults: A comparative sex and developmental perspective", *Logoped. Phoniatr. Vocol* 25(3), 122-132.
- WIND, J. (1970): *On the Phylogeny and the ontogeny of the Human Larynx* (Wolters-Noordhoff, Groningen, the Netherlands).
- WINITZ, H. e IRWIN, O. (1958): "Syllabic and phonetic structure of infant's early words", *Journal of Hearing Research* 1(3), 250-256.
- WRIGHT, J. (1980): The behavior of nasalized vowels in the perceptual vowel space, Report of the Phonology Laboratory 5 (University of California, Berkely): 127-163.
- YANG, B. (1996): "A comparative study of American English and Korean vowels produced by male and female speakers", *Journal of Phonetics* 24, 245-261.
- ZAHORIAN, S. A. y JAGHARGHI, A. J. (1993): "Spectral-shape features versus formants as acoustic correlates for vowels", *Journal of the Acoustical Society of America* 94, 1966-1982.
- ZAJDÓ, K. (2006): Vowel data of early speech development in several languages, Proc. Multiling 2006, Stellenbosch.
- ZAJDÓ, K. y POWELL, S. (2008): The acquisition of phonological vowel length in children acquiring Hungarian, *Pro. 8th Int. Seminar on Speech Production Strasbourg*, 173-176.
- ZAJDÓ, K., WEMPE, T. G., van der STELT, J. y POLS, L. C. (2011): The acquisition of Hungarian high front unrounded short vs. long vowels, *ICPhS XVII*, p. 2252-2255.

- ZIMMER, E. Z., FIFER, W. P., KIM, Y. I., REY, H. R., CHAO, C. R., y MYERS, M. M. (1993): "Response of the premature fetus to stimulation by speech sounds", *Early Human Development* 33, 207-215.
- ZIPF, G. K. y ROGERS, F. M. (1939): "Phonemes and variphones in four present day romance languages and classical latin from the viewpoint of dynamic philology", *Archives Néerlandaises de phonétique expérimentale*, XV, págs, 111-147.
- ZLATIN, M. A. y KOENIGSKNECHT, R. A. (1976): "Development of the voicing contrast: A comparison of voice onset time in stop perception and production", *Journal of Speech and Hearing Research* 19, 93-111.
- ZMARICH, C. (1997): Studio fonetico e acustico del babbling e delle prime parole di un soggetto dai 10 ai 16 mesi d'età. En: BERTINETTO, P. B. y CIONI, E. L. (a c. di). *Ati delle VIII Giornate di Studio del G.F.S, Pisa, 18-20 dicembre 1997, 1998, 212-229.*
- ZMARICH, C. y LANNI, R. (1998): "A phonetic and acoustic study of babbling in an Italian child". En: *Proceedings of 5<sup>th</sup> International Conference on Spoken Language Processing*, Sidney, 1998, vol. 6, 2703-2706.
- ZMARICH, C. y MIOTTI, R. (2003): The frequency of consonants and vowels and their co-occurrences in the babbling and early speech of Italian children, Istituto di Scienze e Tecnologie della Cognizione del C.N.R. , Sez. di Fonetica e Dialettologia, Padova, Italy Università degli Studi di Padova. 15 th ICPHS Barcelona.
- ZMARICH, C. y BONIFACIO, S. (2004): Gli inventari fonetici dai 18 ai 27 mesi d'età: uno studio longitudinale. En: ALBANO, L. F., CUTUGNO, F., PETTORINO M., SAVY R. (a cura di). *Il Parlato Italiano*. D'Auria ed.
- ZWICKER, E. (1961): "Subdivision of the audible frequency range into critical bands (Frequenzgruppen)", *Journal of the Acoustical Society of America* 33, 248.