structure of the internal representations built up by BAR 2 was matched against experimental data obtained from the experiments reported in the Chapter 3 of this work.

In this section the implementation and training of BAR 2 are described. Next, the analyses of the results in terms of learning for both the monolingual and the bilingual training are reported and discussed. The analysis of the internal representations and the comparison with empirical data are included in Chapter 3, since they were performed using the data obtained in the new experiments reported there.

4.1. Implementation

The implementation of BAR 2 followed the same direction as BAR 1, and was intended to solve the problems encountered with the first simulation. As several changes had to be made, both in the architecture of the network and in the training procedure, the new simulation was called BAR 2. The changes in the architecture (i.e., number of units) are a consequence of the new coding used for the input and output of the network. The following subsections describe the implementation of BAR 2.

4.1.1. Learning Algorithm

The learning algorithm used by BAR 2 is, as for BAR 1, the Back propagation algorithm.

4.1.2. Network Architecture

No changes were made in the architecture of BAR 1 to run BAR 2. The structure of the network is the one that can be seen in Figure 3. As in BAR 1, the number of units used for BAR 2 depended on the coding scheme used, which is described next.

4.1.3. Coding

In the previous chapter, the advantages and inconveniences of the coding scheme used for BAR 1 were described at length (see subsection 3.1.3. for the description of McWhynney and Leinbach's (1991) coding; and subsections 3.3.2. (Performance according to word length) and 3.4. (Discussion) for the discussion of results). Two major problems of the coding were identified: the spelling errors produced by the network, and the irregular coding of words of different length due to the templates scheme.

In order to fix the first inconvenience, a new coding for phonology was adopted in BAR 2, incorporating the articulatory features of the phonemes. This new coding tried to eliminate arbitrary substitutions of both letters and phonemes.

For each phoneme, the first feature coded was consonant/vowel (0 and 1, respectively). After this first code, features were different for consonants and vowels. The consonants were coded according to the following features: voiced, labial, apical, coronal, back, nasal, and continuous. Due to the fact that with this coding equivalencies some phonemes had the same code, two extra units were added to the set. These two units represented for each ambiguous pair of phonemes the phonological feature that differentiates them. The vowels were coded according to the following features: back, front, low, middle, high, round, and length. Both types of encoding can be found in Appendix 3: Phonological Coding for BAR 2.

The templates scheme had to be changed as well. In BAR 1, the double template (left- and right- justified) caused short words to be represented twice, whereas long words were not fully represented (see Subsection 3.3.2, Performance according to word length). With respect to the learning of the words, it implied that short words were learnt very fast and long words were mostly not learnt. Thus, for BAR 2 the right justified template representing the last syllable of the word was eliminated, and the left justified template was enlarged to 5 syllables with an extra group of three consonants at the end:

CCCVV CCCVV CCCVV CCCVV CCCVV CCC

This new template scheme avoided repetition in the representation of short words and allowed the full representation of long words. The length of the words in the training set was, therefore, restricted to 13 characters long (see Section 4.2.)

4.1.4. Number of units

These changes affected the structure of the network. With the new coding, 375 units were needed for the input and 372 for the output. The number is bigger than the one used for BAR 1 (237), thus the number of hidden units was increased to 110.

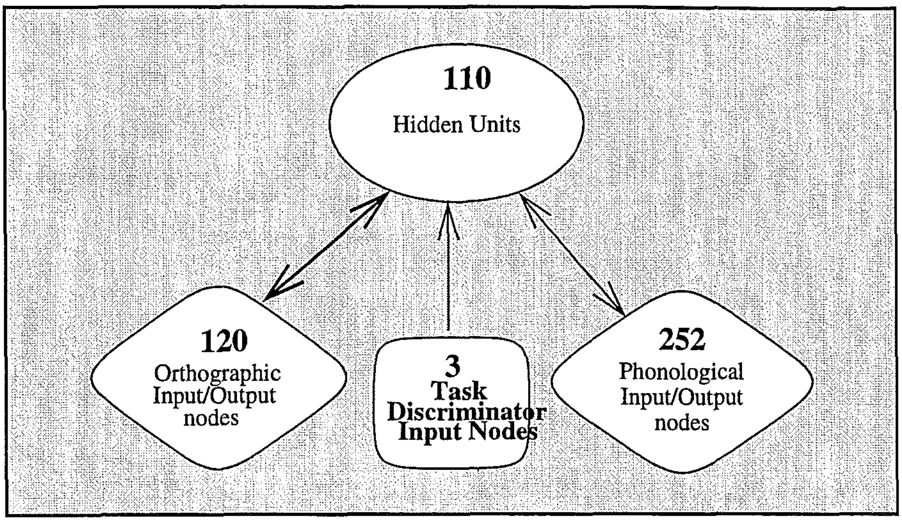

Figure 2.7. Architecture and number of units in BAR 2

4.1.4. Parameters

The learning rate of BAR 2 was set at 0.02. The reason to decrease the learning rate (in BAR 1 it was 0.5) was to prevent the error score from oscillating too much. The momentum parameter was not changed, thus set again at 0.90. BAR 2 was started too with random weights.

4.2. Training

4.2.1. Set of Training Words

Results after the bilingual training with BAR 1 showed that performance was slightly better for English words than for Dutch words, although the results were comparable. Nevertheless this result is not very desirable since the bilingual situation we aimed to model with BAR was that of an adult second language learner; accordingly, performance on Dutch words should have been better than performance on English words. For this purpose, the set of Dutch words used by BAR 2 was bigger than the set of English words. The characteristics of both sets are explained next.

For the training of BAR 2 a bigger set of Dutch words was selected from the CELEX database. Unlike the case for BAR 1, a lexical status criterion was not used. Nevertheless, words with very high frequencies (more than 1370 per million) were eliminated, thus function words were excluded. Words with extremely low frequencies (less than 6 occurrences per million) were not included in the set either. The range of frequency was narrowed with the same logarithmic transformation used in BAR 1 (see 3.2.2. Frequency), to calculate the chances of the word being presented in one epoch of training.

A second criterion for the selection was the length of the words: words had to be from 3 to 13 characters long. Words longer than 13 characters did not fit in the templates. Applying these two criteria to the set of Dutch lemmas, a file with 8074 words was obtained.

With respect to the English set, the criteria for selection were as well the frequency and the length of words. The frequency ranged between the values of 45 and 2073 occurrences per million; and the length varied from 3 to 14 characters long. These criteria yielded to a total training set of 1906 English words.

4.2.2. Pattern Presentation in the Second Phase of Learning

Following the same direction as for BAR 1, during the bilingual training the set of Dutch words was trained together with the set of English words. Due to the difference of size between the two training sets, during this phase of training there were approximately four times more Dutch words than English words. As will be seen in the next section Simulation 2: Results and analyses, this fact reduced drastically the interference of the English words in the already learnt Dutch words.

4.2.3. Software Package

The second simulation was run on a UNIX workstation with a SPARC 10 processor, as with the first one.

For BAR 2 the same modification of the original bp program of McClelland & Rumelhart, (1988) was used. The programs for coding the input and decoding the output of the network had to be new, since the encoding used in BAR 2 was different from that used for BAR 1. Thus, a program to encode the original files from the CELEX database and another to transform the output of the network into readable words were written.

4.3. Version 2: Results and Analyses

The analyses of the results after the simulation of BAR 2 follow the same structure as for BAR 1. First, the learning accuracy during both training phases is evaluated from the evolution of the error scores. Secondly, the performance of the network after monolingual and bilingual training is analysed for the total set of words with respect to word frequency and word length. A short discussion of the obtained results follows.

The analysis of the internal representations and the comparison with experimental data that evaluated the quality of the learning of BAR 2 were carried out on new empiric data from two experiments run within this project. These analyses are described at the end of next Chapter 3, following the experimental report.

4.3.1. Learning Accuracy

The error scores of BAR 1 were steady after 1000 epochs of training, thus the training of BAR 2 was conducted for 1000 epochs for both the monolingual and the bilingual training. Figure 2.8.shows the evolution of these error scores, for both the monolingual and the bilingual training.

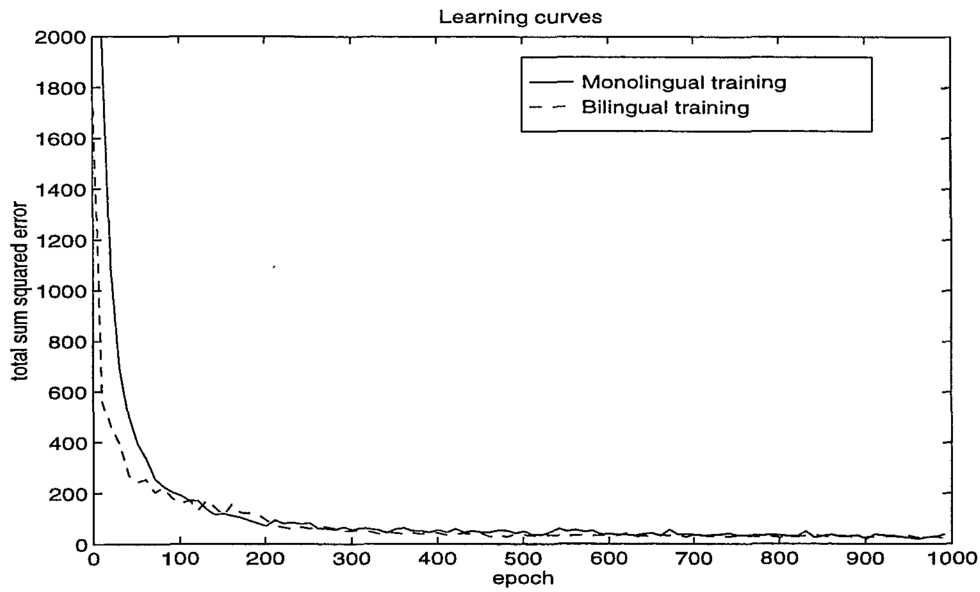

Figure 2. 8. BAR 2 Learning rates per epoch during Monolingual and Bilingual training

The evolution of the error scores in BAR 2 cannot be compared to that of BAR 1 since they are different networks with different number of units. Nevertheless, in both cases the learning of the network follows a similar evolution, showing a decline in the error scores very quickly during the first one hundred epochs of training. The most important result from BAR 2 is that the bilingual error scores are not higher than the monolingual scores unlike for BAR 1, where the bilingual training yields a higher score than monolingual training. This fact confirms the expectations that interference would not occur in the second phase of training, because of the different size of training sets.

The following analyses try to evaluate BAR 2's performance with respect to word frequency and word length, for both monolingual and bilingual training.

4.3.2. Monolingual Training

The following table presents the percentages of incorrect words obtained after the monolingual training of BAR 2. These percentages are very low compared to those obtained after the monolingual training of BAR 1 (see Table 2.1.).

Dutch	Incorrect words (%)
Only Orthography	0.17
Only Phonology	0.36
Orthography+Phonology	0.05
Total Orthography	0.22
Total Phonology	0.41
Total	0.58

Table 2.10. Percentage of incorrect Dutch words after 1000 epochs of monolingual training

After monolingual training, the performance of BAR 2 was almost perfect for all the words, indicating that using a bigger set of words and modifying the coding used were the right measures to improve the learning of the network.

The orthographic errors produced by the network are either a substitution of a character by another or a missing character, as in the examples:

```
0.26754leeftijdgenoot 1 0 0
l eeft i jdge n oot
l eeft i jdfe n oot

0.32134ogenschijnlijk 1 0 0
o g e nsc h i jnli jk
o g e nsc h i jn i jk

0.29198psychoanalyse 1 0 0
ps y ch oan a l y s e
ps y ch oan a l e s e

0.30257tegenspraak 1 0 0
t e g e nsp r aak
t e g e nsp r a k

0.29198wetenschapper 1 0 0
w e t e nsc h a pp e r
w e t e nsc h a pk e r
```

The phonological errors produced by the network also include substitutions or missing phonemes:

```
0.32134dierenarts 1 0 0
d i r @Arts
d i r @ rts

0.25320eeuwenoud 1 0 0
e w @A u t
e w @ u t

0.26754opdrachtgever 1 0 0
O bdrA xtxe v @ r
O bdrA xtke v @ r

0.26754zaterdagavond 1 0 0
z a t @ rd A G a v O nt
z a t @ rd A G a v O n
```

It is interesting to notice that the articulatory information given to the network with the phonological coding seems to have an effect also on the orthographic representations. Both in *psychoanalyse* and *wetenschapper*, the incorrect letters have a similar articulation. The same applies in for the phonological errors, as in the case of *opdrachtgever.*

Next, the analyses of learning according to word frequency and word length are presented.

Performance according to word frequency

Two sets of 200 words, with words of highest and lowest frequencies, were selected from the Dutch training set. The percentages of incorrect words in each set can be seen in Table 2.11..

Dutch	Incorrect low-freq. words (%)	Incorrect high-freq. words (%)
Only Orthography	0.73	0
Only Phonology	2.20	0.28
Orthography+Phonology	0	0
Total Orthography	0.73	0
Total Phonology	2.20	0.28
Total	2.93	0.28

Table 2.11. Percentage of incorrect Dutch low-frequency and high-frequency words after 1000 epochs of monolingual training (BAR 2)

Comparing the percentages with those of BAR 1 (Table 2.2.), it is clear that BAR 2 learned the low frequency words better than BAR 1 (2.93 and 11.50% of errors, respectively). This indicates that the set of correspondences between Dutch orthography and phonology was learnt well enough to be applied also to the words presented fewer times per epoch. Unlike in BAR 1, the highest number of errors were phonological errors (2.20%) and not orthographic errors (0.73%). Again, the bigger number of words might have been an important factor on the learning of the orthographic representation of low-frequency words. The number of phonological errors is still lower than that of BAR 1, except for the 0.28% of incorrect high-frequency words.

Nevertheless, this percentage is very low and can be neglected.

Performance according to word length

As explained in the previous Section 4.1., the coding scheme was modified for the second simulation of the model. The two templates that coded both phonology and orthography were reduced to a single one with a larger syllable structure:

CCCVV CCCVV CCCVV CCCVV CCCVV CCC

This implies that the problems encountered in BAR 1 were solved: on the one hand, short words were not represented twice; on the other, as the selected words for the training set were not longer than 13 characters, long words could be fully represented within the template. The percentages of incorrect long and short words obtained after the monolingual training of BAR 2 are shown in Table 2.12.

Dutch	Dutch shortest words (%)	Dutch longest words (%)
Only Orthography	0	3.19
Only Phonology	0	4.38
Orthography +Phonology	0	1.59
Total Orthography	0	4.78
Total Phonology	0	5.97
Total	0	9.16

Table 2.12. Percentage of incorrect Dutch short and long words after 1000 epochs of monolingual training (BAR 2)

The changes in the coding for BAR 2 are reflected in the results. The differences in learning long and short words are not so dramatic for BAR 2 as they were for BAR 1 (Table 2.3). While 37% of long words were incorrect in BAR 1, only 9.16% were incorrect in BAR 2. The difference still remaining between the learning of short and long words can be explained because some units receive activation only when the word is long, thus these units are more difficult to train. Nevertheless, the difference in percentages between BAR 2 and BAR 1 reflects that the changes made on the encoding scheme improved the performance of the network.

Overall, the analyses of the performance of BAR 2 show that the modifications on the design of the original model resulted in better learning after the monolingual phase of training. The analyses conducted after the bilingual phase of training are reported next.

4.3.3. Bilingual Training

The bilingual training was carried out for 1000 epochs, as for the monolingual training. The total percentages of incorrect Dutch and English words after this phase are summarized in the next Table 2.13.

Dutch & English	Dutch Incorrect words (%)	English Incorrect words (%)
Only Orthography	0.10	0.05
Only Phonology	0.17	1.00
Orthography+Phonology	0.02	0.05
Total Orthography	0.12	0.01
Total Phonology	0.19	1.05
Total	0.29	1.10

Table 2.13. Total percentages of incorrect words in Dutch and English after 1000 epochs of bilingual training (BAR 2)

The percentages of incorrect words in both languages are extremely low. It is remarkable that the network has also improved its performance on Dutch words (0.58% incorrect words after the monolingual training versus 0.29% incorrect words after the bilingual training, see Table 2.10) while learning the words in English.

Because of the different sizes of the training sets, it was expected that English words would be influenced by the features of Dutch words. For this reason, some incorrect words in English were examined after 100 epochs of bilingual training. Four sets of words were tested, corresponding to the high frequency, low frequency, short and long English words. The common mistakes in these sets were the substitution of phonemes, especially the substitution of the final English '*r*' (represented by *R* in the DISC character set) for the final Dutch 'r' (represented as 'r' in the DISC character set) in words as *car*, *door*, *beer*, *where*; and, less frequently, the substitution of the final group '*al*', represented by '*P*' in the DISC character set, for the single phoneme 'l', in words as *conventional* or *professional*. In both cases, the network replaced an English phoneme that does not exist in Dutch with a Dutch one that is similarly pronounced.

Some of these errors persisted after the 1000 epochs of training, as in the following examples:

0.58395international 0 1 0	0.48571literature 0 1 0	0.64392particular 0 1 0
I nt @ n { S @ nP	l It @ r @ J @ R	p @ t IkjUl @ R
I nt @ n { S @ nl	l It @ r @ J @ r	p @ t IkjUl @ r

In general, the highest percentage of incorrect words corresponds to the Phonological errors, both after 100 epochs of bilingual training and after the whole training was accomplished (1000 epochs). Nevertheless, when compared to the total percentages obtained after the bilingual training in BAR 1 (see Table 2.4), the percentages obtained after BAR 2's training are negligible.

In order to check if the network had learnt properly the English orthography to phonology correspondances, a set of new words was tested. This new set was composed by 1172 English words that had the same length as the words in the training set (3 to 14 characters) but frequencies between 25 and 44 per million. After the network was presented with these words, the total percentage of incorrect words was 19.3% (226 words), a percentage comparable to that obtained for English words in BAR 1 after the bilingual training (17.98%). Most of the incorrect words in this test set were caused by phonological mistakes (8.6%). A closer examination of the incorrect words showed again substitution of phonemes, as in the following examples:

0.44609bathroom 0 1 0 b # Tr U m b A Tr U m	0.44917corridor 0 1 0 k Q r I d $ R k Q r I d O R	0.41414cruel 0 1 0 kr U@l kr o@l
0.43634unusual 0 1 0 V nj u ZP V nj u Zl	0.42200undertake 0 1 0 V nd @ t eIk V nd @ t e@k	0.41001explore 0 1 0 I ksp l $ R I ksp l $ r

The type of errors produced by the network is generally caused by Dutch phonology interfering with English phonology. A comparison with data obtained from applied research on second language learning, which is beyond of the scope of this project, would be useful to see whether the model is able to predict the performance of second language learners.

The analyses performed on word frequency and word length are presented next.

Performance according to word frequency

In order to test the interference of the new English patterns in the already learnt Dutch ones, the same test as for BAR 1 was applied. A set of approximately 200 words of the highest and lowest frequencies among the Dutch set was tested after the first epoch of bilingual training.

Dutch	Incorrect low-freq. words (%)	Incorrect high-freq. words (%)
Only Orthography	7.69	0
Only Phonology	8.79	0.83
Orthography+Phonology	2.56	0
Total Orthography	10.25	0
Total Phonology	11.35	0.83
Total	19.05	0.83

Table 2.14. Percentage of incorrect Dutch low-frequency and high-frequency words after 1 epoch of bilingual training (BAR 2)

The interference occurring after one epoch of bilingual training is not as dramatic as was for BAR 1, and it affected only the low frequency words. The network even increased slightly its performance on high frequency words (0.83% of phonological errors, compared to 1% obtained after the monolingual training, see Table 2.11.). Apparently, the change of sizes for both training sets results in the practical disappearance of the interference effect.

The percentages of incorrect Dutch and English words of the highest and lowest frequencies at the end of the bilingual training can be seen in Table 2.15.

Dutch & English	Incorrect Dutch low-freq. words (%)	Incorrect Dutch high-freq. words (%)	Incorrect English low-freq. words (%)	Incorrect English high-freq. words (%)
Only Orthography	0	0	0	0
Only Phonology	1.10	0	6	0
Orthography+Phonology	0	0	0	0
Total Orthography	0	0	0	0
Total Phonology	1.10	0	6	0
Total	1.10	0	6	0

Table 2.15. Percentage of incorrect low-frequency and high frequency words after 1000 epochs of bilingual training (BAR 2)

The results in the table for Dutch words indicate that the effects of interference disappeared completely after the bilingual training, and that the network kept on learning Dutch words during the second phase. The results for English words are extremely good. It should be remarked that all the words in the English set had rather high frequencies (minimum of 45 per million, see Subsection 4.2.1. Set of Training Words for word selection criteria), and consequently, the English words had a high possibility of being presented in each epoch. Thus, low frequency English words were actually trained more than low frequency Dutch words.

Again, the only errors of the network concern the phonological representation. For Dutch words the percentage is very low and does not need further comment. The percentage in English low frequency words (6%) is the highest on the table. This percentage is probably due to the interference of Dutch phonology with the new patterns to learn, as was discussed at the beginning of this subsection.

Performance according to word length

Two sets of long and short words for both languages, Dutch and English, were tested after the bilingual training of BAR 2. Table 2.16. shows these percentages after the bilingual training in BAR 2.

Dutch & English	Dutch shortest words (%)	Dutch longest words (%)	English shortest words (%)	English longest words (%)
Orthographic	0	2.39	0	0.47
Phonology	0	2.39	0	6.98
Orthographic+Phonological	0	0.80	0	0.47
Total Orthography	0	3.19	0	0.22
Total Phonology	0	3.19	0	7.45
Total	0	5.58	0	7.91

Table 2.16. Percentage of incorrect Dutch and English short and long words after 1000 epochs of monolingual training

The long words were the most difficult for BAR 1 and were also those with the highest error rates after the monolingual training in BAR 2. The results in this table show that the same trend persists after the bilingual training of BAR 2. As already mentioned, this higher score on long words is probably due to the fact that long words activate units that are not activated by the rest of the words. Nevertheless the performance on long Dutch words improved after the bilingual training, if one compares the percentages with those in Table 2.12.

4.4. Discussion

The analyses performed on BAR 2 had the same sequence as BAR 1. These analyses show that the changes applied to the first version of the model resulted in a much better performance by the second version. This discussion summarizes the results of BAR 2 and compares them with those of BAR 1.

4.4.1. Error Scores

The error scores of BAR 1 and BAR 2 are not directly comparable because the architecture of the networks is different. It is important to remark that, unlike in BAR 1, the error score after the bilingual training of BAR 2 is not higher than the error score after the monolingual training. This result indicates that the training of new words did not affect the performance of the network on the old words, thus the effects of interference have been eliminated. Subsequent analyses confirm this result.

4.4.2. Monolingual Training

At the end of the monolingual training, only 0.58% of the words in the Dutch training set were not learnt properly; while this percentage for BAR 1 was 7.46%. The difference between the percentage of orthographic errors and phonological errors is not large (0.22 versus 0.41, respectively). A look at the words learnt incorrectly (see examples in Subsection 4.3.2. Monolingual

Training) shows that the errors are mainly missing phonemes or letters in long words. Substitutions of phonemes and letters were the other mistakes, and were less numerous.

BAR 2 produced better results for lower frequency words than BAR 1 (2.93% versus 11.50%, respectively) and a slightly worse result for high frequency words (0.28% in BAR 2 versus 0% in BAR 1). These percentages indicate that the bigger size of the training set enabled a better learning of the correspondences between orthography and phonology in Dutch.

With respect to word length, BAR 2 follows the same trend as BAR 1, having more problems on the learning of long words than on the short words. But again the percentage of incorrect long words is lower (9.16% incorrect long words in BAR 2 versus 37.00% incorrect long words in BAR 1). In BAR 2 the problem of the long words not being fully represented is solved, and this higher percentage of errors in the long words is not attributable to the encoding scheme used. It is possible that these words activate units both in the input and the output layers of the network that are not activated by the rest of the words. Thus, it seems that the problem is generated by the architecture of the network. Nevertheless, the advantage of short words compared to the long words in terms of accuracy is an acceptable feature in a learning model.

4.4.3. Bilingual Training

As already mentioned, the bilingual training of BAR 2 did not result in poorer performance of the network. When comparing the results of Table 2.16 with those of Table 2.4. the advantage of BAR 2 is obvious: while the percentages of incorrect long words for Dutch and English after the bilingual training were around 70%, after the bilingual training of BAR 2 they are 5.58% for Dutch and 7.91% for English. The small advantage of Dutch long words in comparison with English long words is also interpreted as an improvement of the model, which should show less accurate learning of the second language, especially since the set of English training words was considerably smaller than the set of Dutch training words.

The reason for this improvement is the difference in the size of the training sets, that resulted in a null interference in the performance with the Dutch words. The performance on Dutch words was actually better after the second training phase (0.58% incorrect words after the monolingual training versus 0.29% incorrect words after the bilingual training).

In the earlier stages of the bilingual training, the network showed a tendency to replace new English phonemes with similarly pronounced Dutch phonemes. This tendency indicates that the phonological coding used for BAR 2 reduces random substitution of phonemes, since the phonemes used for substitution are those with similar articulatory characteristics. Such a feature makes the model suitable to account for the errors of bilingual speakers.

The percentage of incorrect English words drops drastically at the end of the bilingual training, probably because the frequency of the selected English words was high. That means that the English set, although smaller, had a higher average likelihood of being presented than some Dutch words. This feature explains the low percentage of incorrect English words.

Overall, the analyses of the learning performance of BAR 2 are very satisfactory. The model shows very low scores for incorrect words for both sets of Dutch and English words, and there are no interference effects. Moreover, the observation of the results made during the bilingual phase of training on the set of English words indicate that BAR 2 is suitable for modelling second language learning, although this aspect should be further tested against applied studies.

The analysis of the internal representations of BAR 2 and the comparison of its performance with experimental data are reported in Chapter 3. Because of the different size of the training set, it seemed advisable to use experimental data that explored bilingual subjects with one predominant language, and especially the relationship of words of the first language to those of the second language.

Furthermore, another characteristic of the model is that it stresses the importance of orthographic and phonological similarity in the internal representation of the words. As was said in the Introduction to this chapter, one of the objectives of the model was to account for the cognate effect. The model, though, can account for this effect only if the cause is a similar orthography and phonology, because it does not have any representation for morphology or semantics. Accordingly, the data should explore the role of similar orthography and phonology in producing the cognate effect.

In order to collect data meeting these requirements, two experiments were carried out in this project. Both experiments had the same design and were tested on subjects from the same population, but each of them tested a different priming direction (English-Dutch and Dutch-English). Both experiments were designed to evaluate the importance of word's similar orthography and phonology on the cognate effect.

In the next Chapter 3 both experiments are introduced after a review of previous studies on the cognate effect. Subsequently, the final analyses of BAR 2 are described.

Chapter 3

Empirical Research on the Cognate Effect: Comparing the Performance of BAR 2 with Experimental Data

1. Introduction

Section 3.3. Version 1: Results and Analyses of the previous Chapter 2 presented the analysis of BAR 1. The performance of BAR 1 was compared with experimental data obtained by De Groot and Nas (1991, exp. 3), in order to evaluate the internal representation of words of two different languages that had been built up by the network.We compared it with the reaction times in cross-language priming experiments. The De Groot and Nas experiment used Dutch as the language for primes and English as the language for targets.

The results obtained from the simulation of BAR 2 indicated that its performance was far better than that of BAR 1. Accordingly, we decided to carry out a more extensive analysis of BAR 2's performance with empirical data. Thus, it was necessary to obtain data providing information about the two directions of priming (Dutch primes and English targets; English primes and Dutch targets), which was not available in the literature.

In addition, one of the features of BAR is that it uses orthographic and phonological information to build up the internal representations of words. As will be discussed in detail in Section 2 following, the experiments on the cognate effect have focused on morphology as the main factor for this effect, but the roles of orthography and phonology have not been evaluated. Orthography, phonology and their interaction are important factors in BAR 2, consequently it was necessary to evaluate them from the empirical point of view.

The experiments presented in this Chapter had a twofold objective: on the one hand, to provide new data to test BAR 2; on the other, to evaluate the role of orthography and phonology in the cognate effect.

This chapter is organized as follows: Section 2 contains an overview where former experiments on the cognate effect are presented and discussed. In Section 3 the two experiments are described. Section 4 reports the analysis of BAR 2's performance when compared with the data obtained in these experiments.

2. Empirical research on the cognate effect

The experiments reported in this section are in line with the research introduced in the Section 5.3. in Chapter 1 (The Masked Priming Paradigm in Cross-Language Experiments: The Cognate Effect). Two main findings of this body of research are the relevance of neighborhood on priming effects and the cognate effect observed across languages, both pointing at the role of form similarity as a factor organizing the lexicon.

As already introduced in Chapter 1, the study of bilingual lexical organization has focused recently at the lexical level. The model of the overlapping lexicons (Beauvillain, 1992) has been adopted to describe the relationship between the two lexicons (see figure 1.7., p. 36). In this model, the words that are morphologically similar in both languages (the cognate words) are stored in the overlapping part with a common representation.

In line with the model, the commonly accepted hypothesis is that the cognate effect is a product of morphology, or in other words, as Garcia Albea *et al.* (1985) suggested, the summation of meaning and form. The experiments designed to test this hypothesis focused on discarding other causes of the effect. For instance, Garcia-Albea *et al.* (1985) primed English targets with their corresponding Spanish cognate translations. As the subjects who performed the Lexical Decision Task were monolingual English subjects, these Spanish primes were like non-words for them. Garcia Albea and his colleagues did not observe the cognate effect in these circumstances, hence they concluded that the similar orthography was not the cause of the cognate effect. Nevertheless this result might not be conclusive: one can object that if the cognate effect is a lexical effect it will not appear using primes that are not represented in the subject's lexical memory. Other results obtained in monolingual research contradict those of Garcia Albea *et al.* (1985): Ferrand and Grainger (1992, 1994) obtained facilitatory effects using non-words as primes.

Another study supporting morphology as the cause for cognate effect was done by Alpitsis (1990). The experiments presented until now were conducted with languages using the same alphabetic system (French and English, English and Spanish), thus the visual form of cognate words is similar. For this reason, Alpitsis (1990) worked with Greek/English pairs of words, eliminating the visual form component shared by cognates. The question was whether the cognate effect would be observed across languages where visual forms are dissimilar. Her results confirmed that the cognate effect is present when orthography is different. She concluded that since orthography did not play any role in her experiment, the cause of the effect is the common morphology between cognates. An objection to this conclusion comes from a possibly neglected factor: Alpitsis remarked that her cognate words had a phonological similarity, however the possibility that phonology as the factor responsible for the cognate effect was not explored in her work. The role of phonology in the cognate effect might be important and it deserves more discussion.

First of all it is necessary to remark that the definition of cognate word is not very precise. Generally, it refers to words that share the same meaning and have a similar form. It is easy to see that the similarity of form does not only imply orthographic similarity but in many cases phonological similarity, as in the case of the Alpitsis experiment. Taking examples from Dutch and English cognates, three different types of cognates can be distinguished: Cognates whose orthographic and phonological forms are both similar (Hel/Hell); cognates that have a similar orthography but different pronunciation (Fruit/ Fruit) and cognates with different orthography that are pronounced in a similar way (Muis/Mouse; Voet/Foot). The fact that some cognates share only a similar phonology should lead to the consideration that phonology might be an important factor for the cognate effect. Turning again to monolingual research, several studies support the activation of phonological codes during visual word perception, as for example Seidenberg (1985, 1987), Perfetti & Bell (1991), Van Orden (1987), Van Orden, Pennington, & Stone (1990). Their hypothesis is that during visual word perception the phonological information attached to the word is activated, hence phonology plays an important role in the visual perception of words. Moreover, as mentioned above, Ferrand and Grainger (Ferrand & Grainger, 1992; 1994) obtained facilitation using non-words which were phonologically related to the targets. Hence, the facilitation effect of phonology during visual word perception should not be neglected in cross-language studies.

Experiments in monolingual research have been carried out as well to define the role of morphology in lexical organization. Grainger, Cole, & Segui (1991, exp. 2) used morphologically related primes and orthographically related primes and compared the magnitude of the facilitation to the target. They found that while words with similar morphology facilitate each other, an orthographically similar prime inhibits target processing. Grainger *et al.* (1991) concluded that morphology cannot be reduced to orthography since morphologically related words and orthographically related words produce opposite effects on target processing.

Two methodological points might call into question these results. Grainger *et al.* (1991, exp. 1 & 2) presented the prime in uppercase for 64 msec. The common procedure in masked priming is to present the prime in lowercase and the target in uppercase, so that the target acts simultaneously as a forward mask for the prime. It is not certain that Grainger *et al.*'s subjects were not aware of the prime[1]. Moreover, De Groot & Nas (1991, exp. 3) did not find the cognate effect when using the primes in uppercase. The replication of the same experiment reversing the case for prime and target (prime in lowercase and target in uppercase) reported the facilitation effect (exp. 4).

The second point is that the stimuli Grainger *et al.* (1991, exp. 2) used might have biassed the results. The morphologically related words actually share on average more letters with the target than the orthographically related words (see paper cited, Appendix). More letters in common could favor a bigger facilitation effect for morphological primes. Nevertheless, the inhibitory effect observed using the orthographically related words as primes cannot be explained by this feature. In subsequent experiments (Ferrand & Grainger, 1992; 1994) orthographically related non-words clearly facilitate target identification. Segui and Grainger (1990) explain that orthographic overlap has an inhibitory effect when the prime is more frequent than the target. The experiment of Grainger *et al.* (1991) has been replicated by Sanchez Casas (1995) with four-letter long stimuli. Sanchez Casas primes and targets differ in one letter, thus either morphological as orthographic primes have the same letters in common with the target. She finds a superior effect of morphology versus orthography: while morphologically similar words show a facilitation effect with respect to targets, the facilitation produced by orthographically

1. Moreover, the procedure to check if the subjects saw the prime interrupted the experiment.

related words does not differ significantly from the control words, which have no letters in common with the targets. But unlike Grainger *et al.* (1991), she does not find any inhibitory effect when using orthographic primes. The inhibitory effect of more frequent orthographic neighbors deserves further research.

Some researchers support the alternative hypothesis that morphology is only an emergent property of orthographic redundancy and orthographic/ phonological regularity (Seidenberg, 1987; Seidenberg & McClelland, 1989). Smith (1994) used a small back-propagation network to simulate the learning of words with different morphological features, and concluded that morphology helps in learning new words only if the common morphology implies regular structures, thus minimizing the role of morphology and emphasizing orthographic regularity in lexical organization. Smith points out as well individual differences with respect to the use of morphology during word processing, and interprets these differences as another indication for the relative role of morphology in word processing and learning. Overall, the research concerning morphology has not arrived at a definite conclusion.

Summarizing, in both the theoretical account for the cognate effect and the empirical research supporting it some shortcomings can be found. The innovative point of the model of overlapping lexicons is that language specificity is not longer the factor that organizes the bilingual lexicon. Nevertheless, this model has some theoretical weaknesses: the phonological information is not described; there are not clear criteria about which words are represented in the overlapping part of the lexicon; some words, as the interlexical homographs, present some problems of representation; and there is no description of factors such as frequency and neighborhood size.

Empirical research to support this model comes from a number of studies that used the masked priming paradigm. The conclusion of these studies is that the factor that causes the cognate effect is the common morphology between the cognate words. But it is possible that other factors have been neglected in these experiments, especially the similar phonology of the cognate words and the interaction of phonology and orthography during visual word recognition. On the other hand, the role of morphology in the organization of the lexicon has not been clearly defined by the empirical research. Research on the bilingual lexicon from a different approach might contribute to this debate: studying the cognate effect allows one to address questions about the relevance of morphology in the organization of the lexicon. In the next section two experiments on the cognate effect, from a different approach than those described in the previous sections, are presented.

3. Introducing a new factor in the study of the cognate effect

3.1. Introduction

Until now, in the studies on bilingual lexical organization the stimuli used have been mainly cognate and non-cognate words, thus prime and target have been semantically related. Therefore it is very difficult to draw conclusions about the cause of the cognate effect. For example, De Groot (1992) points to a completely semantic explanation, suggesting that cognate words share the same semantic representations at the conceptual level, whereas non-cognate words have a different representation at the conceptual level. From my point of view, the cause of the effect should not be located at the conceptual level until all the possible factors in the lexical level have been discarded.

From the discussion in the previous section it does not seem clear yet if the cause of the cognate effect at the lexical level is either morphology, orthography or phonology. The purpose of the present work was to further study the role of morphology in the cognate effect. A new factor called Form Similarity was introduced in the experiments reported here, in order to separate the orthographic and/or phonological similarity of words from the similarity due to the common morphology. Thus, Form Similarity is the factor that describes the similarity of form between two words of different languages that do not have any relationship in terms of meaning.

This factor was introduced in a former experiment carried out by the author (Soler, 1995). This experiment used a sample of catalan subjects that had Spanish as a second language. In that experiment the primes were in Spanish and the targets in Catalan. The stimuli set was constructed using two lists of cognate and non-cognate translations Spanish/Catalan, and it was similar to that used by Grainger *et al.* (1991) and Sanchez Casas (1995), allowing the comparison between the effect obtained with primes morphologically related with the target and the effect obtained with primes only orthographically and/ or phonologically related with the target. To each pair in the original Spanish/ Catalan lists, a new pair prime-target was added, where the prime was a word phonologically and/or orthographically similar to the target. These pairs were called false-friend pairs. Thus, for the words bearing a similar morphology, prime and target were cognate translations. For words of similar form, the prime was a word with similar orthography or phonology with respect to the target but not semantically related to it. The task that subjects had to perform was a cross-language lexical decision task with masked priming. The effects of morphology and similarity of form were compared in terms of facilitation obtained for each set of prime stimuli.

The results obtained in this experiment showed a clear facilitation in both lists. The reaction times obtained for cognate primes and for false friends did not differ significantly. In the non-cognate list, the reaction times on non-cognate words were significantly longer that for the false-friend pairs. These results seem to indicate that the factor Form Similarity has a facilitation effect comparable to the cognate effect.

The experiments reported next were carried out at the Institute for Perception Research (IPO) in Eindhoven (The Netherlands). The subjects that participated in the experiments had Dutch as a first language and were proficient in English. In Experiment 1, the primes were in English and the targets in Dutch; in Experiment 2, the primes were in Dutch and the targets in English. The design of these experiments included the introduction of two conditions: repetition priming and control priming. The next subsection includes a detailed description of this design.

A remark has to be made with respect to the effects reported by Segui and Grainger (1990) about frequency and neighborhood. They observed that when the prime was a neighbor of higher frequency than the target, the recognition of the target was inhibited. Unfortunately, word frequency cannot be controlled in cross-language experiments since there is a wide range of variation in the exposure to the second language in bilingual subjects. On the other hand, one of the locus of word frequency is the actual occurrence of words in the use of language (Monsell, 1991). In these experiments, the sample was composed of Dutch speakers that live in the Netherlands and need a high knowledge of English in their professional lives. Thus, it can be assumed that, in spite of their proficiency in English, Dutch words are more frequent than English words for these subjects. Both experiments are described next.

3.2. The experiments

As introduced in the previous section, the aim of the experiments was to compare the facilitation effect produced by the factors Form Similarity and Morphology. Two sets of stimuli were constructed from an original list of cognate and non-cognate pairs of words. The factor Form Similarity was introduced by generating a new pair for each of the targets, where the prime was a word orthographically and/or phonologically related to the target. These new pairs were called false-friend pairs.

One of the features of the cognate effect described in the literature is that the facilitation obtained with cognate words is comparable in terms of magnitude with the facilitation obtained with the repetition of the word (Garcia Albea *et al.*, 1985). In order to evaluate if facilitation effects of both Form Similarity and Morphology, a repetition condition was introduced into the experimental design.

To have a reference for the facilitation obtained in each of the above mentioned conditions, a control condition was added, with primes not related with the targets in terms of meaning or form. Thus, all the possible relations between prime and target, in terms of meaning and form, were considered (see Table 3.1).

Type of pair	meaning related	form related
repetition	yes	yes
cognate	yes	yes
non-cognate	yes	no
false-friend	no	yes
control	no	no

Table 3.1. Relation between primes and target in Experiment 1

In the repetition priming, the prime is the target word[1]. In the cognate priming, the prime is a cognate translation of the target. In the non-cognate priming, the prime is a non-cognate translation of the target. In the false-friend priming, the prime is a non-translation word with similar orthographic and/or phonological form than the target. Finally, in the control priming the prime is a non-translation word with different orthographic and phonological form than the target.

The two priming directions were tested: L2 (English) to L1 (Dutch) in Experiment 1 and L1 (Dutch) to L2 (English) in Experiment 2. Both experiments are reported next.

1. Thus, this is the only within-language condition, where prime and target are from the same language.

3.2.1. Experiment 1

Method

Subjects

A total of 40 subjects participated in this experiment. Some of them were researchers at the Institute for Perception Research (IPO), and others were students working temporarily at IPO on their graduation projects. All of them spoke Dutch as a first language and their ages ranged from 21 to 55 yr.

Materials

Two sets of stimuli were designed following the procedure described. The original list of cognates and non-cognates was extracted from the set of stimuli used by De Groot and Nas (1991, Appendix). Table 3.2 shows examples of the four pairs generated for each target (the complete set of stimuli can be found in Appendix 4).

	primes				**target**
Type of translation word	**repetition**	**cognate**	**false-friend**	**control**	
cognate	appel	apple	appeal	sauce	APPEL
non-cognate	wortel	carrot	worth	bike	WORTEL

Table 3.2. Stimuli words (Primes and target) used for Experiment 1

To elaborate the non-words list, 20 Catalan words were chosen as targets. These words were not known by the subjects. The primes were English words, either similar or dissimilar to the Catalan words. Thus, the non-word list had two conditions: false friend non-words and control non-words. Table 3.3 shows a sample of these lists.

non-word lists	prime	target
false-friend nonwords	rabbit	RIBOT
control non-words	rose	BLEDA

Table 3.3. Non-word stimuli used in Experiment 1

For the stimulus presentation, different lists were constructed in order that each target should appear only once in each list, keeping the conditions balanced between lists and varying the order of presentation for each subject. Thus, each subject had a different list of stimuli.

Procedure

Subjects were tested individually in an office at IPO. The experiment was programmed on PsyScope (Cohen & MacWhinney, 1994) and ran on a Power PC Macintosh. The subjects started the experiment by reading the instructions on the screen, which were in the language of the targets. After hitting a key to start, the training phase followed the instructions, and after a short break (one minute) the experimental phase started. The training phase consisted of eight stimuli, and the experimental phase of 40 stimuli. The sequence of events was similar to that in the experiments of De Groot and Nas (1991): it started with a fixation point (*) that was presented for 500 msec, followed by the mask (six characters #) displayed for another 500 msec. The prime was presented in lowercase for 48 msec, and was followed by the target, displayed in uppercase for 500 msec (or until the subject answered). Subjects pressed a key at the right for the 'word' answer and a key at the left for the 'non-word' answer. The subjects were informed if their answer was correct (GOED), wrong (FOUT) or slow (LANGZAAM). The feedback 'Langzaam' was given when the subject did not answer within an interval of 1200 msec[1]. The feedback remained on the screen for one second, after which a new trial started.

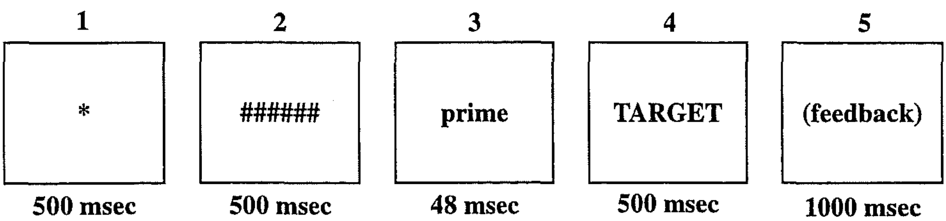

Figure 3. 1. Sequence of events for Experiment 1

Results

A preliminary analysis of the data was performed before the statistical analysis. The subjects with fewer than 30 correct answers (25% of the answers wrong) were removed from the sample (two subjects). Reaction times more than two standard deviation above or below the mean of the correct word answers were trimmed to the appropriate cut-off value to moderate the influence of outliers. This procedure left 735 valid observations for the Word condition on which the statistical analysis was performed. The same procedure was followed for the Non Word condition, which yielded 680 valid observations.

1. The speeded responses also help to avoid semantic interferences (Keatley and De Gelder, 1992)

The mean results for each category of prime for Words, in both conditions, cognate and non-cognate, are shown in table 3.4.

cognate list	R.T.	Sd	non-cognate list	R.T.	Sd
repetition	472.12	80.34	**repetition**	495.30	92.49
cognate	508.49	81.72	**non-cognate**	507.93	82.74
false-friend	550.71	86.59	**false-friend**	524.01	101.28
control	525.80	80.34	**control**	516.47	79.33

Table 3.4. Mean Reaction Times and Standard Deviation per condition in Experiment 1

These results are graphically represented in Figure 3.2.

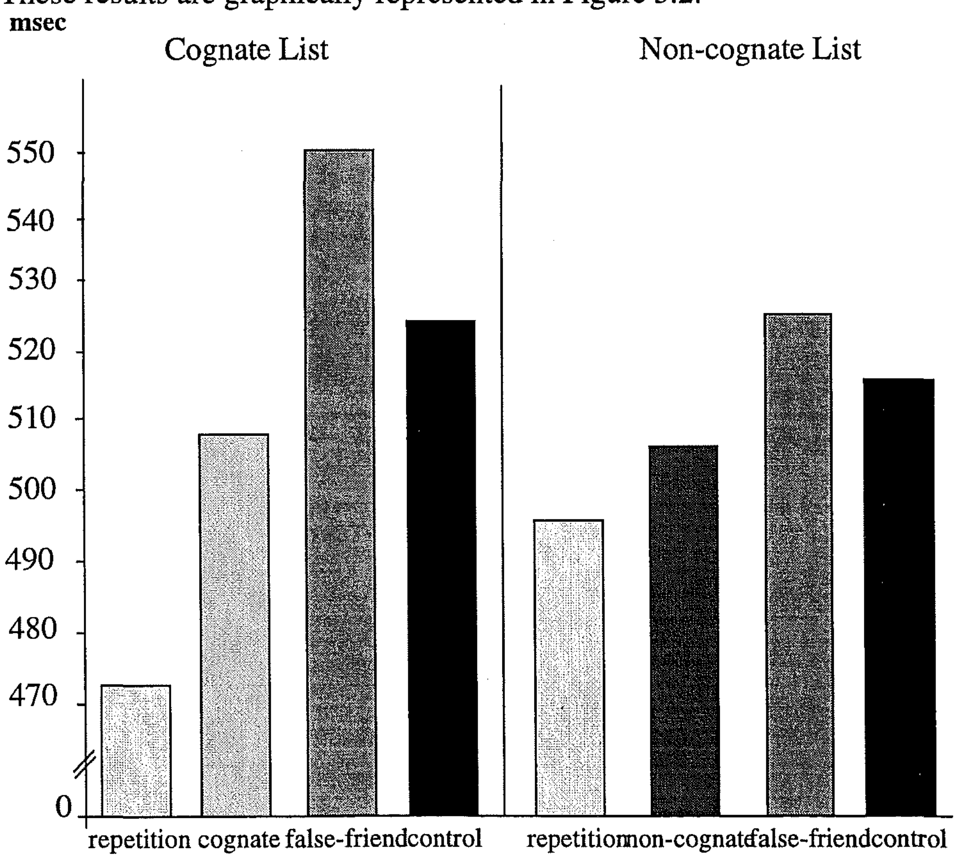

Figure 3. 2. Graphic representation of results for words in Experiment 1 (rt)

The differences between the mean obtained for the primes in the control condition and the means of the rest of conditions give the index of facilitation for each condition, shown in Table 3.5.

Cognate list	Facilitation	Non-cognate list	Facilitation
repetition	-53.68	repetition	-21.17
cognate	-17.80	non-cognate	-8.54
false-friend	+24.69	false-friend	+7.54

Table 3.5. Facilitation observed per condition (condition rt - control rt)

In order to find out if these facilitation effects were significant, t-tests for mean comparison was performed. The results obtained for the cognate list are shown in table 3.6.

Cognate List

	repetition	cognate	false-friend	control
repetition	.	-2.86971 0.0042	-635892 0.0001	-4.49168 0.0001
cognate	2.86971 0.0042	.	-347646 0.0005	-1.62264 0.1051
false-friend	6.358921 0.0001	3.47456 0.0005	.	1.876756 0.0610
control	4.491675 0.0001	-1.87676 0.0610	1.622637 0.1051	.

Table 3.6. T-test values and level of significance for the cognate list (df=690)

As can be seen in Table 3.6, the facilitation produced by repetition primes is