Olga SOLER VILAGELIU

BILINGUAL LEXICAL ACCESS: A CONNECTIONIST MODEL

Tesi Doctoral dirigida
pel Dr. Santiago Estaún i Ferrer

Departament de Psicologia de l'Educació
Facultat de Psicologia
Universitat Autònoma de Barcelona
Any 1996

per la mare, el pare, la míriam i en david

i també per en pau, que ha guanyat la carrera
naixent abans que aquesta tesi

Agraiments

En primer lloc, em cal agrair al Dr. Santiago Estaún el seu recolzament constant en aquests darrers anys. Molt especialment vull donar les gràcies a la Dra. Núria Sebastiàn per introduïr-me a la Psicolingüística (i per tantes altres coses). Gràcies també al Dr. Antoni Castelló pel seu temps i interès, i a la Dra. Rosa María Sanchez-Casas per aconsellar-me respecte als experiments d'aquesta tesi.

També vull agrair al departament de Psicologia de l'Educació que m'hagi permès les estades als Paisos Baixos, on aquest treball ha estat realitzat. I gracies també als meus alumnes, que han hagut de suportar els inconvenients de les meves absències.

No puc oblidar aquí d'altres persones que han estat al meu costat: la meva família (i dins d'ella el Dr. Joaquim Llisterri), i aquells qui no són família però és com si ho fossin. D'entre aquests he d'esmentar especialment la Montserrat Rovira i la Sílvia Zaragoza. A tots, el meu agraïment més profund.

I finalment he de donar les gràcies a dos que no són persones: en Higgins i la Liza, per fer-me la vida molt més agradable.

Acknowledgements

I am indebted to Prof. Don Bouwhuis for many reasons, but especially for taking the risk and giving me the chance to work at IPO. Thanks for your trust, Don. I wish to thank Dr. Rudy van Hoe for his invaluable advice, his moral support, his help on statistics and for the many times he said 'I think it's possible'. I am also indebted to Robert Hofsink, who implemented the second version of the model, for his brains and his patience.

This thesis would not be readable without the help of Dr. Cynthia Grover. Thanks for this and so many other things, Cyn.

To these persons who offered me their friendhip and encouragement (and some hard times too) during my stays in The Netherlands, thank you all.

Finally I have to thank the staff at IPO, especially Nel Evers, Unni van Amelsfort and Han Bolkestein, for their help on solving all these little nuisances.

Contents

Introduction

This project presents a connectionist model of lexical access representations in the bilingual lexicon. The purpose of the project was to simulate the learning of a new vocabulary during second language acquisition. Thus, the literature overview presented in Chapter 1 is centered on the bilingual lexical organization models developed within the theoretical framework of cognitive psychology.

The review of the current models in Chapter 1 revealed that these models were not suitable as a basis for a model aiming to simulate the process of learning of new words. All the models reviewed use a local representation for words, that is, each lexical entry is represented by a single node of information. Such a representational structure implies that a bilingual model developed according to these features would need to grow for each new word learned. It seemed that for the purposes of the present project, an alternative way of representing lexical entries should be used.

Moreover, current models of bilingual lexical organization seem to lack accuracy in their description of the information stored within the lexicon. Two main assumptions are common to all models: the lexical information is stored in two levels (the lexical and the conceptual levels); and lexicons are

language-specific. Information about lexical entries is stored at the lexical level, which is connected to the conceptual level, where the meaning of the words is stored. While the conceptual level is considered to be common to all languages, the lexical level is, as mentioned above, language-specific.

The research reviewed explored the relationship between these language-specific lexicons, and the consequence is that the internal organization of the bilingual lexicons has not been seen as an important issue. Thus, concerning the lexical level, there is a lack of description about important features which are crucial in monolingual models of lexical access and deserve much attention in language processing research, such as orthography, phonology and word frequency.

In order to design a bilingual lexical model that is able to learn, it is necessary to overcome these weaknesses by taking a different theoretical approach. The model should meet two requirements. The first one is that the model should learn new words without increasing its size. This requirement has two different aspects: From a psychological point of view it does not seem plausible that a person able to speak six languages has language-specific local representations of each word in these six languages. This issue is further discussed throughout the pages of the present work. On the other side, it is a practical constraint for the implementation of the model: If the size of the model has to increase during the learning process, the number of words learnt would be restricted by external characteristics such as the amount of memory available in the system used.

The second requirement concerns the information represented in the model. It has already been mentioned that the research reviewed focused on the relationship between the two language-specific lexicons. The trend in this

research has noticeably been to minimize language-specificity as the most important factor in the organization of the bilingual lexicon, and recent research considers morphology to be the main factor structuring the bilingual lexicon. In order to capture this role for morphology, it was necessary to provide the model with orthographic and phonological information. It was challenging to design a model that could treat language-specificity as a consequence of the different orthography/phonology constraints in each language. The scope of the model was reduced to the lexical level, in order to explore the possibilities of orthographic and phonological information as the factors organizing the bilingual lexicon.

The connectionist model Bilingual Lexical Representations Model (BAR), based on Seidenberg and McClelland's (1989) model for word recognition and naming, is presented as an alternative to the models reviewed. Its major features are the distributed representation of lexical entries, and the use of orthographic and phonological information in these representations as the main index for language-specificity. The aim of BAR is to offer an alternative way to describe the lexical level within the bilingual lexicon and to simulate the learning of lexical entries in two different languages, taking into consideration factors such as word frequency and word length. BAR was developed in two phases, BAR 1 and BAR 2, which are fully described in Chapter 2.

The performance of both networks was evaluated in terms of learning accuracy, but also by comparing their performance with empirical data from cross-language experiments. The performance of BAR 1 was confronted with data from cross-language priming experiments reported in the literature (De Groot and Nas, 1991, exp. 4).

The analysis of BAR 1 indicated the changes needed for its performance to become more accurate. A few aspects of the coding of the information supplied to the network were changed, and the size of the training sets of words were also adjusted. This second simulation was called BAR 2.

The internal structure of BAR 2 was tested against data coming from two experiments carried out within this project. In keeping with BAR characteristics, it was necessary to have empirical sources of information about the influence of similar orthography and/or phonology on bilingual subjects performance; besides, it was convenient to use data that explored the relationship between the words of the bilinguals two languages from both directions; in the present case, Dutch words with respect to English words and English words with respect to Dutch words. These requirements led to the design of the two experiments that are reported in Chapter 3.

The two experiments in Chapter 3 consist of a cross-language decision task using the masked priming paradigm in order to explore the Cognate Effect. The cognate effect has been found in recent research on bilingual lexical organization, and it is described as a facilitation effect between the words of the bilingual's two languages: when the prime is a translation morphologically similar to the target, the processing of the latter is faster. This effect has been explained by the interaction of form and meaning of the two words in question. The experiments reported in Chapter 3 introduced a new variable, named Form Similarity, with the purpose of testing the role of orthography and phonology and their interaction as factors in the cognate effect. The results obtained from these experiments were used to test the second version of the model. A discussion about the suitability of the model closes Chapter 3.

The Conclusions of this work point out the different possibilities of the BAR model as a tool for exploring both monolingual and bilingual lexical processing.

Chapter 1

Bilingual Lexical Organization: An Overview

1. Introduction

The general purpose of this chapter is to present a review of the literature in order to examine and evaluate the current models of Bilingual Lexical Organization. Research on bilingualism presents the opportunity to address fundamental issues in language processing, such as memory organization, lexical organization and lexical access. In this section the general outline of the document is presented.

The initial studies on bilingualism took place in the middle of this century. Bilingual subjects were classified according to their proficiency and level of

learning in their different languages. In these classifications bilingual subjects were considered to be exceptional, as opposed to 'normal' monolingual subjects (see Grosjean, 1992, for an extended overview).

Examples of classifications are those of Weinreich (1953/1968), and Ervin & Osgood (1954). Weinreich's (1953/1968) description of the language processing system includes two different types of memory: a conceptual memory where meanings are stored; and a lexicon where the lexical entries and the information attached to them are stored. This distinction is crucial since it gave rise to specific research on lexical organization independently of semantic information.

Weinreich proposed three different types of word knowledge organization in bilinguals (see Figure 1.1)

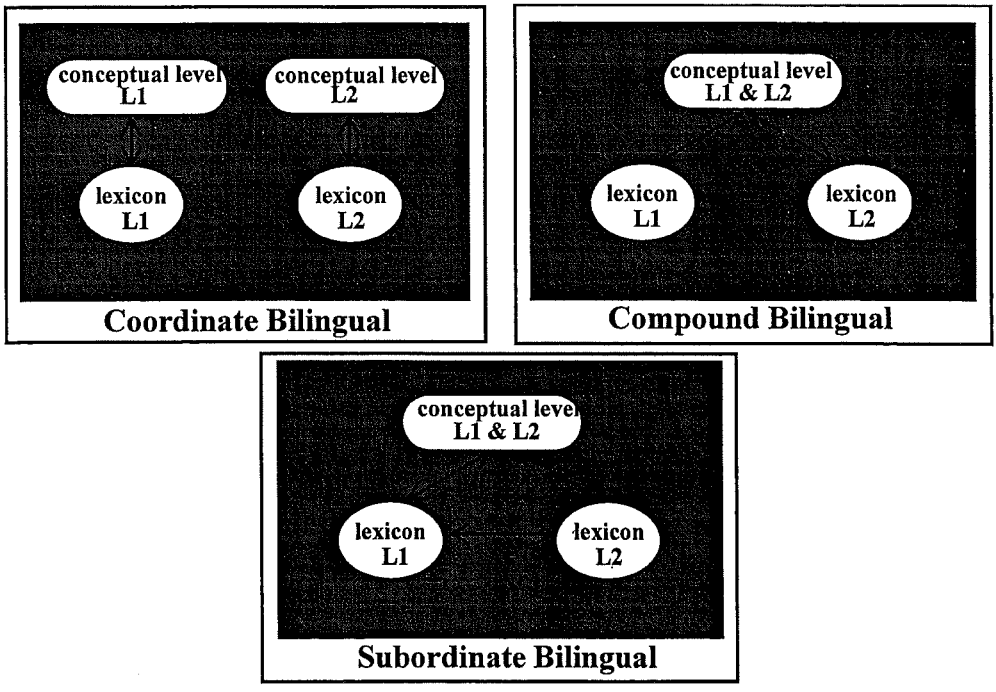

Figure 1. 1. Weinreich's classification of bilingual subjects

For Coordinate bilinguals words and meanings of the two languages are separately structured. For Compound bilinguals one conceptual representation is linked to two lexical systems. Finally, for Subordinate bilinguals the second language is related to the conceptual level only by means of the first language. The classification of Ervin & Osgood (1954) is quite similar to Weinreich's. Due to their behaviorist background, they described the classifications in terms of how languages are acquired, and not from the point of view of internal memory organization. As will be seen in the following sections, these classifications are very similar to the current models of lexical organization.

During the sixties, research in the bilingual field focused on linguistic processing and internal representation of the languages, which can be called cognitive research on bilingualism. According to Keatley (1992) three main research lines developed as a consequence of this new trend. One line of research concentrated on bilingual memory whereas another concentrated on second language learning. A third line developed in the mid-eighties and reoriented the research of the first group. These three lines will be briefly described.

The first group of studies was devoted to the basic processes of human cognition, especially to mnemonic processes related to the knowledge of more than one language. The main subject of the research of this group is memory for meaning in bilinguals, more specifically whether there is a common linguistic system for both languages. These studies are closely related to the above cited classification of Weinreich.

Our literature review takes these studies as a starting point. Kolers (1963) proposed two alternative hypotheses, namely the Independence vs.

Interdependence of languages, which will be discussed in next section. He supported the Independence Hypothesis, claiming that bilingual subjects possess two completely independent language systems, one for each language (as for the Coordinate bilingual in Weinreich's). Other models were developed based on Kolers' work, as MacNamara's Switch-model (MacNamara, 1967a; 1967b; MacNamara & Kushnir, 1971); and the Dual Coding System of Paivio & Desrochers (1980). These models are presented Section 2, **Bilingual Memory Storage**.

The second area of research is more applied, and concerns the studies devoted to Second Language Learning, with the aim of improving teaching techniques and the learner's performance. The main topic is language interference. This line of research focuses on a functional model of language to describe how an inactive language (the one not actually in use) influences the active language (the one the subject is using). These studies are not considered in this review, because their research motivation is mainly applied research on acquiring a second language, and not fundamental research on cognitive processing.

The third line of research appeared in the mid-eighties and it is related to the first one. Hence, the language processing system is described as a two-level system, with a lexical level and a conceptual level. At the lexical level each word-entry is represented by a unique node of information. The main difference with the former research is the idea that languages must share a common substrate at the semantic level. It is assumed that there is a conceptual representation of the meaning of a word which is not language-specific. From this point of view, a single linguistic system copes with bilingual language processing, but the structures of this system are in certain cases language-specific.

The issues which arose were the level at which the language representations are independent and the level at which they are common (Grainger, 1987). The generally accepted hypothesis is a language system consisting of a common semantic level and two independent lexicons connected to the semantic level, but there are two different hypotheses concerning the relationship between these two lexicons: the Word Association Hypothesis (Kirsner, K., Smith, M.C.; Lockart, R.S.: King, M.L. & Jain, M.,1984) and the Concept Mediation Hypothesis (Potter, M.C.; So, K.F.; von Eckart, B.; & Feldman, L.B., 1984). The experimental paradigm most often used to test these hypotheses is the priming paradigm on lexical decision, naming and translation tasks. The two hypotheses and subsequent experiments are reviewed in Section 3, **Models proposing two lexical systems**. This section reviews as well the research done within this framework on learning and language-specific factors (orthography and speech-segmentation) and includes a discussion about the models proposed.

In the present decade two interesting trends have emerged on bilingual lexical organization. First there is the application of the connectionist framework for the description of a bilingual lexical system. For instance, a neural network model for bilingual organization has been proposed by Grainger and Dijkstra (1992), namely the Bilingual Activation Model, based on McClelland & Rumelhart's (1981) Interactive Activation model. It is described in Section 4, **Parallel Access to Bilingual Lexicon.**

The second trend is experimental. The masked priming paradigm turns out to be a very interesting tool to study internal organization in bilingual subjects. In contrast to the priming paradigm, where subjects are aware of seeing the prime, the masked priming paradigm is able to leave out of account subject's

higher-level strategies such as expectation, meaning integration, and episodic traces of memory. For this reason masked priming makes it possible to study the links between words without interference from the semantic level.

Using this paradigm, some authors have found evidence to support a different model of bilingual lexical organization based on morphology. This model also claims the existence of two separate lexicons, one for each language. The results obtained using the masked priming paradigm seem to indicate that a separate kind of organization could apply to cognate words. Cognate words are words that have a similar form and meaning across languages. It is suggested that the two language lexicons have an overlapping part where these cognate words are stored. Both the new model, proposed by Beauvillain (1992), and the related masked form-priming experiments are described and discussed in Section 5, **The Overlapping Bilingual Lexicon.**

Section 6 contains a general discussion and suggests a new direction for research on bilingual lexical organization and access.

2. Bilingual Memory Storage

2.1. The Independence/Interdependence Hypothesis

The research of Kolers (1963) will be taken as a starting point to review the research on bilingual memory. Kolers was the first one to formulate the Independence/Interdependence Hypothesis. The Independence Hypothesis claims that languages are stored separately in the bilingual's memory, thus the words of the two different languages are not connected. Independent storage also means that it is not possible to find between-language effects such as priming in a Lexical Decision Task (henceforth, LDT) or repetition effect, as they occur within languages.

In contrast, the Interdependence Hypothesis claims that languages share a common memory store. Thus, the representations at the semantic-conceptual level should be the same for both languages, implying that semantic priming can occur between languages.

Kolers (1963) found evidence that subjects could associate a word and its translation with different concepts, depending on the language they were using. He interpreted these results as evidence for a separate semantic representation of equivalent words in different languages.

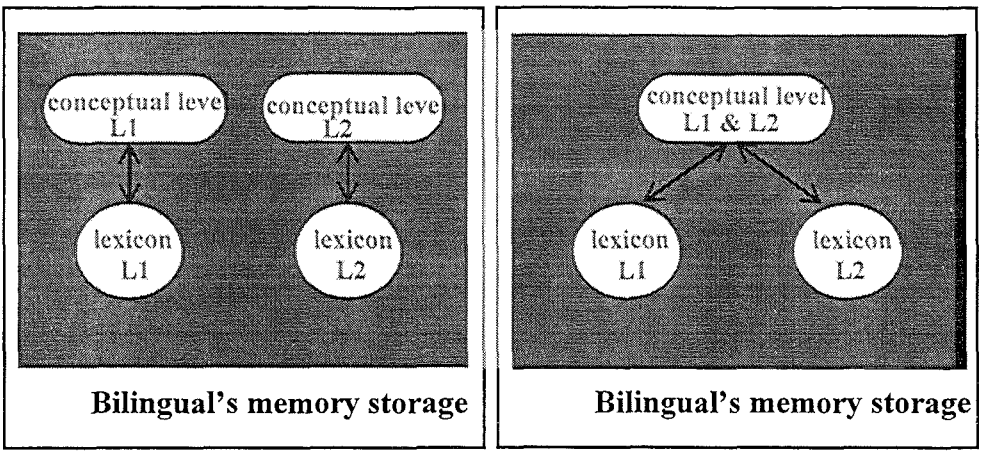

Figure 1. 1. Model of Kolers (1963). The Independent Hypothesis suggests two completely independent linguistic systems for each language in the bilingual's memory. The Interdependence Hypothesis suggests the conceptual information is shared by the two languages[1]

Kolers concluded that there is no common linguistic representation at any level for the different languages of a bilingual speaker; rather memories are stored in the language in which the subject experienced a concrete situation.

In a free-recall experiment with monolingual, bilingual and trilingual lists (English, French and Spanish), Tulving and Colotla (1970) found empirical support for Kolers' hypothesis. Lists of words were presented to each subject and after each list presentation, subjects had some time for recalling as many words as they could from the lists they had seen. The results showed that recall for bilingual and trilingual lists favors first language words. The amount

1.cf. Figure 1.1: the two models match Weinreich's classification on Coordinated and Compound Bilinguals.

of words recalled by subjects from the bilingual and trilingual lists was bigger than the amount recalled from the monolingual list, because more words from subject's first language were recalled. The authors interpreted this result as evidence for separate storage: if all the words of different languages were stored together, it should not be easier to recall words of one language rather than another.

The question that arose consequently was how these two linguistic systems are activated. MacNamara's research was devoted to that topic, and it will be reviewed in next section.

2.2. The Switch Model

MacNamara (1967a/b; MacNamara & Kushnir, 1971) explored how the two systems described by Kolers interacted and maintained their independence. In his experiments subjects had to perform different tasks in monolingual or bilingual contexts. His results confirmed those of Kolers: he found that in a bilingual context the subjects needed more time to solve the task. To explain this phenomenon, McNamara developed the switch model.

According to the switch model, the two linguistic systems of a bilingual speaker are organized in a way that only one of them is accessible at a time. The language is selected according to a decision mechanism that MacNamara called a switch. There are two Switches for each linguistic system: the output switch, controlled by the speaker; and the input switch, controlled by the stimulus. The switch has two possible states: *on* and *off*, equivalent to the active or passive state of the language. When one language switch is *on*, the other language is *off*, both languages are never activated at the same time.

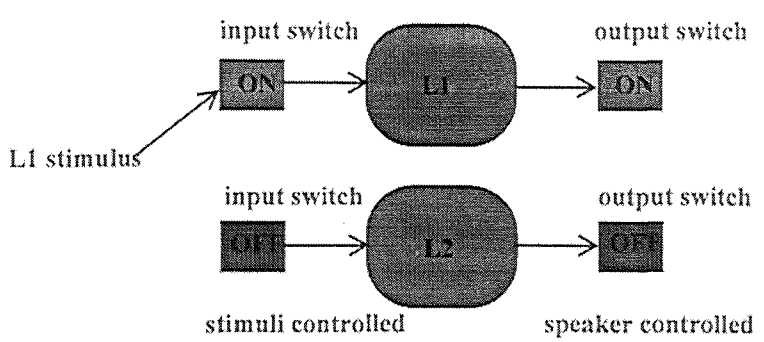

Figure 1. 2. MacNamara's switch model. In the figure, the L1 stimulus turns ON the L1 switch, so L1 is the active language and L2 the passive language.

This model implies that the input mechanism must generate a preliminary categorization of the received signal and identify it as belonging to one or the other language in order to process the proper language. This categorization takes some time: in this way MacNamara explained the reason why his subjects had longer reaction times when the task required the comprehension of more than one language. Subsequent research investigated the processing level at which the stimulus categorization is done.

MacNamara's interpretation of the empirical results has been questioned. Doctor and her collaborators (Doctor, E. A.; Ahmed, R.; Ainslee, V.; Cronje, T.; Klein, D.; Knight, S., 1987) pointed out the time bilingual subjects spend on switching from one to the other language can be justified by the fact that languages differ orthographically, phonologically, syntactically, and semantically. The information about the languages has to be separated to avoid confusion but this does not imply that the languages have to be activated in an independent way.

2.3. The Dual Coding System

The model Paivio & Desrochers (1980) developed for bilingualism takes a position between the Independence and Interdependence hypotheses. This model is based on the Dual Coding System model previously developed by Paivio (1971). The Dual Coding System gives a global framework for verbal and non-verbal processing. It is composed of two associative networks of logogen-type units (for a description of the Logogen Model, see Morton, 1969)).

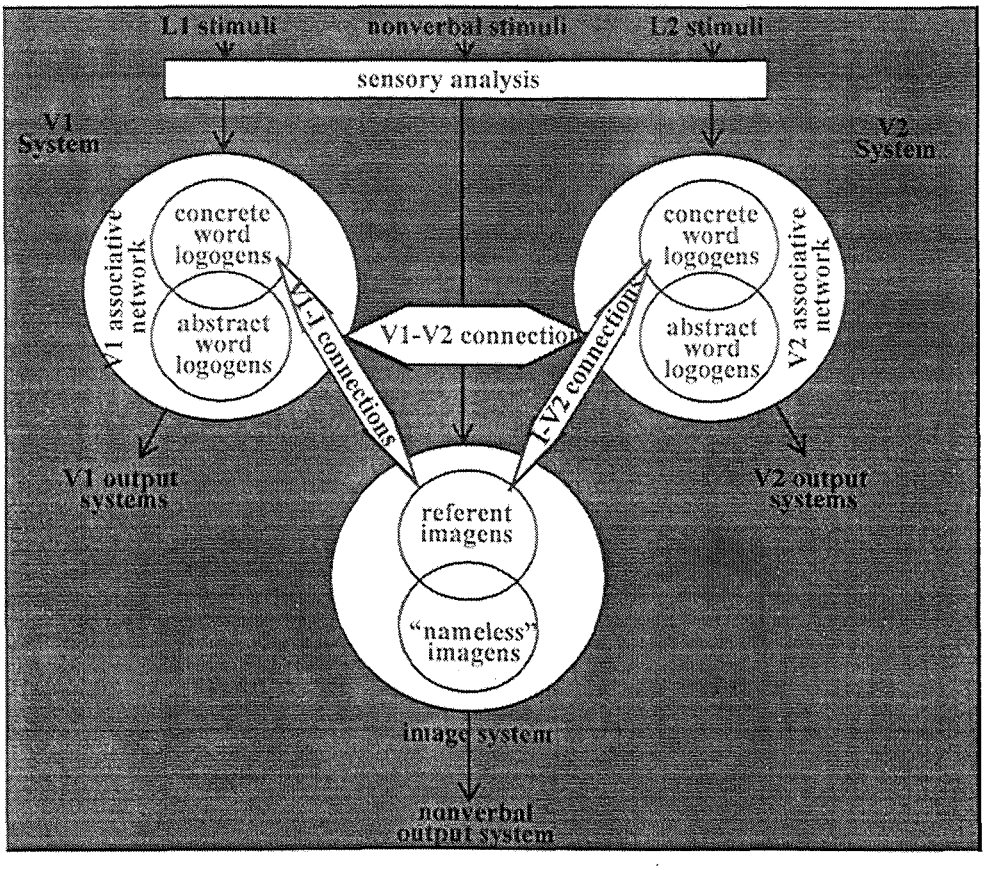

Figure 1. 3. Paivio & Desrochers (1980) Dual Coding System Model

The linguistic associative network is divided in two subsystems composed of concrete-word logogens and abstract-word logogens. This network, in turn, is connected to the image-system network, representing the knowledge of the world, and it is composed of two subsystems: referent *imagens* and nameless *imagens*. The *imagens* are units similar to *logogens*, but specific for image systems. Linguistic and image systems have different input-output mechanisms. In order to adapt this model to bilingual processing, the authors incorporated a new language system (i.e. a new associative network) corresponding to the second language. This network is connected to the original linguistic system and to the image-system as well, and input-output mechanisms are also independent.

Paivio & Desrochers' model is probably too general and it was not tested in subsequent research. The model has some weak points with respect to linguistic processing. Basically it neither explains what information is stored at semantic level nor some particular linguistic phenomena such as phonological representation. In general, the model does not solve the theoretical conflict between the independence and interdependence hypotheses.

Nevertheless, the dual coding system has proved to be useful in explaining phenomena described in the recent literature on bilingual lexical structure research. De Groot (1992) and Keatley, Spinks & De Gelder (1994) used this model because it can account for asymmetrical cross-language priming, showing an unbalanced relationship between the two languages of bilingual speakers. De Groot (1993) analyzed the abstract vs. concrete word effect in bilingual processing tasks and suggested that the internal representation for abstract and concrete words is different in the bilingual lexicon: while

translations of concrete words have very similar meanings, translations of abstract words might have language-independent meanings. De Groot proposes that semantic representation might be distributed. Thus, translations of concrete words would share the same semantic nodes and translations of abstract words would not. This issue is further discussed on Chapter 3.

2.4. General considerations

A common feature of the models developed between the 60s and the 80s is their holistic point of view. The authors focused on language as a special memory system with some specific mechanisms, such as the switch suggested by McNamara. The holistic point of view is especially remarkable in the model of Paivio & Desrochers (1980) just described.

But as new effects were empirically found (especially concerning lexical access), researchers were forced to narrow their attention and devote themselves to more specific topics. Thus, research on bilingual lexical access appears to have been radically separate from research on monolingual lexical access in the years that followed. As will be seen, most of the research concerning visual word perception, for example the discussion on Dual/ Single Route of Lexical Access, was completely ignored by researchers in bilingual lexical organization. The next section describes the models proposed in the 80s that are still supported nowadays. These models are essentially concerned with the relationship between the two lexical systems of the bilingual speaker.

3. Models proposing two lexical systems

The two most important theories in the 80s on bilingual lexical system structure are the ones proposed by Kirsner, Smith, Lockart, King, & Jain (1984); and Potter, So, von Eckart, & Feldman (1984). Both papers reviewed previous literature and intended to summarize the results of different models.

The models have in common that two lexical systems are assumed, one for each language, and a common representational system where concepts are stored[1]. This idea had already been introduced by Caramazza & Brones (1979) and it is intrinsic to the Interdependence Hypothesis. The main difference between the two models is in the relation between the two lexical systems. Both models are described next.

1.It is important to consider, as De Groot & Barry (1992) point out, that in the literature on bilingual lexicon the terms 'meaning', 'meaning representations', or 'concept representations' are used interchangeably.

3.1. The Hypotheses

3.1.1. Concept-Mediation Hypothesis

Potter *et al.* (1984) presented two hypotheses known as the Word Association Hypothesis and the Concept Mediation Hypothesis (see figure 1.5). According to the Word Association Hypothesis, lexicons are connected directly by means of word connections. According to the Concept Mediation Hypothesis model, the lexicons are related only via their connections with the conceptual information, which is common for both languages.

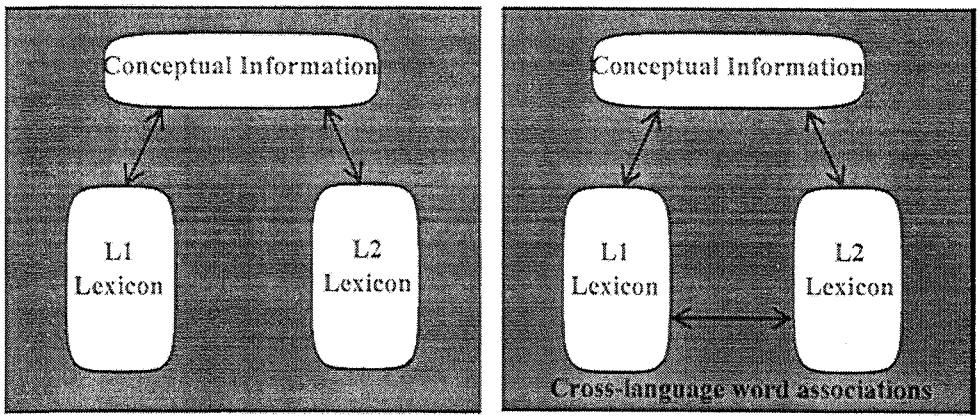

Figure 1. 5. Left: Concept Mediation Model (Potter *et al.*, 1984); Right: Word Association Model (Kirsner *et al.*, 1984). The only difference between the models is the direct association of word entries at the lexical level.

In their research, Potter *et al.* (1984) used three different experimental tasks (reading words aloud, translation, and picture naming) for two subject groups with different levels of proficiency in their second language. They generated predictions that would support one or the other hypothesis, as summarized in the following table. In the table, the different steps needed for each task according to both hypotheses are described. The more steps needed, the longer the Response Time (R.T.). Remark that for the Concept Mediation Hypotheses the steps are the same, so the R.T. should be identical.

Word Association Model		Concept Mediation Model	
Picture-naming	Translating	Picture-naming	Translating
(1) Recognize Image	(1) Recognize L1 word	(1) Recognize Image	(1) Recognize L1 word
(2) Retrieve Concept		(2) Retrieve Concept	(2) Retrieve Concept
(3) Retrieve L1 word			
(4) Retrieve L2 word	(2) Retrieve L2 word	(3) Retrieve L2 word	(3) Retrieve L2 word
(5) Say L2 word	(3) Say L2 word	(4) Say L2 word	(4) Say L2 word

Table 1.1. Predictions according to Concept Mediation and Word Association Hypotheses for Picture-Naming and Translation tasks (Potter *et al.*, 1984).

For both groups of subjects the results showed there was no time difference for the two tasks, suggesting that there is no direct word association between the two lexicons. Potter *et al.* (1984) concluded that words of different languages always relate through the conceptual level of representation.

3.1.2. Word Association Hypothesis

The experimental approach taken by Kirsner and his collaborators (1984) to support the word association hypothesis is different from Potter's. They used the priming paradigm in a lexical decision task to try to find repetition priming effects. The facilitation in the formal (or repetition) priming is an effect useful to check if two different forms of a stimulus have the same internal representation. For example, a word in uppercase is presented as a prime for the same word in lowercase. Kirsner *et al.* used this technique in order to find out if the translations of a word into the two languages of a bilingual subject facilitated each other. If this facilitatory effect between words of two languages was found, then it could be concluded that their internal representations are strongly associated.

The authors carried out five different experiments using both formal and semantic priming. They found facilitatory effects between languages using semantic priming, confirming the existence of a common semantic level for both lexicons. They also found facilitatory effects of a very short duration using formal priming. Thus they concluded that words of the two languages might be associated but that the links between the two lexicons were weaker than links from lexical to semantic level.

Since Kirsner *et al.* (1984) used as translations words that were not cognates (that is, not similar in their phonology and/or orthography) they pointed out that maybe another type of organization could account for the organization of the cognate words. They did not explore this relationship further, but their remark is very important, as will be seen in the next pages.

Although the work of Kirsner *et al.* (1984) seems to treat the bilingual lexical organization from a wider point of view, taking into account the previous research on memory to draw conclusions, Potter *et al.*'s (1984) Concept Mediation model has received more empirical support. As important examples of some of this research focusing on the convergence of the two languages at the semantic level, we mention Schwanenflugel & Rey (1986), Frenck & Pynte (1987), and Chen and Ng (1989). The last two authors evaluated facilitatory effects between languages, and they found that these effects are larger when the prime is a translation from the other language that when it is a semantically related word of the same language. This result is predicted by the Concept Mediation Hypothesis, because when presenting a word and its translation only one conceptual node should be activated, while presenting two related words two conceptual nodes need to be activated. In a different experiment they found that facilitatory semantic effects are comparable to the priming effect found with picture primes. In their conclusions they claimed that the mental processes implied in the priming paradigm are universal and independent of the differences between languages.

In next pages the language-specific factors in bilingual language processing will be reviewed. Next, some of the research devoted to learning as a factor in bilingual lexical organization will be summarized as well.

3.2. Factors in Bilingual Lexical Organization and Access

3.2.1. Language-Specific Factors

Some important research has focused on language-specific factors in order to find out if linguistic differences in a bilingual's languages would imply language processing differences. Obviously, the existence of language-specific factors concerns the internal organization of bilingual languages and the bilingual lexicon. As Aaronson and Ferres (1986) pointed out, when language-specific differences imply differences in cognitive linguistic information processing, it is necessary to find where they are in order to know specifically in which level of processing they can produce their effects.

In this section only the factors that concern the lexical level of processing, namely orthography and speech segmentation, will be presented.

A typical focus of research has been orthography. Katz & Feldman (1983) developed the orthography depth hypothesis, claiming that in languages with deep orthography (like English) written language is processed analogically, while in languages with shallow orthography (like Spanish or Serbocroatian) the readers use articulatory coding to access the lexicon. Other experiments supported this hypothesis (Bentin, Bargai, & Katz, 1984; Bentin & Frost, 1987; Chen & Juola, 1982; Chitiri, Sun, Willows, & Taylor, 1992); but there is research showing that shallow-orthography readers not only use articulatory decoding specifically for reading, but also read in an analogical

It is interesting to bring these two findings together: Phonological segmentation appears to be a language-specific process, whereas there is evidence for orthographical decoding as an universal process. The different acquisition of both representations might play an important role.

3.2.2. The learning factor

This aspect has been studied by Chen and his collaborators (Chen & Leung, 1989; Chen, 1990; Chen, 1992). Chen proposed different models of lexical organization for different learning stages during the acquisition of a second language and for the different developmental stages of the subject when learning the second language. In Chen & Leung (1989), evidence is provided by some experiments comparing the performance of beginner adults, beginner children, and those proficient in both languages. The authors propose that in the first stages of learning a new language this new language is accessed through formerly built up cognitive systems. Thus, an adult will use the first language to acquire the new one, and the language organization in memory can be described by the Word Association Hypothesis[1]. Children instead will use a prototypical-image system as the mediating system, since their first language system is not developed enough to support another language. Their internal organization then can be described according to the Concept Mediation Hypothesis.

1.Note the similarity of this proposal to Weinreich's description of the *Subordinate Bilingual.*

Other authors, like Favreau & Segalowitz (1982) and Frenck & Pynte (1987) corroborated Chen's conclusions, finding differences between proficient and beginner subjects. Abunuwara's work (1992) on trilingual subjects' internal organization gave support to the hypothesis of different lexical organization according to the stages of learning. This author showed that the languages of a trilingual speaker are related to each other in different ways. His subjects were speakers of Arabic, Hebrew and English, and he found out that while Arabic and Hebrew (L1 and L2) were strongly related at the lexical level, as the Word Association model would predict, English and Hebrew (L3 and L2) seemed not to be connected at all, as suggested by the Concept Mediation Hypothesis.

Recently Kroll and her colleagues (see Dufour & Kroll, 1995, for a review of her different works) supported Chen's model by describing a different effect called Translation Asymmetry[1]. According to Kroll, translation from L1 to L2 takes more time than translation form L2 to L1 in the early stages of second language learning. This difference decreases as proficiency in second language is achieved. The proposed model explains translation asymmetry: translation from L1 to L2 requires conceptual access, whereas translation from L2 to L1 can be accomplished directly by the links between words in the two languages at the lexical level.

In spite of the different learning strategies, Chen pointed out that the newly acquired language gradually develops an independent status, and the fluent bilingual lexicon is organized according to the Concept Mediation Hypothesis. Overall the results point out the importance of learning stages in bilingual lexical organization.

1.Keatley, Spinks & De Gelder (1994) reported the same effect.

3.3. Discussion

As De Groot (1992) pointed out, the models for bilingual lexical organization are usually extremely simple and sometimes these models fail to account for the complexity of reality. As has been seen in the previous sections, the authors within the bilingual research field distinguish two representational levels for bilingual memory organization: the lexical and the conceptual level. A description of the organization for each level is as follows:

1. At the lexical level, each entry is represented by a single node. The information this node accounts for is normally not specified. It can be assumed that the node represents all the information about the word except for the meaning, the latter being represented at the conceptual level. Usually the modality of the representation is not specified in bilingual models. That means these models do not specify if the representation at the lexical level refers to the orthographic representation of the word, the phonological representation, or both of them.[1] Hence much research being devoted to the automatic vs. strategic activation of phonological representations of words during visual word perception (Coltheart, 1978; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Seidenberg, 1985; Seidenberg, 1987; Tannenhaus, Flanigan, & Seidenberg, 1980; Van Orden, 1991, among others) seems to be ignored by researchers working on the bilingual lexicon. Another aspect that is never described, not even in the research on the learning factors just introduced, is how the new nodes for L2 representations are created.
2. The meanings of the words are stored at the conceptual level. How this

1.See Doctor & Klein (1992) for an exception: the authors describe a model partially based on the Double Route model and they specify Phonological and Orthographical Lexicons.

meaning is represented is generally not specified. In general, the conceptual level is supposed to be a semantic network where the related meanings are linked together. These links explain the Associative Priming Effect within and between languages. As this project is concerned with the lexical level, the conceptual level will not be further discussed.

4. Parallel Access to Bilingual Lexicon

Grainger & Dijkstra (1992) took a different point of view to describe the internal organization of the bilingual lexicon. Grainger and Dijkstra consider that there is a language network for each of the bilingual's languages, formed by local nodes representing the words. The authors propose a theoretical framework to explain bilingual lexical access, named Bilingual Interactive Activation, based on the Interactive Activation Model proposed by McClelland & Rumelhart (1981).

This theoretical framework (the authors insist on the fact that it is not a model, only a framework to set further research) describes a three level network like McClelland & Rumelhart use. The first level is composed of nodes identifying letters, the second of nodes identifying words and last one contains language-specific nodes. Grainger & Dijkstra's proposal tries to explain why lexical access is not language-selective: language is not identified before the words have been accessed. Thus, word-nodes would be connected with the proper language-node with a one-way connection; and this would explain why lexical access is not controlled by superior processing

In a subsequent article, Grainger & O'Regan (1992) suggest that a feedback connection should exist between language-nodes and word-nodes. They describe an activating connection from language-node and the correspondent word-nodes, and an inhibitory connection from language-nodes to the opposite language word-nodes. .

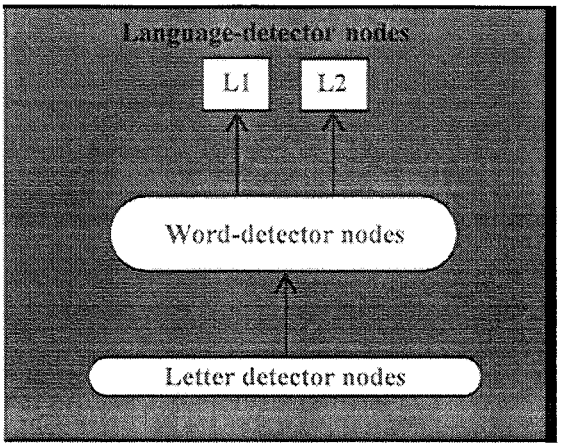

Figure 1. 6. Bilingual Interactive Activation Model (Grainger and Dijkstra, 1992)

Grainger & Dijkstra (1992) introduced the use of connectionist framework to explain the specific phenomenon of language identification in lexical access in bilinguals, but as they used local representations, their point of view is not far from the lexical-instance models that have been reviewed in previous sections. Perhaps their most interesting suggestion is that language information is accessed post-lexically. Some of Grainger's previous works had shown empirically that language identification occurs after word identification (Grainger & Beauvillain, 1987; 1988; Beauvillain & Grainger, 1987). This implies that language specification might not be important at the lexical level, thus questioning the need for language-specific representations. The next section concerns this point.

5. The overlapping bilingual lexicon

This section introduces the theoretical tendency on bilingual research developed during the last years. The main feature of this orientation is that the bilingual lexicon is not described as two independent lexicons, but as a single lexicon that contains the representations of the words of the two languages. This bilingual lexicon is partially language-specific and partially common: some word representations are represented in the language-specific parts, and some word representations are shared by the two languages. The underlying concept for this lexical structure is, as introduced in last section, that language identification occurs after lexical access. Therefore, the bilingual lexicon should be organized by factors other than language of the words. Accordingly, some researchers (Beauvillain, 1992; Cristoffanini, Kirsner, & Milech, 1986; Sanchez-Casas, Davis, & Garcia-Albea, 1992) consider that lexical representation is independent of bilingual experience, and so lexical access, either within as between languages, is governed by orthographic or associative principles. These authors suggest that although mental representations of lexical forms are necessarily language-specific, lexical organization is not governed by language (Beauvillain, 1992). The model suggested to account for this organization is described in the next Section 5.1.

A number of empirical studies supporting this theoretical approach have been centered on exploring the cognate effect. The cognate effect is a facilitatory effect reported in cross-language experiments using the masked priming paradigm. Both the masked priming paradigm and the cognate effect are described in the next pages. The empirical research on the cognate effect will be introduced and discussed in next chapter, where the experiments done for

this project are also reported.

5.1. Accessing a Common Lexicon

This new approach to bilingual lexicon organization was first suggested by Cristoffanini, Kirsner & Milech (1986). Cristoffanini *et al.* reviewed contradictory results in the literature: using as a prime the translation of the target word in a lexical decision task produced in some experiments a clear facilitation effect, but not in others (Kirsner *et al.*, 1984; Chen & Ng, 1989). These results can be explained assuming the complete independence of the lexicons, as some authors did; but they can also be explained by a different lexicon structure, based on the common morphology of words. Facilitation then takes place when the translations share orthographic features. Thus the authors concluded that morphology rather than language identity controls lexical functioning.

The experiment carried out by Beauvillain (1992) tried to show that lexical access in bilingual subjects is governed by the same principles as monolingual lexical access: basically, the frequency of morphological form, as Forster (1976) described. Beauvillain carefully chose five-character English and French stimulus-words, considering the frequencies of the bigrams and trigrams that formed them. The combinations of letters that occur most frequently in one or the other language were considered language-specific, while the words with combinations equally frequent in both languages were considered non language-specific. The subjects, with equal proficiency for English and French, needed more time to recognize non-language specific

words, independently of the language to which they belonged. Instead, when the task was performed by monolingual subjects in both languages, the reaction time was the same for specific and non-language specific words. These results show the sensitivity of lexical access to the orthographic features of the words.

Beauvillain concluded that there is a strong superposition of both language lexicons in bilingual subjects (see figure 1.7), and that the superposition is achieved by storing together words that share the same morphology (i.e. non-language-specific). Language does not govern lexical access in bilinguals, although this does not imply that language information is not represented in the lexicon.

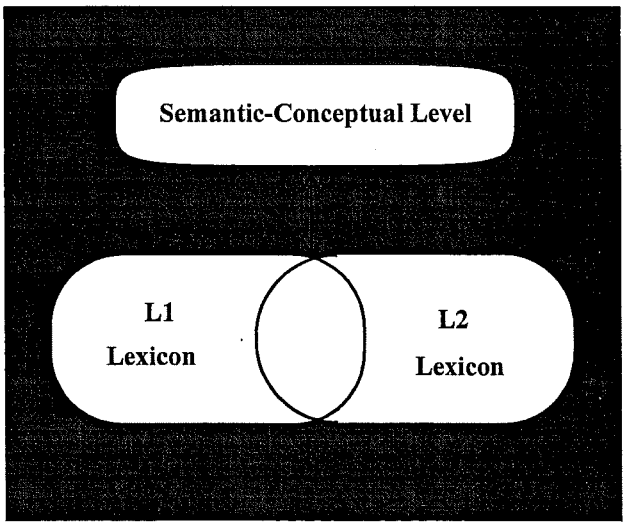

Figure 1. 7. Beauvillain (1992) partially overlapping specific lexicons

Recently, the masked priming paradigm has been used to test between-language facilitatory effects at the lexical level, bringing up some more evidence for a common lexicon. These empirical contributions are reviewed in the next subsection.

5.2. Masked Form Priming Paradigm

Form priming has been considered useful for exploring the organization of lexical information. This priming paradigm implies presenting a word as a prime with similar form to the target, as for example when a word is presented in lowercase (word) as a prime for the same word as a target in uppercase (WORD). The facilitation obtained with this technique is expected to be due to the similar form of the words and not to any meaning relation. Hence, the experimental results can show how the words are related at the lexical level, not making implications about the conceptual or semantic levels.

Segui & Grainger (1990) pointed out that in word recognition very little research has been carried out at the level of formal representations of words in memory, compared to the amount of research devoted to semantic and conceptual organization. The same applies for bilingual word recognition: form priming has been little used in bilingual research compared to semantic priming. Moreover, until the present decade the studies reporting form priming effects in cross-language experiments presented contradictory data (Kirsner, Brown, Abrol, Chadha, & Sharma, 1980; Kirsner *et al.*, 1984). In recent years, the results obtained with form priming have been more consistent and reliable since the masked priming paradigm was applied.

The advantages of masked priming were described by Forster. Forster's studies on Monolingual Form Priming (Forster, 1987; Forster & Davis, 1984) showed that the facilitation obtained with form priming could be due to subject's expectancies and integration of the meaning, especially when the target is presented a long interval after the prime (long Stimulus Onset

Asynchrony.). When the SOA is long, subjects have enough time to retrieve semantic information, thus the facilitatory effect observed is due to semantic association and not to the relation of forms at the lexical level. In other studies the form priming effect was not found (see Forster & Davis, 1991, for a review) and this lack of effect could be due to the inhibition generated at the semantic level.

To eliminate any interference from the semantic level, Forster & Davis (1984) used the masked priming technique, described by Evett & Humphreys (1981) for tachistoscopic identification. The masked priming technique involves a three part display. First, a presentation of hatch marks (##########) appears for 500 ms and acts as a backward mask for the second item, the prime, presented in lowercase for a very short duration (typically, 30 to 60 ms, but it can also be shorter). The third item, presented in uppercase, is the target which is to be classified by the subject. The target acts as well as a forward mask for the prime.

When the prime is presented in this way the subject is unaware of its presence. Forster & Davis (1984) suggest that in that way episodic traces are relatively inaccessible, thus emphasizing the lexical component in lexical access. Hence, masked form priming appears to be a better tool than form priming for investigating the role of form in lexical organization.

5.3. The Masked Priming Paradigm in Cross-Language Experiments: the Cognate Effect

Recently, masked form priming has been used to study bilingual lexical organization. The specific advantage of masked priming in bilingual studies is that since the primes are not detected by the subjects, they perceive the experiment to be in the language of the targets (Alpitsis, 1990).

Garcia Albea, Bradley, Sanchez Casas and Forster (1985) compared the facilitation effect obtained by repetition priming (the same word presented as prime), cognate priming (the translation of the target as prime, with similar form) and non-cognate priming (the translation as a prime, with a different form). They found that both repetition priming and cognate priming facilitated target processing. Moreover, their results showed no differences between the magnitude of the effect obtained with repetition priming and cognate priming. This facilitatory effect of cognate words, comparable with the effect obtained with the repetition of the same word, is known as the Cognate Effect.

The cognate effect was also reported by De Groot and Nas (1991, experiment 4). De Groot and Nas did not find associative priming effects for non-cognate translations, but they found strong effects under masked priming conditions for cognate translations. Sanchez-Casas *et al.* (1992) also reported strong priming effects for cognate translations in a cued translation task.

Overall this evidence shows that the cognate effect under masked priming is very robust between languages. The model that has been proposed to account

for the cognate effect is that of Beauvillain (1992) described in Section 5.1. (see fig. 7). As already mentioned, the words that are morphologically similar in both languages (the cognate words) are stored in the overlapping part with a common representation. Since cognate words share the same internal representation, the facilitation produced by using as prime the cognate translation of the target is the same as the facilitation obtained using as prime the target itself, because in both cases prime and target activate the same representational node.

This model contains some weaknesses that we will try to summarize here. First of all, and in line with other models of bilingual lexical structure, the model does not include any description of how phonological information is represented. Secondly, words with common morphology within one of the languages may be difficult to place: if the basic rule is that words with the same morphology are stored together in the overlapping part of the lexicon, then morphologically similar words within and between languages should share the same representational node in this part of the lexicon. Thirdly, interlingual homographs (words with exactly the same orthographic form but completely different meanings) present some representational problems as well: with respect to the conceptual level, these words are language-specific, whereas with respect to their form they are not. Last but not least, there is an important omission with respect to factors such as frequency and neighborhood size. This information is by definition language-specific and there is no description of how it would be represented in a single node for two different words.

Chapter 3 of this work discusses the empirical research that support this model, both from monolingual and bilingual research, and reports about the

two experiments done within the present project.

Next section presents a discussion about the suitability of the theoretical models introduced in this chapter and proposes a new theoretical framework to describe the bilingual lexical organization.

6. General Discussion

We will try to summarize different research lines in the literature and propose a new approach to the study of bilingual lexical organization.

It is not difficult to see that existing models on bilingual lexical organization tend to merge the lexicons. From the first proposals of Kolers, suggesting a complete different linguistic system for each language, up to most recent ones suggesting an overlapping lexicon, language specificity loses its place as the most important organizational factor for internal lexical representations.

An important body of evidence, coming from a different area of research, concerns the interaction of orthography and phonology during visual word perception. This evidence seems to indicate that the activation of both codes occurs simultaneously during word perception. A new line of research centers now on the study of perceptual units smaller than words, for example, the syllable (Alvarez, 1995; Dominguez y Cuetos, 1995) or parts of the syllable (Taft, 1995). This evidence supports the activation of phonology during reading, but also indicates that sublexical units play an important role in visual perception of the word.

Using the masked priming paradigm, two findings are especially remarkable: the importance of the neighborhood on priming effects (Segui and Grainger, 1990), and the cognate effect observed across languages (Garcia Albea *et al.*, 1985; Alpitsis, 1990; Sanchez-Casas *et al.* 1992). Both findings refer to the similar morphology as a decisive factor in lexical organization and bilingual lexical organization.

Overall the existing evidence seems not to be explained by current models, since these models lack a proper description of the interaction of orthography and phonology, and they do not account for the Cognate Effect, as discussed in the previous section. As evidence seems to support a representation of lexical entries smaller than words, maybe a different theoretical framework should be adopted to describe bilingual lexical organization.

Seidenberg & McClelland (1989) suggested a new theoretical framework for word recognition and naming with a different internal lexical representation. Their model is described in detail in Chapter 2. They propose a distributed representation for lexical entries. According to the model, each lexical entry is not represented by a single node but by a pattern of activation through a number of nodes. In their model there is no lexicon in the sense that has been previously described. The lexical level is formed by a set of hidden units that are connected to two sets of input and output units. One of these two sets acts as encoder/decoder of the phonologic information of the words and the other as the encoder/decoder of the orthographical information of the words. These two sets are connected to the hidden units that represent the lexical level. The activation of the hidden units is a product of both phonologic and orthographical information, thus the pattern of activation obtained across the hidden units is a bimodal (orthographic and phonological) representation of each word.

The application of this kind of representation could be useful in a bilingual model. The models presented in this chapter represent each lexical entry with a single node of information; consequently, learning words of a second language implies the creation of new word-nodes in a new lexicon. Instead, using a distributed representation for lexical entries implies that there is no

need to add more representational nodes when the words of a second language are learned, because an infinite number of representations can be held by a pattern of activation across a limited number of nodes. Thus, language specificity as an organizational factor for lexical entries disappears and becomes a by-process, since the internal representations will be more alike when words are similar, independently of the language they belong to.

This type of representational model could account for the Cognate Effect, since cognate representations would have a similar representational pattern. This priming effect could be explained by the similarity of activation patterns for prime and target words.

As connectionist models are easy to simulate, a bilingual lexical model with the described features could be programmed. This characteristic represents another advantage over the classical models described in the previous pages because it means the model can be actually be tested. A new model for bilingual lexical organization has been designed and tested in this project. The model is called Bilingual Access Representations (B.A.R.) and two versions of it have been developed, which are presented next in Chapter 2.

Chapter 2

Bilingual Access Representations Model: Developments

1. Introduction

This chapter describes the development of a connectionist model for bilingual lexical organization. The Bilingual Access Representations model was designed after reviewing the current models of bilingual lexical organization. These models, as indicated in the previous chapter, were not suitable for simulating the learning of words during second language acquisition.

The Bilingual Access Representations Model is based upon the Model for Word Recognition and Naming developed by Seidenberg and McClelland (1989). This model used distributed representations for lexical entries, a feature that allowed the learning of new words without enlarging the model. Besides, it offered a way to include in the model orthographic and

phonological information. The working hypothesis for the implementation of BAR was that with this information the model could account for the cognate effect reported in the literature of crosslanguage research (see Chapter 1, Section 5: The Overlapping Bilingual Lexicon).

This chapter consists of three sections. Section 2 is a theoretical introduction to the model, and it reviews connectionist models of language processing as they relate to word recognition. Section 3 introduces the first version of the Bilingual Access Representations model (BAR 1), describing the design of the network, the training, and the obtained results. The analyses of BAR 1's performance resulted in the specification of the changes necessary for the second version. This second version (BAR 2) is described in Section 4.