*Psychologica Belgica*
1998, 38-3/4, 177-196.

# SIMULATING CROSS-LANGUAGE COMPETITION WITH THE BILINGUAL INTERACTIVE ACTIVATION MODEL

Ton DIJKSTRA and Walter J.B. VAN HEUVEN
*University of Nijmegen, The Netherlands*

Jonathan GRAINGER
*CNRS, Aix-en-Provence, France*

The BIA model is an implemented language-nonselective access model for bilingual word recognition which assumes an integrated lexicon for words of different languages. In this paper, simulations with the BIA model are presented that account for the effects of lexical competitors from the same and another language on target word recognition in two different paradigms. First, the model simulates recent lexical decision results obtained with French/English bilinguals in two masked orthographic priming experiments (Bijeljac-Babic, Biardeau, & Grainger, 1997). Differences in the result patterns for subject groups varying in second-language (L2) proficiency (monolinguals, beginning, and more advanced bilinguals) are simulated by assuming a reduced frequency range for L2-words. In demonstrating that the monolingual participants in this study very likely knew some L2-words, the model displays considerable heuristic value. Second, the BIA model simulates the effects of intra- and interlingual neighborhood density on target word recognition in blocked and mixed progressive demasking (PDM) with Dutch/English bilinguals (Van Heuven, Dijkstra, & Grainger, 1998). The model accounts for a shift in result patterns observed in high-proficiency subjects over parts of the blocked PDM experiment by varying the degree of asymmetric cross-language inhibition from language node to word level. Finally, the model simulates the over-all result pattern in the mixed PDM experiment by manipulation of both cross-language inhibition and the frequency range for L2-items. A comparison with some other nonselective and selective access models indicates that the BIA model produces better overall simulation results because the two mechanisms allow the model to capture the asymmetric cross-language effects observed in unbalanced bilingual participants.

## Introduction

How word forms of different languages are stored in the mental lexicon of bilinguals and how they are accessed during word recognition have long been major issues in research on bilingualism (Dijkstra & Van Heuven, 1998; Grainger 1993; Keatley, 1992). With respect to lexical storage, a distinction has been made between the hypotheses of language-independent and language-dependent storage. The first type of hypothesis proposes that bilinguals possess

Correspondence concerning this article should be addressed to Ton Dijkstra, NICI, University of Nijmegen, P.O. Box 9104, NL-6500 HE Nijmegen. Electronic mail may be sent to dijkstra@nici.kun.nl.

one integrated lexicon for both their languages, while the second type assumes two separate lexica, one for each language. With respect to on-line processing, an analogous contrast has been postulated between language-selective and nonselective access views. Here the issue is whether word forms of both languages are activated during word recognition or only word forms of the language that is targeted in the communicative setting.

In the last 25 years, a large number of studies have collected evidence that word recognition in one language may be affected by knowledge of words from the other language, while a few have found support for the language-selective access hypothesis (see Brysbaert, this volume, for an overview of the empirical evidence). The prevalence of studies supporting language-nonselective access suggests that a selective access hypothesis is no longer tenable. The seemingly conflicting results in the experimental literature can be reconciled by assuming that bilingual word recognition basically involves a language-nonselective processing system which may function more or less selectively depending on experimental circumstances. In other words, the state of activation of a bilingual's languages is not rigidly fixed in either a monolingual or a bilingual setting, but fluctuates depending on (as yet unknown interactions between) "top-down" factors (task demands, type of instruction, and so on) and "bottom-up" factors (stimulus list composition in terms of types and language of items, etcetera) (Dijkstra, Van Jaarsveld, & Ten Brinke, 1998; Grosjean, 1988, 1997; Li, 1997).

A recently implemented language-nonselective bilingual word recognition model that can in principle account for such cross-experiment differences in task demands, stimulus list composition, and subject second-language proficiency, is the Bilingual Interactive Activation (BIA) model (Dijkstra & Van Heuven, 1998; Grainger & Dijkstra, 1992). As well as being nonselective, the BIA model also assumes language-independent storage, in the form of an integrated lexicon for the words of different languages.

It might be assumed that this model characteristic follows from a "natural" relationship between the language-nonselective access hypothesis and the language-independent storage hypothesis. However, matters are not that simple. Another theoretical option is to assume separate storage for words from two different languages, but parallel access to all words irrespective of language on the basis of an input string (Smith, 1997, p. 149, 164). For instance, an English input word could simultaneously and in parallel activate lexical competitors (alternative word candidates) from both Dutch and English. Just like the integrated lexicon hypothesis, this hypothesis of language-independent lexica but parallel activation might account for both the absence and the presence of cross-linguistic effects in experiments. Whether or not the non-target language generates interference effects would depend on the extent to which non-target language activation influences a later decision process.

Crucial to testing the two hypotheses against each other is to show that cross-language effects arise during the process of identifying the input word, and not (only) later during a decision stage. In other words, the activation level of a lexical candidate from one language (e.g., CORK) should be affected not only by the presence of form-similar candidates (or "neighbors") from the same language (e.g., COOK), but also from the other language (e.g., VORK).[1] One way to investigate this issue is by experimentally manipulating the lexical context of a particular target word. In the first study discussed below (Bijeljac-Babic, Biardeau, & Grainger, 1997), this was done by briefly presenting, just before the target word appeared, form-related or unrelated masked orthographic primes from the same or the other language. However, it could be argued that this type of lexical context is not normally present during word recognition. Another type of manipulation, originating from the monolingual domain, may provide more convincing evidence here. It involves an examination of the effects of cross-linguistic neighborhood density within and across languages on target word recognition. In the second study to be discussed (Van Heuven, Dijkstra, & Grainger, 1998), the number of neighbors of a target word in the two languages (many or few) was varied. To the extent that the effects observed in these studies reflect cross-linguistic effects during an identification stage, they provide evidence against the hypothesis of independent lexica but parallel access.

However, the major focus of the paper will be on demonstrating to what extent the BIA model mentioned above is able to simulate the experimental results in these studies. The BIA model will be shown to do a better job than either strictly selective access models (which behave like monolinguals with respect to the target language) and nonselective access models that assume an integrated lexicon with two equally strong language subsets (and thus behave more like balanced bilinguals). The better performance of the BIA model is due to the incorporation of several mechanisms to simulate characteristics of unbalanced bilinguals as well as the cognitive flexibility of bilinguals having a high L2 proficiency. As we shall see, asymmetric data patterns across the first- and second-language in an unbalanced bilingual may originate from (1) subjective frequency differences between L1 and L2 (modeled in terms of differences in resting-level activations of word units) and (2) top-down feedback effects from each lexicon as a whole to the word level (modeled by means of so-called language nodes). The paper will end with an evaluation of some strong and weak points of the current version of the BIA model.

---

[1] Note that it is hard, if not impossible, to distinguish the independent and integrated lexicon hypotheses by means of experiments involving cross-language homographs or cognates. Precisely because homographs and cognates share their form representation across languages, parallel activation of lexical candidates from two languages would follow.

## The BIA Model for Bilingual Word Recognition

The BIA model, illustrated in Figure 1, shares its basic architecture with the monolingual Interactive Activation (IA) model (McClelland & Rumelhart, 1981, 1988). In extending this widely supported model for monolingual word recognition to the bilingual domain, we have adhered to important modeling principles such as *parsimony* and *generalizability across domains* (Jacobs & Grainger, 1994). Furthermore, the BIA model as such incorporates the IA model as a special case. This modeling strategem, called "*nested modeling*" by Jacobs and Grainger (1994), has as one important advantage that simulations done with the IA model can be exactly reproduced by the BIA model.

As can be seen in Figure 1, the BIA model consists of four layers of nodes, corresponding to different types of representations: letter features (14 for each letter position), letters (26 for each position in a word), words (a combined lexicon of 1323 English and 982 Dutch four-letter words), and languages (one node for each language). As in the IA framework, there are connections between nodes at each level as well as between nodes of different levels. All units at the word level are negatively interconnected, i.e. they can mutually inhibit each other's activation. A letter unit activates all words that contain that letter in the correct position, and inhibits them (reduces their activation) if there is a mismatch. Activated words also feed activation back to their constituent letters. The parameters which regulate these interactions all have the same settings as in the original IA model (McClelland & Rumelhart, 1981, 1988).

If a letter string (word or nonword) is presented to the model, letter features become active for each position of a letter in the string, and these features subsequently activate letter nodes at their respective positions. Letter nodes in turn activate the word nodes they are connected to. Next, the word nodes activate their language nodes and also send activation back to the letter level. In addition, active word nodes inhibit other word nodes, and language nodes send top-down inhibition to word nodes from the other lexicon. Thus, after a few time steps, a complicated stream of activation arises between nodes at different levels. Gradually, the word candidate that best matches the input will become most active. Recognition is assumed to take place as soon as a word's activation exceeds a particular threshold level. How long this takes depends on several characteristics of the word itself (e.g., its frequency) and its relationship to other words in the lexicon (e.g., the number and frequency of similar words or neighbors). Given the complexity of the activation process, simulations are needed to establish the actual recognition times for different stimulus sets. Intuitions are often deceptive here. As McClelland and Rumelhart (1981, p. 382) themselves noted: "Although the model is essence quite simple, the interactions among the various nodes can become complex, so that the model is not susceptible to a simple intuitive or even mathematical analysis."

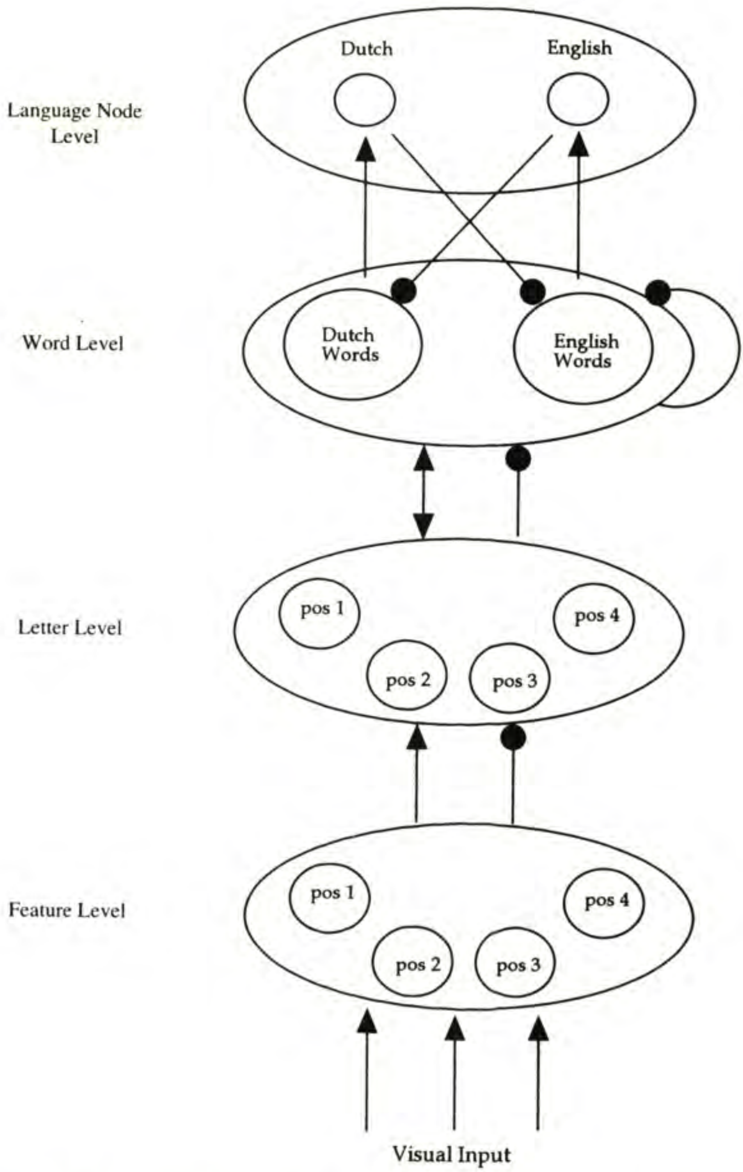



*Figure 1.* The Bilingual Interactive Activation (BIA) model. Excitatory connections are indicated by arrows (with arrow heads pointing in the direction of activation spreading), inhibitory connections by ball-headed lines.

*Word Frequency in Mother Tongue and Second Language*

As in the IA framework, differences in word frequency within and across languages are implemented in the BIA model in terms of differences in the (range of) relative resting-level activations of word units. Higher frequency words have a higher resting level activation when the recognition process starts and therefore a head-start in their ability to suppress other (less frequent) items. To the extent that differences in first and second language proficiency are reflected by differences in (subjective) frequency of usage of words from the two languages, they can appropriately be implemented as variations in resting-level activation. Simulations with the BIA model accept the same input words as human subjects react to in experiments, and the frequency data from English and Dutch corpora as incorporated in the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993) can be used as an estimate of their subjective frequency, at least for highly proficient bilinguals. The continuum of proficiency between a pure monolingual and a balanced bilingual may be simulated by varying either the number or the relative frequencies of the words in the two lexica, or both. Future analyses of the development of "real" human lexica over time are needed to determine how frequency differences can best be implemented.[2]

*Asymmetric Cross-Language Effects From Language to Word Nodes*

Special to the BIA model (apart from the incorporation of two lexica) are the language nodes, one for English and one for Dutch. The language nodes collect activation of all words from one lexicon and suppress all words in the other lexicon. The extra level of language nodes requires the setting of two additional parameters: word to language-node excitation and language-node to word inhibition. However, this last parameter may be differentiated in one parameter for English and one for Dutch. In this way, the top-down inhibition effects from the two language nodes can differ in size (asymmetric inhibition). Note that the BIA model would probably also be able to produce similar behavior without language nodes if there were asymmetric inhibitory links between words of different languages (e.g., L1-words might inhibit L2-words more than vice versa). However, the language nodes are theoretically useful in stressing the functional unity of a lexical system; in addition, sending (more or less) inhibition from the language nodes to the word level provides a mechanism

---

[2] The currently applied method assumes a reduced frequency range of words from the second language, implemented as changes in the resting-level activation of such words. For more information and details on the simulation studies (e.g., with respect to parameter settings), see Dijkstra and Van Heuven (1998).

enabling the simulation of context effects on the recognition of the currently presented word. For instance, the effects of task demands or words presented previously can be simulated by assuming (pre-)activation of the language nodes by sources external to Figure 1.

### *Model to Data Comparison*

In this paper we will show how manipulation of L2-frequency and variation of cross-language inhibition allows the simulation of a wide range of empirical results with the BIA model. How can the simulation data be compared to the empirical data? Several types of comparison are possible (Dijkstra & Van Heuven, 1998). First, and simplest of all, is the ordinal comparison. As soon as the average recognition cycle of an input word or of all words in a stimulus category is determined (either stochastically or deterministically), a comparison can be made with mean RTs. If the pattern of average cycle duration across conditions "looks similar" to that of the RTs, the model is assumed to "fit" the data reasonably well. Second, an interval comparison between simulation and experimentation data can be done by computing the Pearson correlation between cycle and RT means. Third, the empirical data (RT means or medians for items in different conditions) and the simulation data (cycle means for items in different conditions) can both be subjected to an analysis of variance by items. Subsequently, it can be examined it the two analyses result in similar patterns of main and interaction effects. Finally, a strong test of model fit can be done by computing a chi-square on the basis of the minimal squared deviation between simulation and experimental data. If the observed chi-square is not significant, this indicates the model cannot be rejected on the basis of the empirical data. We will apply the last (chi-square) and the first ("qualitative comparison") techniques to the empirical data discussed in the next two sections.

## Simulating Cross-Language Masked Orthographic Priming Effects

A first approach to studying cross-language effects of form similarity on target recognition is to collect lexical decision times to target words preceded by masked orthographic prime words from the same or a different language. Bijeljac-Babic, Biardeau, and Grainger (1997) performed two such experiments with French/English bilinguals. In their first experiment the target language was English, in the second experiment it was French. The target words in both experiments were low-frequency words or nonwords that were orthographically legal and pronounceable. The prime words were high-frequency words (presented for 57 ms) that either belonged to the same language as the target words or to the other language. Both experiments showed that when prime and target words were from the same language, target word decision times were inhibited for orthographically related primes relative to orthographically dissimilar primes. For example, when the target language was English, the prime-target combination "real - HEAL" led to slower RTs than the combination "roof - HEAL". This finding suggests that any facilitation effects that might arise due to form priming (overlap in letters) were annihilated by inhibition effects due to lexical competition. More relevant, however, is that when prime and target were words from different languages ("beau - BEAM"), inhibition effects were found as well (see Figure 2a). This indicates that lexical knowledge from the other language affected target recognition.

Experiment 2 showed that the size of this cross-linguistic inhibition effect increased as a function of the participants' level of proficiency in the prime word's language. The response patterns of high-proficient bilinguals showed a considerable within-language inhibitory priming effect (28 ms) and a trend towards even larger cross-language inhibition (43 ms). Since the primes had very short durations and were not consciously recognized, a tentative conclusion is that the target was affected by the prime during the identification stage and not a later decision stage. If primes and targets had activated two different and independent lexica, no cross-language inhibition would have been expected to arise under these circumstances.

### *Simulating Within and Between Language Effects of Masked Orthographic Primes*

To simulate the results of this study, a French and an English four-letter lexicon were incorporated in the BIA model. Subsequently, simulations with a fixed recognition threshold were run for most prime-target pairs tested in Experiment 1.[3] At the first processing cycle, the prime word was presented to the model and at the third cycle it was replaced by the target word. The results of this simulation are presented alongside the experimental results of Experiment 1 in Figure 2a/b. The simulation not only captures the longer average RTs

[3] The simulations presented here extend those reported by Bijeljac-Babic et al. (1997). In our simulations we eliminated a few prime-target pairs that consisted of morphologically complex items (e.g., "GUNS") or could be considered as interlingual homographs (e.g., "CHAT"). However, the general result pattern turned out to be quite robust and independent of the specific selection of items. Furthermore, in our BIA simulations we could not incorporate all items of the study's second experiment, because they consisted of both four- and five-letter words. However, simulations with the subset of four-letter words resulted in the same data patterns as for Experiment 1.

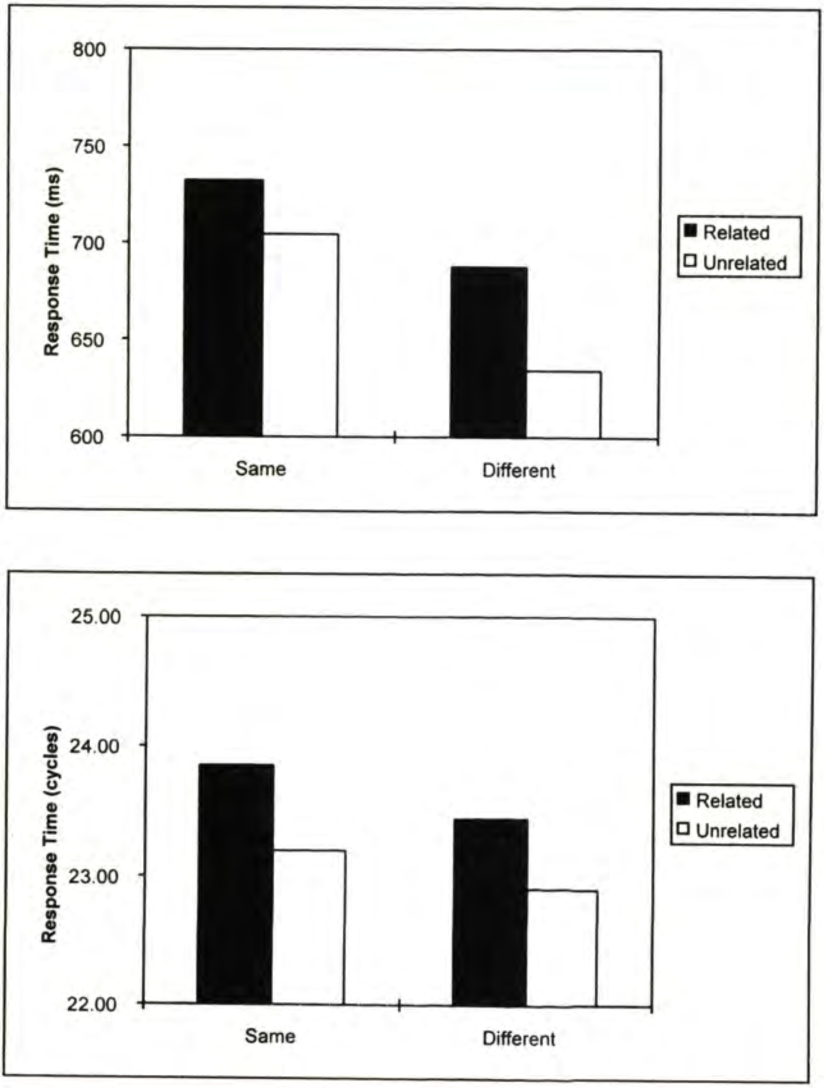


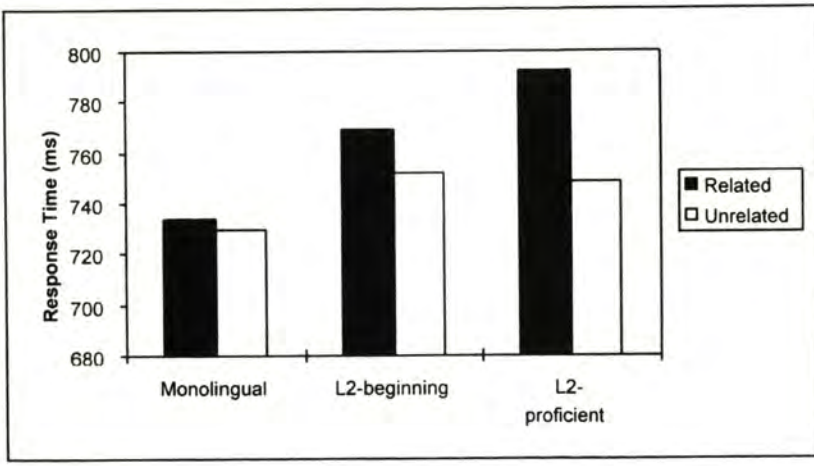

*Figure 2a and b.* Empirical (a) and simulation (b) data for English target words preceded by orthographically related or unrelated English (same) and French (different) prime words from the lexical decision study by Bijeljac-Babic, Biardeau, and Grainger (1997).

in the within-language prime condition but also the trend towards larger inhibition effects in the between-language prime condition.

*Simulating Masked Orthographic Priming Effects for Different L2-Proficiency Groups of Subjects*

Subsequently, we examined whether the model could simulate the different data patterns for the three L2-proficiency groups run in Experiment 2 by Bijeljac-Babic et al. (1997). The result pattern for the more advanced bilinguals was the same as in Experiment 1, for which BIA simulations were presented in the previous section. The results for the monolinguals were simulated as follows. First, we removed the second-language lexicon from the model, which turns the model into the well-known monolingual IA model. Next, we followed the same procedure as above, presenting the primes (which are now nonwords to the model) for two cycles, followed by the target items for the remaining cycles until recognition. Interestingly enough, the results of the simulation, presented below the empirical results in Figure 3a/b, deviate from those of the experiment in the sense that the model predicts facilitation for the related condition rather than null-results. According to the model, the overlap in several letters between nonword prime and target word should have resulted in faster target recognition. In fact, the model predictions are in line with several studies from the monolingual literature. For instance, nonword primes produced facilitation effects on target words in studies by Ferrand and Grainger (1992, 1993), and Grainger and Ferrand (1996).

The simulations suggest an explanation for the deviating experimental data in the present study. It may be that the monolingual subject group participating in the experiment was, in fact, not completely monolingual. Examining the stimulus material, this option does not seem unlikely. It would appear that even so-called French monolinguals might know English words such as FIRE, MILE, or SALT. To test the likelihood of this possibility, we ran another simulation with the same stimulus material, but now incorporated the four-letter stimulus words as items in a second-language lexicon with a frequency that was much lower than their original frequency. Indeed, it proved very easy to replicate the empirical pattern of results obtained by Bijeljac-Babic et al. in this manner. The simulation result for the intermediate subject group is also presented in Figure 3b. We are currently preparing an experiment in which we will attempt to induce the expected data shift across proficiency levels (reflecting a transformation from nonwords into low-frequency words) in one group of subjects by training them in an artificial language. Thus, in the current set of simulations the BIA model has demonstrated considerable heuristic value (cf. also Dijkstra & Van Heuven, 1998).



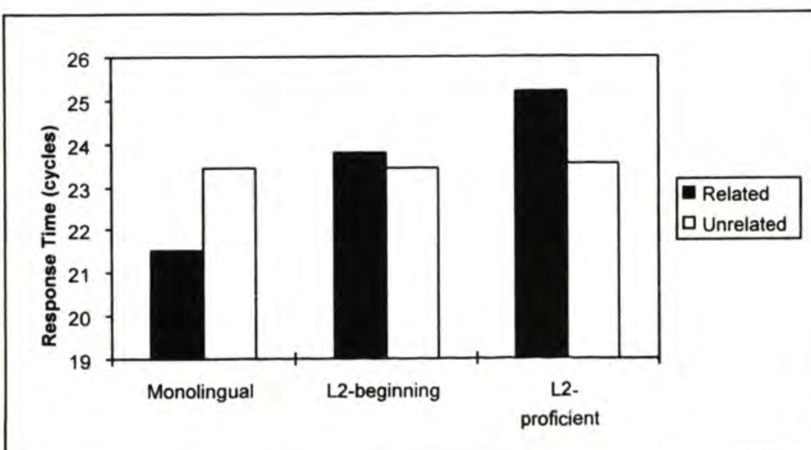




*Figure 3a and b.* Empirical (a) and simulation (b) data for the three L2-proficiency groups of subjects of Bijeljac-Babic, Biardeau, and Grainger (1997): Monolinguals (English lexicon only, 1622 words), beginning bilinguals (English and French lexica, with strongly reduced French word frequencies), and more L2-proficient bilinguals (English and French lexicons, 1622 and 792 words).

## Simulating Cross-Language Orthographic Neighborhood Effects

In the masked orthographic priming study just described, interference was observed in the related prime-target condition (e.g., "beau - BEAM") relative to the unrelated condition (e.g., "donc - BEAM"). The interference was ascribed to competition during the target identification process, suggesting that the view of two independent lexica must be rejected. However, this conclusion may not be totally convincing. It could be argued that, even though the prime is never consciously seen by the subject, it "naturally" activates the second lexicon, because there is a perfectly matching word form in that language. The word candidate activated by the prime ("beau") could then enter into a competition with the target word ("BEAM") during a late decision process. This possibility is excluded in a second paradigm (progressive demasking) that we will now discuss, because it is applied to letter strings that form words in one language only. Here lexical competitors of the target word (from the other or the same language) are never actually presented to the bilingual subject, who just has to identify a presented target word. Evidence for lexical competition exerted by the non-target language which is obtained with this paradigm will provide even stronger support for the hypothesis of an integrated lexicon.

Following up on Grainger and Dijkstra (1992), Van Heuven, Dijkstra, and Grainger (1998) examined cross-lingual effects on items that are unique to one language but overlap with items in the other language in all letters but one. In their study, they extended monolingual research on neighborhood effects to the bilingual domain. This research started from the definition of neighbors first proposed by Coltheart et al. (1977). According to this definition, a word's orthographic neighborhood is the set of words that share all but one letter in the same position with the target word. Thus, the words LORD, WOOD, and WORK are all neighbors of a target word WORD (see Grainger and Jacobs, 1996, for an overview of monolingual neighborhood density studies). For the bilingual domain, one can distinguish neighbors of a target word from within a language and those from the other language (Grainger & Dijkstra, 1992; Grainger, 1993). For instance, the word WORD also has a number of Dutch neighbors, such as BORD, WOND, and WORP.

Van Heuven et al. (1998) performed a series of experiments in which the number of neighbors from within the target language and from the other language was systematically varied. The experimental paradigms used were progressive demasking and lexical decision. In the progressive demasking (PDM) task, the presentation of the target word is alternated with that of a mask (a checkerboard pattern). During this process of alternation the target presentation time slowly increases while that of the mask decreases. The subject's task is to push a button as soon as the target word is identified. In two PDM-experiments, Dutch subjects identified both Dutch and English words which were presented, respectively, in two successive language blocks (Dutch-English or English-Dutch; Experiment 1) or during a mixed Dutch-English presentation (Experiment 2). The number of neighbors a test word had within and between languages was orthogonally varied. Dutch or English target items could have many or few neighbors in English (number of English neighbors: English N), and many or few neighbors in Dutch (number of Dutch neighbors: Dutch N).

The results of both experiments indicated that the identification of English target words was always affected by them having Dutch neighbors. The overall pattern is graphically illustrated by the data of the mixed experiment in Figure 4. The results seem to support at least two conclusions. First, a language-selective access model must be rejected. During access, the bottom-up input must initially activate words from both languages (in other words, in these task situations the subject is indeed in a bilingual language mode). It is hard to see how otherwise the presence of neighbors from another language could have affected target recognition times. Note that for the type of stimulus material here, the input word itself belongs exclusively to one language. This stands in contrast to interlingual homographs, where the input letter string is a word in more than one language. Thus, it is not due to the special characteristics of the homographs that language-nonselective results have been often obtained in the past.

In addition, the data strongly support the view that the two lexica of the bilingual do not form two independent systems. The combined influence of Dutch and English neighbors on word recognition indicates that lexical candidates from the other language affect the target identification process itself, even when the target words are presented in separate language blocks. The only way to save the nonselective-access independent-lexica hypothesis would be to make the decision stage sensitive to neighborhood density characteristics. However, this would in fact amount to an incorporation of cross-linguistic influences exerted by the non-target lexicon into the word recognition process.

In the next two sections we discuss how the specific results of the two experiments can be simulated by the BIA model, first considering the blocked experiment and then the mixed experiment.

### *Simulating Shifting Neighborhood Effects Across Experimental Parts in Blocked PDM*

In the blocked progressive demasking experiment, target items were presented in blocks of Dutch or English items. Some subjects saw the Dutch items first, others saw them only after the English items. Interestingly enough, the pattern of RTs produced by high-proficiency subjects was found to change over the

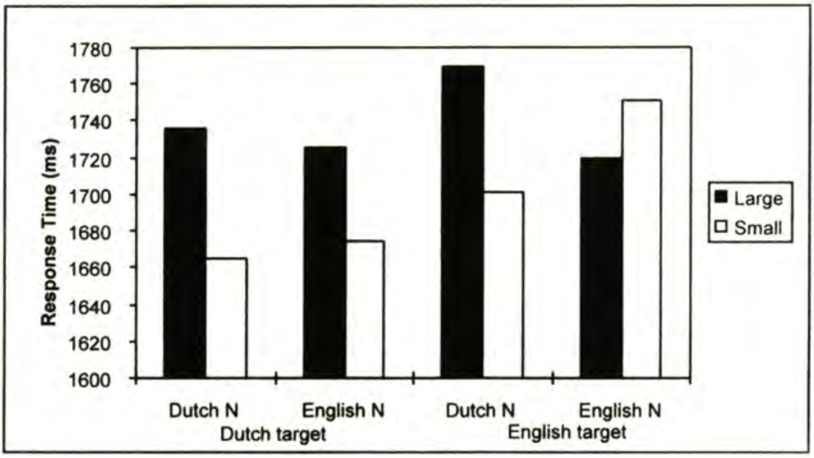



*Figure 4.* Reaction times to Dutch and English target words with large or small numbers of neighbors in English (English N) and/or Dutch (Dutch N), as obtained in the mixed progressive demasking experiment.

four parts of the experiment, as graphically illustrated in Figure 5a. An initial inhibition effect of English neighborhood size on Dutch target identification was observed, which decreased when the English items had been presented earlier. Thus, the Dutch RTs were strongly inhibited by the number of English neighbors when the Dutch words were presented in the first block. Closer examination revealed that the effect was especially strong in the first half of this first block, but was reduced in the second half. For the group of subjects who saw the Dutch words only after the English words, the effect of English neighbors was reduced even further and turned into facilitation in the last half of the second block.

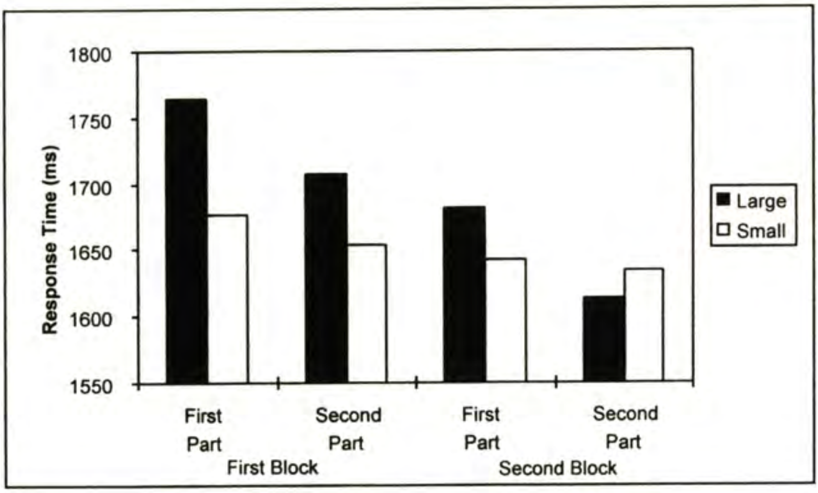
This finding suggests that high-proficiency subjects can in some way modulate the relative activation of their two languages. When they arrive in the experimental situation, they expect that their English proficiency will be tested (they know they are participating in some kind of bilingual experiment), and will therefore keep their English lexicon in an active state (Grosjean, 1997). However, after they have been tested with respect to their English word knowledge, they can focus on the Dutch task at hand and the relative activation of English may drop. This account is supported by simulations with the BIA model. As Figure 5b indicates, the model is able to simulate the pattern of results over time by varying the relative amount of (asymmetric) inhibition spreading from the Dutch language node to English words.

### *Simulating Cross-Language Neighborhood Effects in Mixed PDM*

In the mixed progressive demasking experiment, subjects identified English and Dutch words in lists composed of words from both languages. Figure 4 presented earlier shows the global results of the mixed experiment for English and Dutch target items. As will be remembered, the figure shows the effects of orthogonally varying the number of neighbors (large/small) of a target word in English (English N) and Dutch (Dutch N). A first glance at the figure already shows that both English and Dutch target items were affected by neighbors from the same and the other language, but not in the same way. While the effects of neighborhood density were inhibitory across languages for both Dutch and English target words, there was an inhibitory within-language effect in Dutch and a facilitatory effect in English. What type of bilingual word recognition model could account for such a complex pattern of results? Clearly, a language-selective word recognition model (such as the IA model, having only one lexicon), can never account for the combined results of the Dutch and English items obtained in the same experiment. A language-nonselective access model is required to mimick the behavior of these unbalanced bilinguals.

It can be argued that such a model needs (at least) two provisions. First, the subjective frequency of the English items must be lower than that of Dutch items to reflect the different levels of exposure of Dutch-English bilinguals to these two languages. We have already seen how a reduction in the frequency range for English words in the model can be effective in simulations. Figure 6 (left most panels) shows a simulation run on a variant of a language-nonselective access model with only such a *reduced frequency range*. Second, the model must be able to account for effects that depend on the experimental circumstances. For instance, differences in stimulus list composition and in task demands are known to affect the way subjects use their word recognition system (Grosjean, 1997; Dijkstra, Van Jaarsveld, & Ten Brinke, 1998). Such context-dependent effects can be simulated by varying the *asymmetric top-down inhibition* from the language to word level. We saw an example of this above when the shift in data patterns for high-proficiency subjects in the blocked PDM was simulated by assuming a gradual change in top-down inhibition in the course of the experiment. It turns out that, in the context of the mixed PDM, an optimal fit of the simulations with the empirical data can be obtained by introducing a limited amount of top-down inhibition. This can be seen by comparing the simulation data in the left panels in Figure 6 (model variant without top-down inhibition) and in the middle (variant with limited inhibition) with the empirical data for the Dutch and English target items in the right panels.

In sum, the BIA model suggests that the neighborhood density effects in Dutch (inhibitory) and in English (facilitatory) in these PDM experiments can be accounted for by a combination of two manipulations: reducing the L2



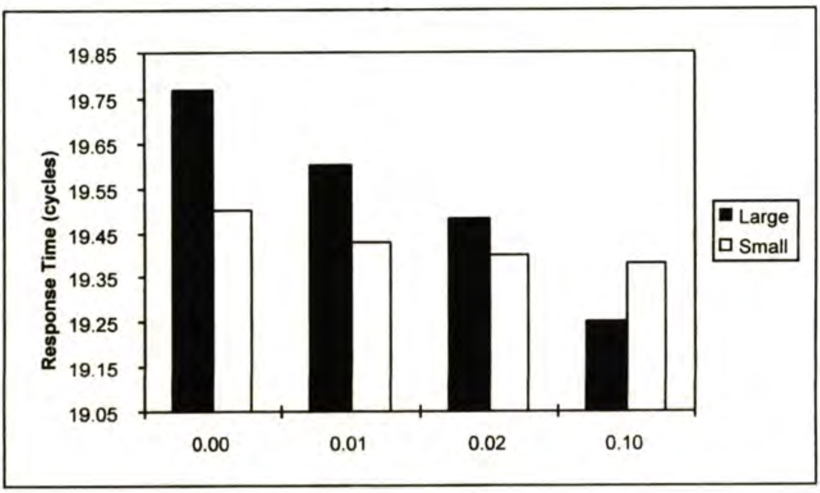




*Figure 5a and b.* Development of effects of English neighbors on the recognition of Dutch target words (a) for high-proficiency subjects over four parts of the blocked progressive demasking experiment and (b) for BIA simulations involving an increasing amount of cross-language inhibition from the Dutch language node to English word nodes. The Dutch words either had a large English neighborhood (many neighbors) or a small one (few neighbors).

frequency range and introducing (experiment dependent) asymmetric top-down inhibition from language node to word level. That combining these two mechanisms indeed resulted in a good fit was evident from a chi-square test on the results of the mixed PDM experiment and the BIA model with limited inhibition and reduced L2-frequencies (Figure 6, right and middle panels). A chi-square was computed on the basis of the eight pairs of empirical RTs and predicted number of cycles for conditions with large and small numbers of neighbors within and between languages with respect to Dutch and English items. For six degrees of freedom (eight minus two free parameters, one for the reduced frequency range and one for the top-down inhibition), the resulting chi-square of 12.15 was not significant ($p>.05$), indicating that the model data did not differ significantly from the empirical data. In the development phase of a cognitive model it may be difficult to obtain non-significant chi-squares, but (as in the case of correlational analyses) a comparison of model variants in terms of the relative size of chi-squares may already be quite informative (for an example, see Baayen, Dijkstra, & Schreuder, 1997).

## Conclusions

In this paper, we have examined how an implemented model of bilingual word recognition, the BIA model, simulates lexical competition effects between orthographically similar words from different languages in two different paradigms (masked orthographic priming and progressive demasking). Especially the empirical results on neighborhood density effects and the corresponding simulation work indicate that cross-language competition arises during the lexical identification process and not during some late decision process. As such, the results not only support the hypothesis of language-nonselective access, but also that of an integrated lexicon for bilingual word recognition.

Furthermore, we have discussed in some detail (1) how the BIA model simulates cross-language effects of masked orthographic priming; (2) how the model can do so for bilinguals differing in their L2-proficiency; (3) how it simulates neighborhood density effects that change in the course of an experiment; and, (4) how one and the same model can capture different patterns of effects for items from two languages presented in a mixed language list. These different simulations were successful due to two special characteristics of the BIA model (apart from the standard IA mechanisms of lateral inhibition between word units and top-down feedback from words to letters): the possibility of introducing a reduced frequency range for L2-words, and the mechanism of top-down cross-language inhibition exerted by the language nodes onto the word level.

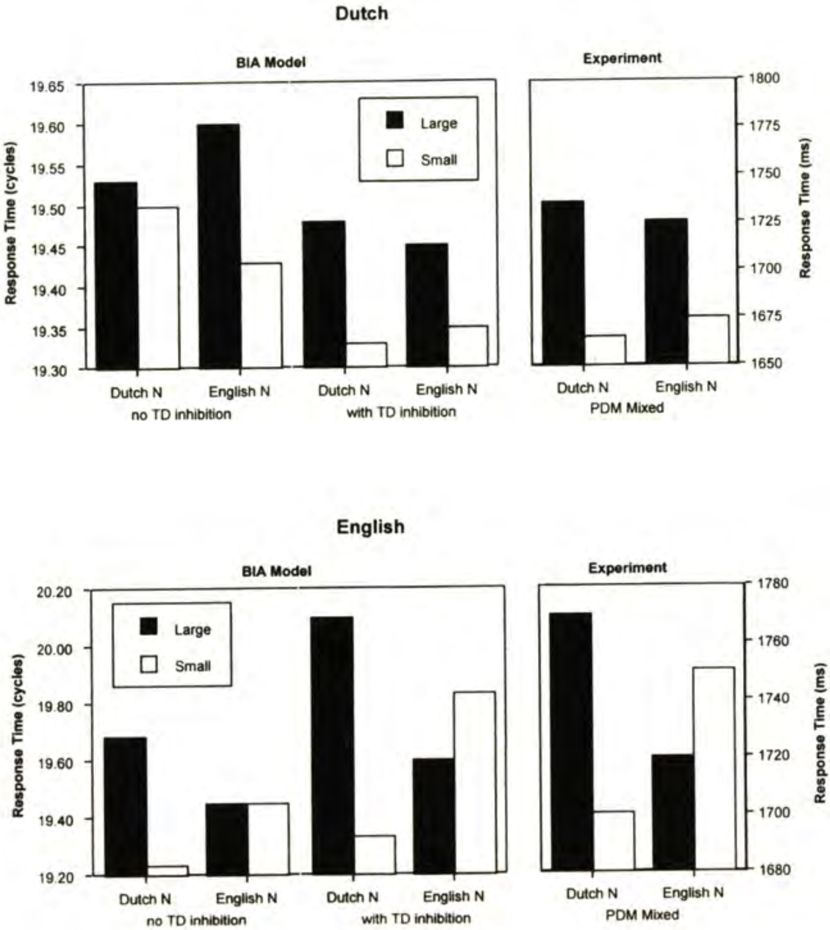



*Figure 6*. Simulations of the Dutch and English data from the mixed progressive demasking experiment with two variants of the BIA model involving a reduced English word frequency range. The first model variant (left) did not assume any asymmetric top-down (TD) inhibition from Dutch language nodes on English words, while there was limited top-down inhibition in the second variant (middle). The empirical data (right) are the same as those presented in Figure 4. English N: number of neighbors in English. Dutch N: number of neighbors in Dutch.

Taken together, these results indicate that the BIA model in its current form already possesses a certain *descriptive adequacy* in that the model's behavior captures relevant aspects of bilingual word recognition. This is true especially in a qualitative respect, while quantitative comparisons between data and model still need to be extended. The model was also shown to display some *horizontal generality* in that it could be used to simulate the results of two different experimental paradigms while all the basic parameters and other model assumptions remained unchanged. Finally, the model displayed considerable *heuristic value* in helping us to interpret some of the empirical data in a new way and generating ideas for new experiments to test the interpretation.

Despite these apparent successes, the model is still in the first stage of its development. Currently, it incorporates only orthographic representations, which implies that the success of any simulation of word recognition processes depends on the extent to which semantic and phonological aspects can be ignored. In this respect, the *horizontal generality* of the present version of the model is certainly limited. Furthermore, it is gradually becoming clear that bilingual word recognition is context-dependent to an extent that requires any processing account to be complemented by a "meta"-account that explains performance changes due to task demands and language context. We are at present developing such a meta-account for the BIA model, which attributes an important role to the second mechanism explored in this paper: top-down cross-language feedback (see Dijkstra, Van Jaarsveld, & Ten Brinke, 1998).

## References

Baayen, H., Dijkstra, A., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37*, 94-117.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition, 25*, 447-457.

Brysbaert, M. (1998). Word recognition in bilinguals: Evidence against the existence of two separate lexicons. *Psychologica Belgica, 38*, 163-175.

Dijkstra, A., & Van Heuven, W.J.B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A.M. Jacobs. (Eds.), *Localist connectionist approaches to human cognition* (189-225). Mahwah, NJ: Erlbaum.

Dijkstra, A., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism, 1(1)*, 51-66.

Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked nonword priming. *Quarterly Journal of Experimental Psychology, 42A*, 353-372.

Ferrand, L., & Grainger, J. (1993). The time-course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31*, 119-122.

Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 11-26). Amsterdam/Philadelphia: John Benjamins.

Grainger, J., & Dijkstra, A. (1992). On the representation and use of language information in bilinguals. In R.J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 207-220). Amsterdam: Elsevier .

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language, 35*, 623-647.

Grainger, J., & Jacobs, A.M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103*, 518-565.

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3*, 233-274.

Grosjean, F. (1997). Processing mixed language: Issues, findings and models. In A. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (Chapter 8, pp. 225-254). Hillsdale, NJ: Erlbaum.

Jacobs, A.M., & Grainger, J. (1994). Models of visual word recognition: Sampling the state of the art. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 1311-1334.

Keatley, C.W. (1992). History of bilingualism research in cognitive psychology. In R.J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 15-49). Amsterdam: Elsevier.

Li, P. (1996). Spoken word recognition of code-switched words by Chinese-English bilinguals. *Journal of Memory and Language, 35*, 757-774.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception, Part 1: An account of basic findings. *Psychological Review, 88*, 375-405.

McClelland, J.L., & Rumelhart, D.E. (1988). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol 3, pp. 203-239). Cambridge: Bradford Books.

Smith, M.C. (1997). How do bilinguals access lexical information? In A. de Groot & J. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic perspectives* (Chapter 5, pp. 145-168). Hillsdale, NJ: Erlbaum.

Van Heuven, W.J.B., Dijkstra, A. & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458-483.

Manuscript received May 1998