Practical Modelling of Bilingual Word Recognition: Bilingual Interactive Activation vs Interactive Activation

Miriam Gorra (gorram@carleton.edu)
300 N College St
Northfield, MN 55057 USA

William Maxon Bremer (bremerw@carleton.edu)
300 N College St
Northfield, MN 55057 USA

Abstract

Word recognition is a central component of human language cognition, and more specifically of cognitive activity during the fundamental human task of reading. We implement here two spreading activation models of word recognition, the Bilingual Interactive Activation (BIA) model and the Interactive Activation (IA) model on which it is based. We test them extensively on both isolated word cases and sequential sets of input, using a combined lexicon of Dutch and English words. In a comparison of the two models and a human study of word recognition response time in bilingual Dutch-English speakers (Dijkstra & Van Heuven, 1998), we find that the models do return moderately accurate results in the isolated word cases, but do not necessarily produce useful data in the sequential case.

Keywords: Word recognition; Bilingual Interactive Activation.

Overview

Introduction

Word recognition is so fundamental a part of our internal language processing that we almost never stop to recognize it, despite its incredible functionality. We process texts accessed visually at an average rate of 300 words per minute (Cook & Bassetti, 2005) or 5 words per second, and the central function in that process is visual word recognition. During the process of reading we perform not only the visual processing of line segments to letters and letters to words, but also the grammatical and phonological processes necessary to contrive sense from the passage (Cook & Bassetti, 2005)

Bilingual word processing and recognition is interesting in particular due to some of the questions it raises. The primary two are as follows: first, when word candidates similar to the input string are activated, are candidates from both languages considered, or from only one? Secondly, is the internal lexical "database" stored as one single integrated lexicon or in two distinctly separate ones? These two possibilities, that of distinctly separate linguistic processing and competition and that of more or less integrated competition, are generally referred to as language selective and language non-selective access hypotheses respectively \cite{DijkstraVanHeuven1998a}. All models and data referred to in this work operate under the assumption of language non-selective access.

In this work, we utilize an implementation of the Bilingual Interactive Activation (BIA) model as well as the more monolingual Interactive Activation (IA) model with varying sets of bilingual input words. The input words were chosen based on a reference set of human data in which subjects were presented with those same words. Our languages of choice are Dutch and English, as they align with the human data. In a final test, we present words to the model sequentially without resetting activation values, in order to assess the model's viability as a representation of word recognition in context. In terms of the isolated tests, the model performed well: the BIA model functioned as a more accurate representation of the natively Dutch bilingual subjects of the human study than the IA model, both becoming certain more quickly in the case of the native language and slightly more slowly in the case of the non-native language. On the other hand, the sequential test revealed little useful data, though that in itself lends us insight on the nature and usefulness of the BIA model.

Background

The Bilingual Interactive Activation (BIA) model was presented in 1998 (Dijkstra, Van Heuven and Grainger) as a language non-selective access model for bilingual word recognition concerned with processing from the visual feature level to the letter level. It was based on the Interactive Access (IA) model constructed in 1981 (McClelland and Rumelhart, 1981). BIA was reasonably successful at the time, and Dijkstra and Van Heuven would later go on to write a modified version called BIA+ in their 2002 work "The architecture of the bilingual word recognition system: From identification to decision."

The implementation became a staple of modelling bilingual word recognition, and although many other developments would occur in the field it would remain a reference point for many. The 2005 article "Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service?" by Brysbaert and Duyc mainly concerned itself with errors in the then-prevailing RHM, but cited BIA as an example of building a successful bilingual model on the foundation of an existing monolingual model.

The PDP handbook, a further work by McClelland, (2015) actually heavily based its model implementation portion around the IA model, a fact that has both been a great aid to this work and further shows the standard set by this model and its successor. The BIA model repeatedly demonstrates its own viability and modern relevance, and is a worthy subject of closer examination.

Model

Interactive Activation

We'll start with a description of the IA model, as it serves as the necessary foundation for the BIA model. The IA model is a specialized spreading activation network designed to represent visual word recognition. Note that each individual model only processes words of a preset length: It utilizes additive ("excitatory") and subtractive ("inhibitory") connections It features three separate levels of units: the feature level, the letter level, and the word level. Each level's behaviors are fairly self-explanatory in terms of reasoning, and are elaborated here for the purpose of clarity.

The feature level can for all intents and purposes also be treated as an input layer. It takes in 4 sets of 14 binary inputs, each of which represents the presence of a certain line segment in that letter position. Fig 1 depicts the position and appearance of the 14 segments detected by the feature level. As an example of one letter positions input set, the letter "A" would be represented as 11111010100000. The processing of visual input is assumed, and the model is directly passed the binary lists of present features. Outgoing connections from the feature level are fairly straightforward: each feature unit excites any letter units in the same letter position that possess that feature, and inhibits any letter units in the same letter position that do not possess the feature.

The letter level behaves similarly in some ways: each letter unit has an excitatory connection to words that have that same letter in the same position, and an inhibitory connection to words that do not. However, each letter unit also possesses an inhibitory connection to each other letter in the same spatial position as itself.

Finally, the word level possesses a feedback connection back to the letter level, such that each word unit excites and inhibits letter units in the same position as they appear in the word itself. Similar to the letter level, each word unit possesses an inhibitory connection with all other words. The units that appear in the word level are determined by an imported lexicon file, generally an extensive list of words of the required length.

All excitatory and inhibitory connections are additive in relation to a given units activation level, scaled based on by the *alpha* and *gamma* input parameters to the model respectively. Note that there are not two single universal *alpha* and *gamma* parameters, but rather a separate *alpha* parameter exists for each inter-level excitatory connection and a likewise a *gamma* parameter exists for each inhibitory connection.

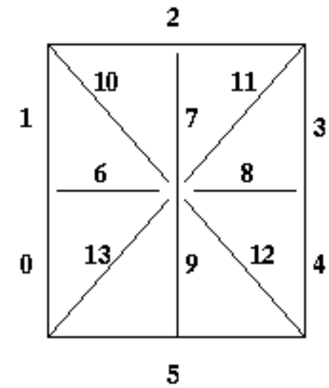


Figure 1: Diagram depicting the 14 letter features the IA and BIA models take as input as a binary list on the feature level.

With each cycle, each unit's activation level is updated based on the level of the incoming excitatory and inhibitory connection. If any given words certainty level, the calculation of which is given in the implementation section, rises above a certain point, then the model determines that word to be its result. The "correctness" of the truly correct word is self-reinforcing: the word excites the activation of the letters contained within, and the letters reinforce the word in question. Using a formula described in the Implementation section, the model takes the running average of the activation level of any given word each cycle and uses it to determine the certainty of the model and the probability of correctness.

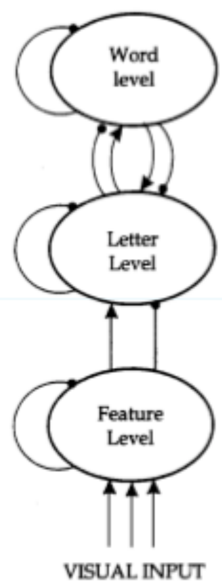


Figure 2: A general visual representation of the IA model. Circle-headed arrows represent inhibitory connections, while triangularly-headed arrows represent excitatory connections.

Bilingual Interactive Activation

The BIA model is fundamentally similar to the IA model, possessing each of the levels and features that the latter does. However, as depicted in Fig. 3, the BIA incorporates a new level into its design: the language level.

The language level possesses only two units, one for each language tested for. Hypothetically, there could be an arbitrary number of nodes incorporated for an arbitrary number of languages, but given a specific interest in bilingualism we will discard the possibility. Each word in the lexicon from a given language possesses an excitatory connection to the corresponding language node, and each language node serves only to inhibit each word in the other language. The inhibitory connections to the word level can differ depending on the language node: for example, in the Van Heuven dataset the bilingual speakers in question are native to Dutch, meaning that it is much more strongly inhibitive to English than vice versa.

Although words are contained in one bilingual lexicon rather than two separate ones, words are each "aware" of the language they belong to via said excitatory connection to their given language node, which can then slightly inhibit words of the opposite language. Thus, rather than the model selecting a language in a binary manner at any point, words in a given language simply reinforce the likelihood that the correct word is not in the opposite language.

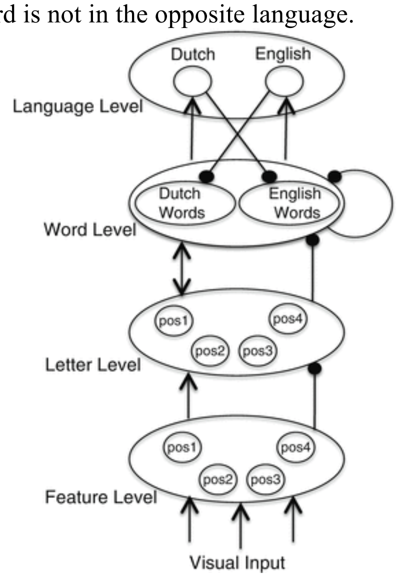


Figure 3: A simple visual representation of the BIA model (Dijkstra & Van Heuven, 2002)

Implementation

Here we include the implementation details for the sake of maximal clarity. The specific values of the model's various parameters are key to its accuracy and to proper understanding of this specific set of trials.

The following parameter value choices are based on recommendations from "The PDP Handbook" (McClelland, 2015), with the distinct exception of values unique to the BIA model. Specifically, those are the word to language alpha, the language 1 to word gamma and the language 2 to word gamma. Those were obtained from the Van Heuven et al. study (1998).

Excitatory (alpha) values:	
Feature to letter	0.005
Letter to word	0.07
Word to letter	0.3
Word to language	0.3

Inhibitory (gamma) values:	
Feature to letter	0.15
Letter to word	0.04
Word to word	0.21
Letter to letter	0.00
English to word	0.00
Dutch to word	0.03

There are a couple of interesting value choices here worth discussing. First of all, in accordance with the PDP handbook implementation (McClelland, 2015) the IA-relevant alpha values scale up by a power of ten with each level ascended, and no letter to letter inhibition takes place. This balances the natural downwards scaling of activation values as we ascend the levels in the first case, and allows for proper letter scaling in the second (McClelland, 2015).

The non-IA relevant values (Van Heuven et al., 1998) are chosen for a specific reason pertinent to the study at hand. Because all subjects of the study from which we receive our human data were native Dutch speakers, we attempt to imitate the way in which Dutch would be dominant in their word recognition by giving a small value to the Dutch to word inhibition, and no value to the English to word inhibition.

In order to determine the likelihood that a given unit is the correct response, we require the following successive formulas (All of which are referenced from McClelland, 2015).

The response strength of each unit is given by:

$$s_i(t) = e^{k\overline{a_i}(t)}$$

Wherein k is the scalar parameter *oscale* and $\overline{a_i}(t)$ is the running average of the activation of unit *i* at time *t*, which is given by:

$$\overline{a_i}(t) = (orate)a_i(t) + (1 - orate)\overline{a_i}(t\text{-}1)$$

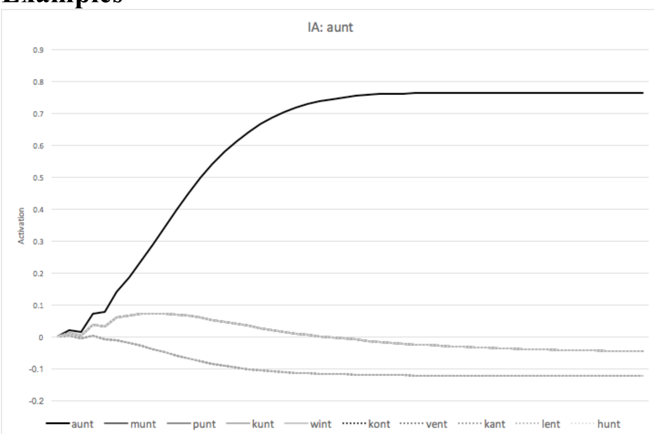
The probability of choosing this particular unit as the response is:

$$p(r_{i,t}) = \frac{s_i(t)}{\sum_{j \in C} s_j(t)}$$

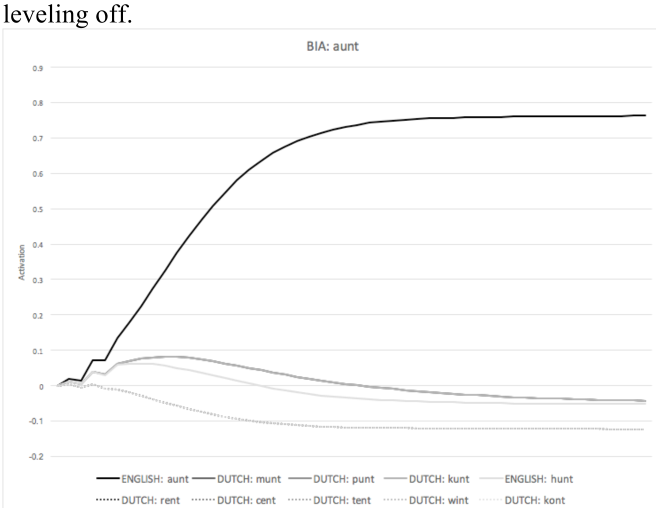
It is worth noting that there are certain features of the IA and/or BIA that we chose to leave out. For example, the original IA model features an "absence unit" for each feature in the feature level (McClelland & Rumelhart, 1981) This is meant to allow for incomplete or damaged inputs to be processed by the model, as the missing features could be filled in as neither present nor absent. We chose not to include these because ambiguity of input was irrelevant to the tests we wished to perform, and thus the feature was not needed.

In addition, the original model included a parameter and encoding for word frequency inputs, and the weighting of words based on these word frequency inputs (McClelland & Rumelhart, 1981). The reason we did not include this feature was twofold: first, the tests we were conducting were on a small enough scale that weighting of word frequency would've been essentially negligible. Secondly, we did not have access to the same frequency datasets used in the human tests, and thus had no applicable frequency data for the context.

Examples



Here is an example of activations over time of various word units in the IA model when given the input "aunt". You can see how certain words begin to rise immediately when excited by the matching letters, only to fall once inhibited unmatching by the correct word as its activation rises. The correct word, which matches all the input letters, rises quickly due to its lack of activation, before eventually leveling off.



In the BIA version, activations follow a similar pattern, with one exception. The english word "hunt" has a lower activation than in the IA model from almost the beginning, falling below the Dutch words it matched in the IA model, before rejoining them at the end. This is due to that initial rise in activations for several Dutch words having an inhibitory effect on "hunt", the only non-correct English word in this sample. This inhibitory effect fades once the Dutch words stop being activated and the language node decays, leading to the activations rejoining each other near the end.

Experiments

An advantage of the BIA and IA models is that they are very versatile and give a variety of types of outputs, which allows for greater analysis. We implemented several different output methods, and ran different experiments to demonstrate each.

Experiment #1

Our primary experiment compared response times for our two models and human data taken from Walter J. B. van Heuven and Ton Dijkstra's 1998 paper "Orthographic Neighborhood Effects in Bilingual Word Recognition". In it they tested a group of bilingual Dutch-English speakers, with Dutch as their native language, on word recognition in both Dutch and English. The experiment presented subjects with a word in either Dutch or English, covered by a mask of visual noise, which gradually disappeared. Subjects were asked to respond with the word as soon as they could discern it, and the response times measured.

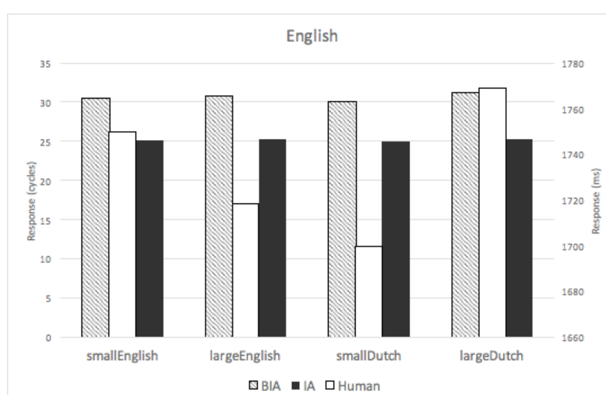
The words presented were grouped according to the number of neighbors (words that closely resemble them) they had in both English and Dutch, and the response times averaged for each of these four categories for both languages.

Method

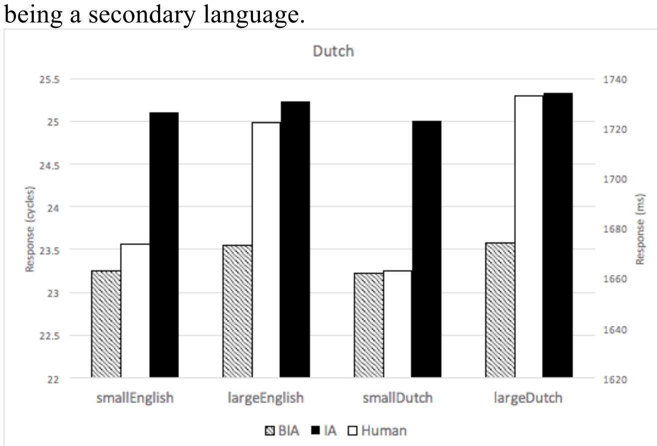
Van Heuven and Dijkstra also implemented a BIA model to compare to their human data, and gave suggested parameters of having the Dutch language suppress English words by 0.03, and giving English words a lower resting activation of -0.3. This was done in order to mimic the fact that English was a secondary language for the subjects.

Our BIA model used those suggested parameters, as well as the same input words used in the human experiment. We ran cycles of the model until it gave a 90% probability to the correct word, then averaged the number of cycles for each word in the category. We did the same for out IA model, using a combined lexicon of both English and Dutch words.

Results



For the English test words, we saw much less variability in our results than in the human data. This is potentially just a function of the difference between measuring in milliseconds and measuring in cycles. The human data shows that English words with either a large number of English neighbors, or a small number of Dutch neighbors, are easier to identify. Our models instead both show small numbers of neighbors in either language as slightly faster. Also, the IA model is consistently faster than the BIA, which makes sense considering that the BIA model is suppressing the likelihood of English words to account for it being a secondary language.



We ran the same tests, but with Dutch input words, where we saw more variability. Here the human data had smaller numbers of neighbors in either language as faster, which is the same pattern we modeled for both languages. Both of our models also captured that the number of Dutch neighbors had a larger impact than the number of English neighbors. For this input, the IA model was consistently slower than the BIA, which makes sense given that the response time was calculated based off the number of cycles needed for the correct word's response strength to be a significant fraction of the overall response strength, and the BIA model lowered the English response strength through lower resting levels and language-level suppression, thus making Dutch words more likely.

Experiment #2

Under normal circumstances, the IA and BIA models are reset in between words, always starting from baseline resting activations. However, when people are shown words, they generally arrive in sequence, which may affect how quickly people understand them. To mimic this, we ran un-reset versions of our models on various inputs.

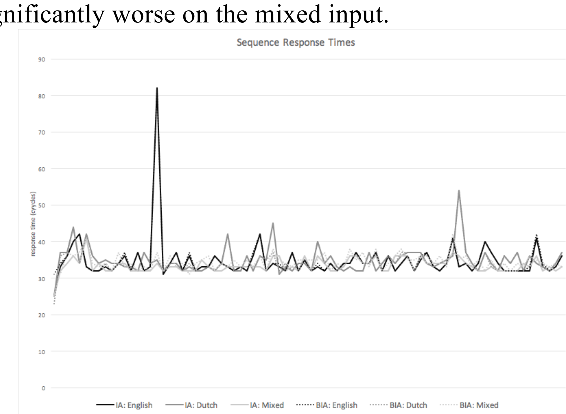
Method

We assembled three different data sets, all of equal size, one consisting of Dutch words, one of English, and one of mixed. The alternating words were chosen by alternating a selection from the Dutch set and the English set, in order to not bias due to alphabetical or neighbor orderings in the original data sets. These sequences were run through both an IA and BIA model with parameters as discussed in Experiment 1, running each word until it reached 90% probability. The difference from the previous experiment is that we didn't clear the network activations between words, and recorded the cycle number for each input. This made the network have to adapt to a new word from a network already biased towards certain letters, words, and potentially language.

Results

	IA	BIA
English	34.6375	34.85
Dutch	34.7375	34.6375
Mixed	33.4	34.475

In all our trials, the models took considerably more cycles than seen in earlier networks which were cleared between words. We also saw the BIA model doing slightly worse than IA on English and slightly better on Dutch, as well as significantly worse on the mixed input.



When looking at the response times throughout the input, we can see that in general the IA and BIA models follow each other exactly, or with the IA model slightly faster. In general, all of the trials remain steady at around 35 cycles, but there are occasional exceptions, most notably a spike up to above 80 in the English inputs.

Discussion

Our results suggest that not clearing the network increases response time by around ten cycles (when compared to our results from Experiment 1). However, this increase is much less in the English BIA, which was the slowest of the cleared networks, but here has increased the least, putting it in the middle of the pack. However, the relative ordering between the four single language trials remains the same, with English BIA the worst, followed by Dutch IA and then the other two. This suggests that clearing the network has little effect on the model's similarity to human data, just its overall runtime. This has good implications for the use of the IA and BIA models, since they tend to be used with clearing, while intuition would suggest that humans do not.

Of particular interest in this experiment is the use of a mixed input set of alternating languages. This forces the BIA model to switch between languages, while the IA model does not have to. Intuitively enough, the BIA model is around a cycle slower than the IA model on average, implying that adjusting to a new language lengthens response time. This delay when code-switching is a major difference between the IA and BIA model, and an argument for using the BIA model when dealing with bilingual input.

There are several noticeable spikes in the data, in particular one in the English input. An examination of the data reveals that that particular input was “nude”, which had been immediately preceded by “noon”. This could potentially be due to the similarities in the words, as the input “nude” would continue to excite that initial N, as well as many features in a second U and third D. It is possible that these similarities kept exciting “noon” even as it begun to be inhibited by the final E and new features, and thus increased the amount of cycles until “nude” had a 90% probability, since the response time is based on correct response strength divided by overall response strength.

General Discussion

In general, our results were not very close to the human data. There are many factors that could explain this: flawed parameters, skewed lexicons, un-implemented supplemental features, or simply a lack of broader human data. Our models were heavily reliant on corpuses whose most accurate and useful versions required payment, and therefore we were unable to get full lexicons or frequency data. Additionally, the constraints of the IA model, which primarily models response time, as well as our bilingual focus, made it difficult to find experimental data to base our model off of, leading to our only human data being taken from one relatively small study of one language combination. The BIA model has also been extended more recently into a BIA+ model, which we chose not to implement since it uses phonetic and semantic data we don’t have, but does indicate a general consensus that the original BIA model is flawed.

This being said, we can still compare our two models against each other. In general, the IA model ran faster, which we feel indicates that it does a poorer job of modeling human computation than the BIA. The human data we did see, as well as our personal experience with bilingual language, has indicated that humans do have slowed response times when, for example, they are forced to switch suddenly between languages, even when fluent. This result is a strong indication that the BIA model might also be correct in its assumption that bilingual humans draw upon both lexicons (with varying amounts of suppression) even when only using one language. Even though we were unable to match human data in Experiment #1 enough to firmly state a conclusion, our other experimental results support the BIA model over the IA model when modeling bilingual word recognition.

References

Brysbaert, Marc. Is it time to leave behind the Revised Hierarchical Model of bilingual language processing after fifteen years of service? Bilingualism: Language and Cognition 13 (3), 2010, 359–371, Cambridge University Press 2010 doi:10.1017/S1366728909990344

Dijkstra, Ton and Walter Van Heuven J.B. (2002). "The Architecture of the Bilingual Word Recognition System: From Identification to Decision." Bilingualism 5, no. 3 Cambridge Press.

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), Scientific psychology series. Localist connectionist approaches to human cognition (pp. 189-225). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Keuleers, E., Brysbaert, M. & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. Behavior Research Methods. Retrieved from http://crr.ugent.be/programs-data/subtitle-frequencies/subtlex-nl

Leech G, Rayson P, Wilson, A. (2001) Word Frequencies in Written and Spoken English: based on the British National Corpus. Retrieved from http://ucrel.lancs.ac.uk/bncfreq/lists/1_2_all_freq.txt

LI, P. (2013). Computational modeling of bilingualism: How can models tell us more about the bilingual mind? Bilingualism: Language and Cognition, 16(2), 241-245. doi:10.1017/S1366728913000059

McClelland, J. L., & Rumelhart, D. E. (1988). Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises. Cambridge, MA: MIT Press. Chapters 2 and 7.

McLelland, J E, Rumelhart, D E. (1981). An Interactive Activation Model of Context Effects in Letter Perception: Part 1. An Account of Basic Findings. Psychological Review, Volume 88 (Number 5). Pages 376-407.

Second Language Writing Systems, edited by Prof. Vivian Cook, and Benedetta Bassetti, Multilingual Matters, 2005.

Van Heuven WJB, Schriefers H, Dijkstra T, Hagoort P. Language Conflict in the Bilingual Brain. Cerebral Cortex (New York, NY). 2008;18(11):2706-2716. doi:10.1093/cercor/bhn030.

Walter J.B. van Heuven, Ton Dijkstra, Jonathan Grainger. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. Journal of Memory and Language, Volume 39 (Issue 3). Pages 458-483.