

BACHELOR THESIS

ARTIFICIAL INTELLIGENCE

Radboud University



**‘Do you like pizza?’
How Multilink can be applied to Italian-English word
recognition**

Author:

Axel van Buren
S1037864

First supervisor:

Dhr. Prof. Dr. A. F. J. Dijkstra
Donders Institute for Brain,
Cognition and Behaviour
ton.dijkstra@donders.ru.nl

Second supervisor:

Dhr. A. G. A. J. van Geffen
Research Services
aaron.vangeffen@ru.nl

Second reader:

Dr. F.T.M Léoné
Donders Institute for Brain,
Cognition and Behaviour
f.leone@donders.ru.nl



Abstract

Theories of word recognition have become so specific that they allow the creation of computational models that are capable of predicting and generating results similar to empirical data. Multilink is such a recently developed successful model of word processing. It is a localist-connectionist model that can account for monolingual and bilingual word recognition in different task situations. The model has been proven to be quite precise when simulating English and Dutch research results. However, until now, for Italian the word recognition performance of monolingual and Italian-English bilinguals has not yet been simulated. In this paper, we will evaluate the quality of Multilink's simulations on the word recognition performance of Italian-English participants in a recent study (Cangelosi & Dijkstra, in preparation). To this end, the available English-Dutch lexicon is extended with Italian. It is shown that Multilink simulations with this new lexicon mimick the empirical data from an Italian-English lexical decision task to a promising extent.

Table of contents

Chapter 1: Introduction.....	3
Prior research	6
Cross-language similarity.....	6
Cognate facilitative effect and varying orthographic overlap.....	7
Cognate varying cross-linguistic overlap	8
Chapter 2: Multilink.....	9
The integrated lexicon.....	9
Assumptions of the Multilink model	9
The model's architecture	10
Resting level activation.....	12
The activation function.....	13
The current parameter settings	14
Chapter 3: Simulating the Italian lexicon	16
Justifying the correlations	17
Chapter 4: Simulating the Italian empirical data.....	18
Simulating the lexical decision task.....	18
Considering cognate pairs varying in Levenshtein distance.....	19
Simulations with a generalized lexical decision task.....	20
Identical cognates	21
Non-identical cognates.....	22
Control words.....	24
Conclusion	25
Chapter 5: General discussion.....	26
Appendix.....	28
A: Multilink parameters.....	28
B: Stimulus composition list experiment 1.....	29
References.....	30

Chapter 1: Introduction

An amazing skill of language users is that they can store and use several different languages almost simultaneously. When we consider the performance of such ‘bilinguals’ or ‘multilinguals’, a number of interesting questions pop up. How, for instance, do you handle words that are similar in a foreign language and your own? Take the Italian word “talento”. This word means exactly what you think in English. In this case, not only the form of the word, but also its meaning is similar across Italian and English. Such a word is called a ‘cognate’. Another Italian word is “magazinno”. An English speaking person would probably think that this Italian word translates into the English word “magazine”, which is a collection of articles on paper. However, this would be wrong, because the correct translation would be “warehouse” (in Dutch it would be “magazijn”). Words like these, that sort-of betray you when you encounter them, are called false friends.

The existence of words that overlap in form only (false friends) or in form and meaning (cognates) raises the interesting question of how we relate the words of different languages during processing. In this study, we consider how Italian-English bilinguals process words in their first language, Italian, and how they translate words from their second language, English, back into Italian.

One way to translate the Italian word ‘ragazza’ would be to use its orthography (spelling) to directly activate its translation ‘girl’. This is called ‘word association’ (Kroll & Stewart, 1994). The second way to translate ‘ragazza’ would be to look up its meaning after which the associated word form ‘girl’ is retrieved from the mental lexicon (Potter, So, Von Eckardt, & Feldman, 1984). This is called ‘concept mediation’.

The revised hierarchical model (RHM) by Kroll and Stewart (1994) combined both ways to explain the pattern of translation latencies in bilinguals from their first (L1) to second (L2) and second to first language. However, in 2009 Brysbaert and Duyck argued that this verbal model was no longer providing a useful description of how bilinguals process words. In their view, the implemented localist-connectionist Bilingual Interactive Activation Plus (BIA+) model provided better explanations for available data for word recognition, where BIA+ expands on the previously established BIA model, depicted in figure 1.

The Bilingual Interactive Activation model (Dijkstra & van Heuven, 1998) is a clear example of how an existing computational model of monolingual processing can be adapted to account for language processing from two languages. It took its base from the Interactive Activation model created by McClelland and Rumelhart (1981), but added nodes to the word level for L2 input. It also contained a fourth layer that had two language nodes.

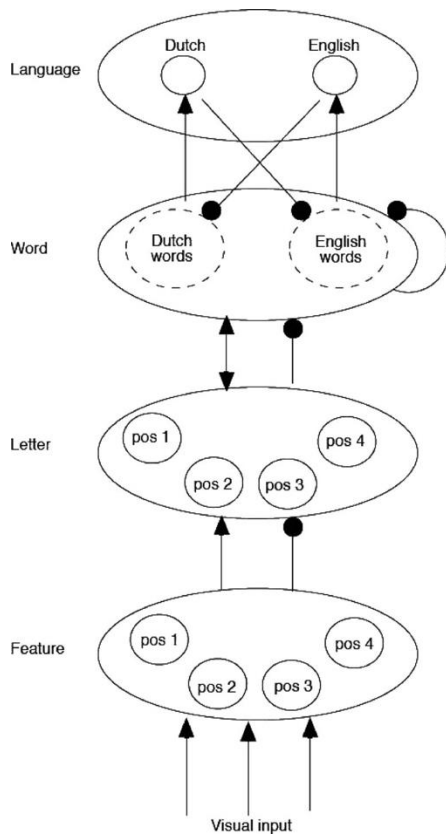


Figure 1. The Bilingual Interactive Activation Model (BIA), shows how a computational model of monolingual language processing can be adapted to include input from two languages. Note. Arrows indicate excitatory connections, whereas circles indicate inhibitory connections. This figure was taken from “The architecture of the bilingual word recognition system: From identification to decision.,” by Dijkstra & Van Heuven (2002).

The BIA+ model incorporates the BIA model as a special case, but adapts the function of the language nodes. It also incorporated a task/decision system in line with the ideas by Green’s (1998) on task schemas and task control, as graphically displayed in figure 2. Just as in the BIA model, the bilingual lexicon is integrated across languages and is accessed in a language non-selective way. The BIA+ assumes that bilingual word recognition is not only affected by cross-linguistic orthographic similarity effects, but also by cross-linguistic phonological and semantic overlap. To implement this possibility, the BIA+ model incorporates a lexicon specifying phonological and semantic representations besides orthographic ones (Dijkstra & Van Heuven, 2002).

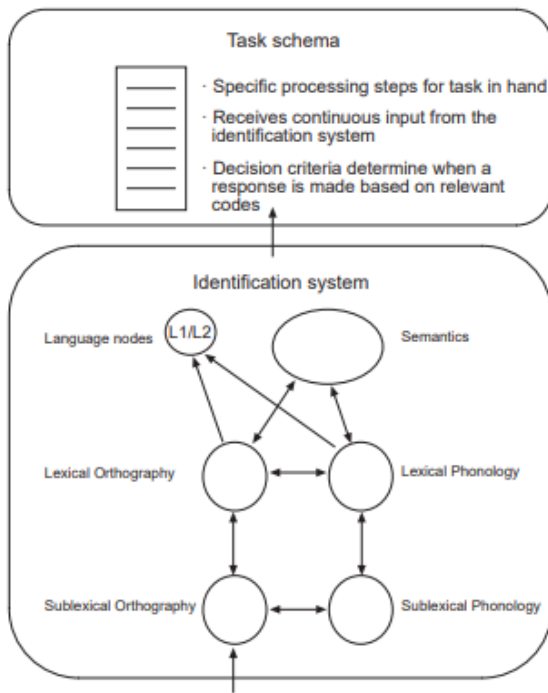


Figure 2. The Bilingual Interactive Activation+ Model (BIA+), incorporating the task schema. Note. This figure was taken from “The architecture of the bilingual word recognition system: From identification to decision.,” by Dijkstra & Van Heuven (2002).

The BIA+ notion of interconnected units that are activated in different ways for different item types can be related to the theoretical view of the brain in terms of processing networks of nervous cells. As these nervous cells networks are built from neurons which collect signals from others through their dendrites, gaining activation/information from several connected sources and passing this onto other neurons, which is the typical behavior of interconnected units.

Another computational model for bilingual word retrieval that implements this view is the localist-connectionist model Multilink (Dijkstra & Rekké, 2010). This model simulates both monolingual and bilingual word recognition, and word translation. In particular, the model accounts for the performance of bilinguals in lexical decision and various other tasks. The focus for the present study will lie on the simulation of human word recognition performance in the lexical decision task. In this task, the participant has to decide whether or not a presented letter string is an existing word or not by pressing either a ‘correct’ or an ‘incorrect’ button.

Multilink tries to reduce the shortcomings of RHM by extending it using BIA+ (Dijkstra & Rekké, 2010). Multilink shares the basic assumptions of BIA+ about Interactive Activation, the distinction between orthographic, semantic, phonological, and language membership representations, and lastly the presence of a task-decision system. It also assumes that L1 and L2 lexicons may differ in size and that the links between word form and meaning representations may be different for L1 and L2, as assumed by the RHM model. Multilink has elaborated on the assumptions behind these models, and also incorporates an integrated lexicon containing semantic and phonological representations. The lexicons containing this data are based on the Dutch and English lexicon projects, which are large databases containing lexical decision data for more than 140000 words from lexical decision tasks performed by two different groups of British participants (DLP and ELP; Keuleers, Diependaele, & Brysbaert, 2010). For all words, orthographic, phonological, and semantic representations are

present in the lexicon. Using those representations, simulations can be done to estimate the capabilities of Multilink when simulating lexical decision tasks. The Dutch and English lexicon projects provide empirical reaction time data that were used to evaluate the performance of Multilink. Later in this study, we will do the same for Italian-English.

By means of simulating empirical reaction time data, several interesting aspects of bilingualism can be studied. Those aspects are highlighted when the model is confronted with different types of words, such as cognates, neighbors, and false friends (Vanlangendonck, 2012). The first item type of interest are cognates. In this study, cognates are considered words that are form similar translation equivalents. An example of a cognate is the word 'pizza'. The second type of interesting words are interlingual neighbors, which are words from two different languages that differ slightly by orthography. The English word 'number' is a great example of this as the Italian equivalent is 'numero', varying slightly orthographically. A final type of words to be simulated are false friends, words that are orthographically similar but semantically different. An example we used earlier was the Italian word "magazinno". Using lexicons containing many types of such stimuli, we can test the quality of Multilink simulations for empirical data gathered from a lexical decision task performed by Italian participants.

Prior research

To set the stage for the research reported later in this thesis, we will now discuss relevant literature starting with the research that laid the ground rules for Multilink. As previously mentioned, the localist-connectionist model Multilink was created on the basis of RHM and BIA+. After Brysbaert and Duyck (2009) argued for the RHM to be abandoned in favor of the BIA+ model, Dijkstra and Rekké evaluated the RHM and BIA to investigate if an implemented model of word translation could be built by combining their assumptions. This new model, Multilink, was already capable of performing successful simulations on cognate recognition and word translation in both first or second language (Dijkstra & Rekké, 2010).

Compared to the current lexicon of Multilink, the lexicon used in the pilot studies was far smaller, consisting of only 892 English-Dutch word pairs. This poses the question if the amount of used data was sufficient to ensure valid simulation results. The theoretical argument is that if empirical results are in line with model simulations and go beyond the results of a verbal model (such as RHM), this can be seen as validation proof. To ensure a sufficient basis for validation here, we will first simulate performance of a larger lexicon with a total of 1459 words.

Cross-language similarity

Dijkstra et al. (2010) investigated how the cross-linguistic similarity of translation equivalents affects the bilingual word recognition process. Reaction times of participants were gathered while they were performing one of three tasks. Dutch-English bilinguals processed cognates with varying degrees of form overlap. The results showed task-dependent patterns that provided evidence in favor of a localist connectionist account (Dijkstra et al., 2010).

The stimuli in this study consisted of a list of words on a continuum ranging from completely identical cognates to varying form or even completely different form words. The first experiment involved an English lexical decision task, in which the words were translation equivalents varying cross-linguistic similarity. Responses became faster when the similarity increased, even indicating a relatively large cognate facilitation effect for identical cognates. The second experiment was a bilingual, Dutch-English language decision task, based on a subgroup of 157 English and 157 Dutch words. The results

showed that the effect of orthographic similarity reversed, based on which the experimenters concluded that orthographic similarity had inhibitory effects on the reaction times of the participants. The final experiment was an English progressive demasking task, in which the participant is gradually exposed to a word or nonword, arising from behind a mask. The results of this experiment are in strong contrast with the reaction times patterns of the previous two experiments, because they seem to not be dependent on orthographic semantic similarity ratings. The results showed no orthographic similarity effect for non-identical cognates, but a semantic similarity effect arose (Dijkstra et al., 2010).

The results of these experiments were interpreted in the localist-connectionist account of bilingual word recognition. First off, by assuming form overlap between two readings of the cognates leads to a frequency-dependent parallel activation of these readings. Secondly, the readings representations activate a shared semantic representation and separate language membership nodes. When task-specific processing assumptions are also added the findings of the final experiment can be explained within the same orthographic-semantic network. The interpretation must take into account a strong effect of task specific demands, as language decision, lexical decision and progressive demasking are all done uniquely. Originally, the BIA+ model formed the theoretical framework for interpretation, but later simulations with Multilink were also reported (Dijkstra et al., 2009).

The processing of identical and non-identical cognate words has also been explored in Experiment 1 of Dijkstra et al. (2010). Unbalanced Dutch-English bilingual performed an English lexical decision experiment involving English words with varying Levenshtein distances to their Dutch translations. Having used the implementation of a cognate representation based on the 2010 study good fits between the model and empirical data was gained (Dijkstra 2018). This led to the conclusion that when it comes to word recognition and production, there are multiple factors that interact in a complex fashion. However Multilink forms a promising basis for further development of such computational models for word recognition or production.

A limitation of these empirical data and simulations is that they were only concerned with Dutch-English bilinguals. Further validation of lexical decision results on different items types, in particular cognates, is necessary for other language combinations. This will be done for Italian later in this study.

Cognate facilitative effect and varying orthographic overlap

Other empirical data focussing on cognates are available, yet not for Italian. However, the available data still provide valid theoretical insights and outcomes for bilingual word recognition. While the previous study already discussed the variation of cognates, another study, performed by Comesaña et al. in 2015, considered the facilitative effect of cognate words vanishing when reducing the orthographic overlap. The aim of the study was to further explore the issue of identical cognates and non-identical cognates, by manipulating for the 1st time cross-linguistic similarities of identical and non-identical cognate words as well as stimulus list composition.

By having Catalan-Spanish bilinguals perform the lexical decision task in Spanish, Comesana et al. found modulations in cognate processing as a function of their degree of orthographic and phonological overlap, confirming the prior findings mentioned before, regarding the processing of cognates taking into account cross linguistic similarities (Comesaña et al., 2015). The results show that non-identical cognate processing varies as function of orthographic and phonological overlap, as well as a function of semantics.

Cognate varying cross-linguistic overlap

However, while the study of Dijkstra et al. 2010 focused on English-Dutch bilingualism, a recent Italian-English study by Cangelosi and Dijkstra (in preparation) examined the effect of different degrees of cross-linguistic overlap by determining item differences not in terms of ratings, but by varying Levenshtein distances. In addition, this study involved Italian-English bilinguals performing an Italian-English lexical decision task. Analogous to the earlier study, the experiment considered the effects for English-Italian cognates that differ in cross-linguistic overlap. Their English (L2) lexical decision task, with cognates varying in cross-linguistic similarity, consisted of 50% English words and 50% English nonwords. The prediction for the experiment is more cognate facilitation for cross-linguistically more similar cognates (in terms of Levenshtein distance).

The results of the study by Cangelosi and Dijkstra will be compared to the simulation results of Multilink for the same stimulus materials. This comparison will be the main focus of chapter 4.

To set the stage for the simulations with Multilink, we will discuss the underlying ideas behind Multilink in Chapter 2. Next, in Chapter 3, a total simulation of the Italian-English lexicon is done and analyzed. Finally, the above mentioned study on English-Italian cognates that differ in cross-linguistic similarity.

Chapter 2: Multilink

Multilink has been implemented as a computational cognitive model for bilingual word recognition and translation. It can be defined as a precise, operationalized and quantitative representation of reality in a restricted domain of human information processing, in particular, monolingual and bilingual word retrieval. The model simulates the cognitive processes underlying language use in different experimental tasks, such as a lexical decision task, by applying computations to certain theoretical constructs such as the symbolic representations for word forms and word meanings (Dijkstra 2018).

The integrated lexicon

Word representations are assumed to be stored in a database that is referred to as the Mental Lexicon. At present, the Multilink model correctly recognizes, produces, and translates English and Dutch words, while processing in other languages is currently under investigation. All words are represented in an integrated lexicon. Initially, a base lexicon was used in two Simulation Series assessing the capacities of the model for simulation with use of several Resting Level Activation (RLA) functions. This base lexicon was formed by retrieving all 3 to 8 letter words present in the English lexicon project (ELP; Balota et al., 2007) and the Free Association Database (Nelson, McEvoy & Schreiber, 1998). For these words, the Dutch-English translation pairs were acquired from the Dutch lexicon project (DLP; Keuleers, Diependaele & Brysbaert, 2010) and the English and British lexicon projects (Keuleers et al, 2011). For a word to be included in the lexicon, three conditions had to be met. The first was that the Dutch translation should also be between 3 to 8 letters long. Furthermore, the word pair could not consist of identical cognates, and, finally, the word had to have a reaction time value recorded in either lexicon project. In addition to this information, the word frequency in occurrences per million tokens (opm) were also provided. After the two simulation series showed the capability of Multilink at correctly simulating empirical data, an enriched lexicon was developed from this base lexicon. Cognate pairs and other translation pairs were added to the lexical database.

For the Italian-English study to be presented, the integrated Dutch-English lexicon was extended upon with Italian additions. This led to the creation of the integrated lexicon containing over 1459 translation pairs, containing orthographic and ipa-style phonological representations. The frequencies per million were also acquired from CoLFIS, Corpus e Lessico di Frequenza dell'Italiano Scritto (CoLFIS, 1995).

Assumptions of the Multilink model

To enable accurate simulations of language processes, the Multilink model relies on several assumptions regarding the challenges that computational cognitive models face. In earlier work by Brysbaert and Duijck (2010, p. 368), five challenges to the RHM model were established.

To start, there is the problem of whether the words from different languages are stored separately or together. Convincing evidence in favor of integrated lexicon is that interlingual neighbors affect target word recognition (Van Heuven, Dijkstra & Grainger, 1998). Additional evidence is that word candidates that are morphologically related to a target word are activated even when they belong to another language (Mulder, Schreuder, & Dijkstra, 2013). However, it could be that two lexicons are

functionally separate but with parallel access; sublexical activation could then create resonance among shared lexical features. Sharing of letters and resonance with higher levels might result in language non-selective effects even when the lexicons are functionally separate (Kroll et al. 2010). In the Interactive Activation (IA) approach, languages could be seen as separate to the extent that the size of lateral inhibition differs within and across languages (for instance, it could be weaker during languages). In its 2019 implementation, Multilink assumes there is no lateral inhibitory effect between words at all, neither within nor between languages. However, one could argue there is an element of competition implemented by the frequency element. When the frequency of L1 words is generally larger than that of L2 words, the L1 would have an advantage (Dijkstra, Wahl, Buytenhuijs et al 2018).

With respect to evidence for language (non)selective access, there is agreement that following an input word, lexical candidate sets from multiple languages begin to compete for recognition. In sum, Multilink assumes language non-selective access and parallel activation of word form neighbors. Co-activated orthographic neighbors of the input word are phonologically recoded. This leads to word form neighbors in the target language that may compete with the activation of the correct translation (Dijkstra 2018).

In RHM, the orthographic representations for words are directly linked to their translation equivalent through word association. Where Brysbaert and Duijck argue that such a strong connection between translations should result in strong cross-linguistic priming effects, the evidence for said effects is minimal. Including excitatory connections between lexical translation equivalents risks impeding word recognition, as irrelevant words could be activated in the model's network (Brysbaert & Duijck, 2010). Therefore, in Multilink, word translation is not done by word association but by conceptual mediation instead. Word forms from different languages are only connected on the basis of their meaning (Dijkstra 2018).

Another issue with the RHM model is that the connections between L2 words and their meanings are stronger than perceived in RHM. RHM proposed that beginning learners are more likely to pay more attention to orthography rather than semantics. However, more recent evidence claims learners of a foreign language (L2) quickly link up orthographic representations to their meanings (Maede et al. 2017). To make up for the imbalance of bilinguals with varying L2-proficiency, differentiating in vocabulary usage, Multilink has implemented the frequency-dependent resting level activation.

Besides differences in L1 and L2 proficiency, there is also the issue of language-dependent and language-independent semantic features. Representations of word meaning could have features that are either language-dependent or -independent. An example of this are the French words 'balle' and 'ballon', where depending on which one you use this also describes the features of the ball. It will be either small or hard, while in other languages this distinction is not made (Paradis, 1981). Full meaning equivalence does not hold for all translation pairs, but in the early phase of implementation, Multilink assumes holistic meaning representations that are entirely shared or separate across languages (Dijkstra 2018).

The model's architecture

Making use of these assumptions, a standard network architecture was created. The purpose of the model is providing a general implemented account of word form and meaning retrieval during word

recognition and production (Dijkstra & Rekké, 2010). Combining many characteristics of models, such as BIA+ and RHM, the model was written in the programming language Java, using an object-oriented approach (cf. Roelofs, 2008).

These characteristics can be clearly seen in the figure below, where orthography, phonology, and semantics all play part in spreading input activation throughout the network. As you can see, the written input, represented by the blue bar, activates several orthographic representations. These nodes then continue to activate the semantic and phonological nodes associated with that orthographic representation, while also activating the language membership representations (Dijkstra 2018).

However, the models connections work both ways. While the activation spreads from orthography to phonological representation, it also works in the opposite direction and towards other orthographical similar nodes. To prevent too much activity from happening, the model has a resting level activation (RLA) and an activation threshold, which specifies the amount of activation required for a decision is made. In other words, this threshold indicates at what moment in time the simulated task has been performed. These terms will be further elaborated on in the following sections.

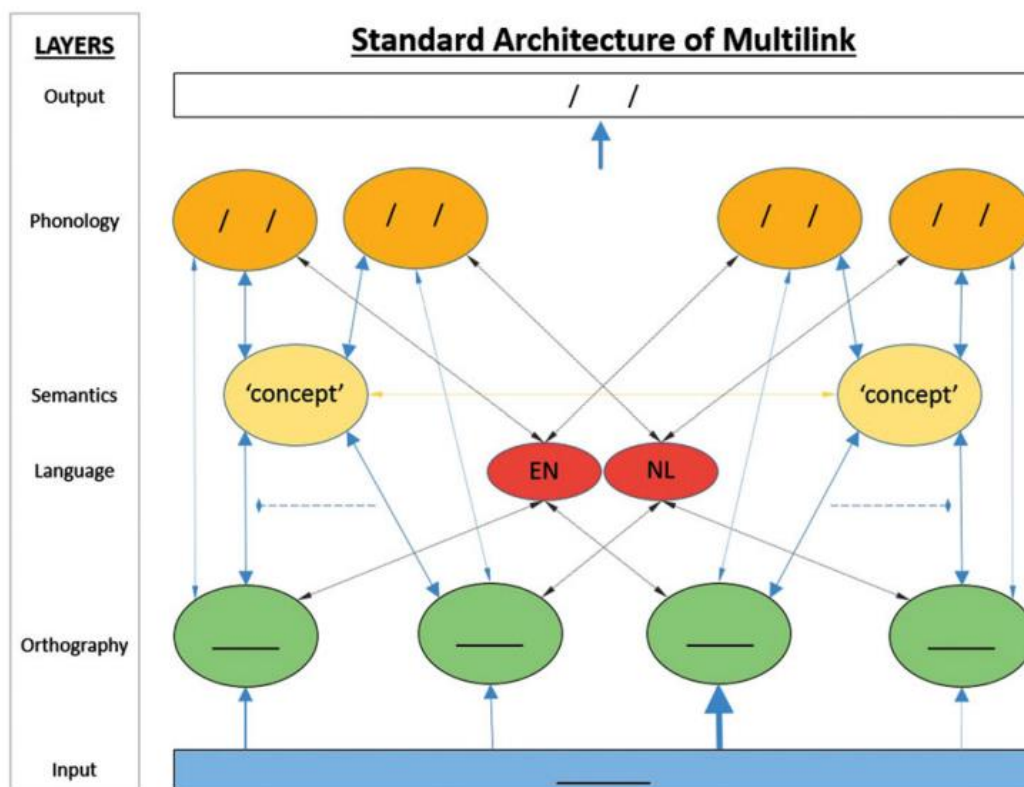


Figure 3. Standard network architecture of Multilink. Note. Input is indicated by blue underscore, orthographic (O) representations as the green, phonological (P) representations by orange slashes. Language as the red nodes, the semantic (S) representations as yellow, and the final output node as white slashes. The dashed line between the connection from O to S indicate that the activation is summed after taking half of the second node's activation input. The output of the structure is task dependent. This figure was taken from "Multilink: a computational model for bilingual word recognition and word translation.," by Dijkstra et al. (2018).

Resting level activation

Resting level activation (RLA) is the base activation of words when they are not in use. It has a value below zero and is dependent on the word's frequency of usage. To be more specific, in the IA and BIA/BIA+ models the RLA depends on the rank of each word. The rank of the word is achieved by sorting the lexicon from high to low frequency. However, assigning rank to the frequencies causes a drastic difference in the word frequency distribution as these frequencies are mapped to the RLA domain. Take, for example, three words that are next to each other in rank, but have different frequencies, using 1, 10 and 100 per million as frequencies. If these would be ranked, this would lead to the same difference in RLA as words with 1, 2 and 3 per million as their frequencies. This, however, does not make sense as the actual frequency differs a lot between the two sets. Besides this, the RLA can also be influenced by changing the lexicon size as the ranks would shift, causing the RLA to also shift.

When comparing model performance for monolinguals and bilinguals, the RLA difference might become problematic as the frequency of L2 words for the bilingual would be lower, ranking them lower, leading to a lower RLA than with the monolingual (Dijkstra, 2018).

A better approach for RLA would be having the RLA distribution mirror that of the reaction time (RT) distribution. Note that RT distributions are very different from rank and frequency distributions. This different approach follows the idea that the activation function can give rise to a sigmoid growth curve for the activation of a node over time. Taking word recognition as an example task, the target word is only recognized when the orthographic node crosses the activation threshold.

Mirroring the reaction time distribution using frequency is achieved by applying a log transformation on the word frequencies. This changes the positively skewed shape of the frequency distribution to become less skewed, leading to a more Gaussian looking distribution. But a remaining problem is that the RT distributions for word recognition are more positively skewed. This problem can be solved by reversing the skew of the word frequency distribution.

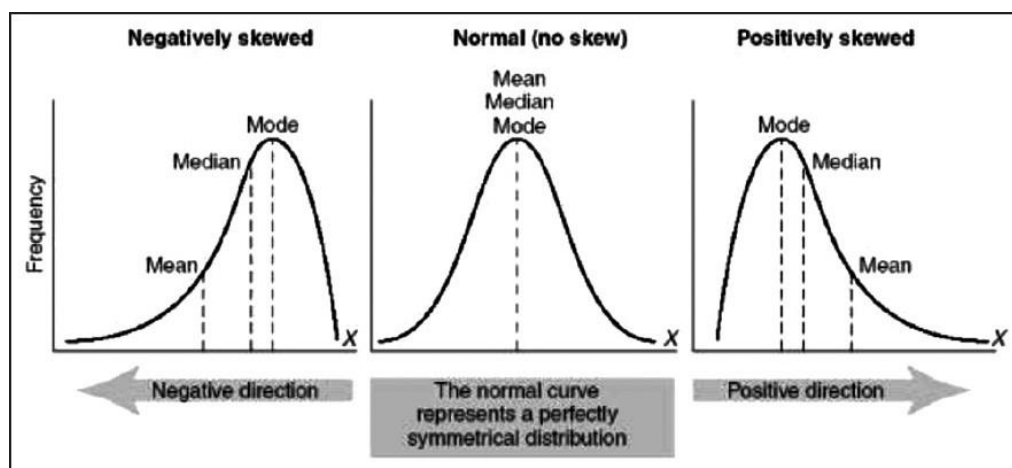


Figure 3. Skewedness of a graph, depicting negatively (left), normal (middle) and positively (right) skewedness. Note. The distributions as described in the text go from right to middle, by applying log transformations. This figure is adapted from 'Descriptive statistics' by South Alabama Edu (2014).

In sum, having the frequency distribution mirror that of the RT distribution and from this the RLA distribution was achieved by applying two equations, where the first equation transformed the word frequencies so that the distribution approximated that of the RT distribution.

$$TF_i = \frac{-1}{\sqrt[3]{opm_i}}$$

The following, second, equation rescaled the outcome of the first equation to an RLA domain between -0.2 and 0.

$$RLA_i = \frac{0.2*(TF_i + |TF_{max}|)}{TF_{max} - TF_{min}}$$

The activation function

Activating the orthographic, semantic, and phonological nodes in Multilink is implemented in a similar way as for the IA model (McClelland & Rumelhart, 1981, 1988) and BIA+ (Dijkstra & Van Heuven, 1998). The activation functions used by Multilink are directly applied to the lexical level, completely skipping the sublexical levels of letters and syllables. The use of the Levenshtein distance creates the opportunity to simultaneously activate words of different lengths (Dijkstra & Rekké, 2010).

Any word retrieval task that is simulated happens in a certain number of time steps or in the case of Multilink, in time cycles. At each cycle the activation of each lexical representation, be it orthographic or phonological, is updated. The following equation shows the net input to a node.

$$n_i(t) = \sum_j k_{ij} e_j(t)$$

This updated value is achieved by adding the damped effect of the current net input from all other connected nodes, including the presented word representation, to its activation at the previous cycle step as the equation below shows.

$$\epsilon_i(t) = \begin{cases} n_i(t) * (M - a_i(t)); & \text{if } n_i(t) > 0 \\ n_i(t) * (a_i(t) - m); & \text{if } n_i(t) \leq 0 \end{cases}$$

This is done while also subtracting the decay.

$$a_i(t + \Delta t) = a_i(t) - \Theta_i * (a_i(t) - rla_i) + \epsilon_i(t)$$

Depending on the similarity, all stored orthographic representation gain activation based on the overlap. The recognition of a word is reached when the activation of a candidate word reaches the 0.72 threshold leading to a final response. If the activation threshold is not met after a certain cycle time, in our case 40 cycles, the task is ended and the final response will be corresponding to no recognition.

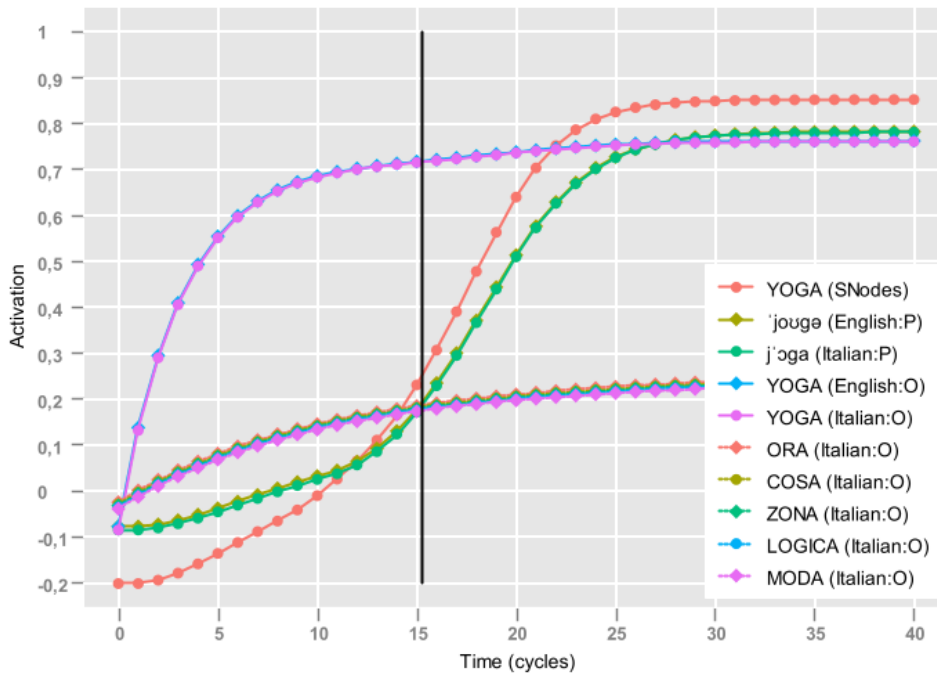


Figure 4. Multilink output graph, simulation of the word Yoga. Note. After 15.1 time cycles the activation reached the threshold of 0.72 leading to a positive response for YOGA (English:O).

Figure 4 shows how different lexical representations are activated over time cycles. The black line indicates the point at which the point of recognition is achieved, which is when the activation reaches the threshold of 0.72. As can be seen the different representations are competing with their activation in the model.

The word 'yoga' is an Italian-English cognate, which means they are orthographically similar and have the same meaning. The Levenshtein distance is Multilink's the solution to word similarity. Levenshtein distance is a marker for the word similarity of orthographic representations. It is mainly used in studies to explore the relation between orthographic, phonetic, and cross-linguistic similarity (e.g., Heeringa 2004; Kessler, 2005; Levenshtein, 1966).

By normalizing the Levenshtein distance, the activation of words with different lengths is possible. The Levenshtein distance describes the number of transitions a word has to undergo to change into the other word. Even though similar words are also activated, the target word is still the one that is recognized. Therefore, Levenshtein distance can be viewed as quite the useful measure to simulate activation of neighbors of a target (Dijkstra et al., 2018).

Referring back to the problem of words of similar orthographic representations, in this study the main focus lies with cognates. When considering identical cognates, these words have a Levenshtein distance of 0, and non-identical cognates which have distances of 1 to 3. Using the normalization of the Levenshtein distance allows for the simulation of said words.

The current parameter settings

How activation spreads throughout the network is affected by the parameter settings of the model. The settings used for the simulations performed with Multilink are added in the appendix, see appendix A. The parameters of interest are those of the PO_alfa and OP_alfa, with a value of 0.03,

which explain the behavior of the bridge between the orthographic and phonological nodes and vice versa. However other parameters are also taken into account. When considering cognates, the activation from two orthographic nodes converges at one specific semantic node. Simply adding this activation together would cause a large overestimation of the cognate effect. Therefore, the input activation is reduced from the least activated node by a factor of 2. The parameter regulating lexical competition is set to zero, implying there is no lateral inhibition caused by lexical items. This variable is still being analyzed with respect to its importance, but this lies outside of the scope of research for this study.

Chapter 3: Simulating the Italian lexicon

To allow a simulation of lexical processing by Italian-English bilinguals, we first need to add an Italian lexicon to Multilink (English is already available). The lexicon used for the simulations contained a total of 1459 English-Italian word pairs. These words range from 3 to 8 letters in length. The purpose of simulating the entire Italian lexicon is to establish the capabilities of Multilink when faced with Italian words compared to English. Multilink will run a simulation with the entire lexicon as the stimulus set. The cycle times achieved by Multilink will be used to calculate a correlation coefficient with the frequencies of the word pairs represented in the integrated lexicon, keeping in mind that a subtitle database like SUBTLEX-UK is not available for Italian.

The expectation is that the orthographic frequency will most likely negatively correlate with the simulated cycle times, meaning that the cycle times are expected to be lower when the frequency tends to be higher. This correlation effect is expected for both languages, even though they might differ slightly. But before this could be researched changes had to be made on the lexicon. As the lexicon is required to be of certain standards.

For starters, the phonetic notation was set to the International Phonetic Alphabet notation (IPA), as this notation is more concise and more accurately interpretable by new users. Word frequencies were gathered by use of the "Corpus e Lessico di Frequenza dell'Italiano Scritto" (CoLFIS). These Italian frequencies were gathered from excerpts from newspapers, published between 1992 and 1994, magazines, and books, including those relating to professional interests (CoLFIS, 1995).

Correlation between frequencies and cycle times

After completing the lexicon, the simulation of the entire lexicon in Multilink was done. The parameter settings were the same as discussed in the Introduction. Of all 1459 words in the stimulus set, all 1459 returned a correct response, which means all words had attained an activation of 0.72 within 40 time cycles. The simulation performed by Multilink shows a strong negative correlation of -0.937 (N=1459) between the achieved cycle times and the frequencies. However, this correlation should be compared to the correlation achieved by simulating the English words and their frequencies. This is done to ensure there is no difference in behavior between the languages and their respective frequencies. Of all 1459 words in the stimulus set, 1453 returned correct responses and led to a correlation of -0.919 (N=1453).

These correlations confirm the hypothesis that the frequencies and cycle times are negatively correlated. Plotting the frequencies against the simulation results would result in the following figures, which clearly represent the simulated behavior of Multilink mimicking that of the word frequencies.

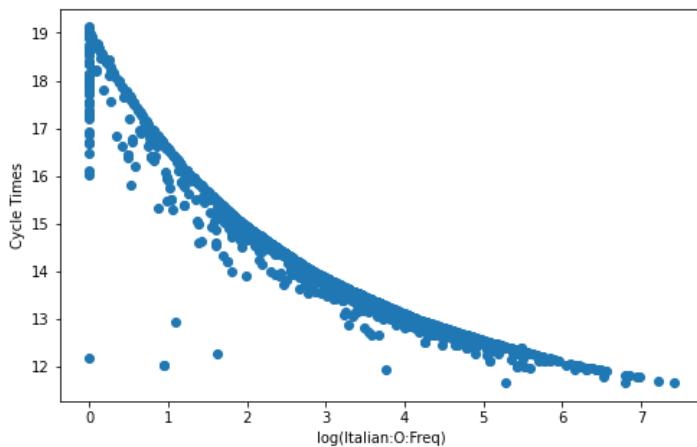


Figure 5. Depicting the cycle times against their respective log(Italian frequency) for each stimulus N=1453 Note. The same shape is formed when using the English frequencies, only differing slightly in cycle times due to frequency value difference discussed later on.

As can be seen the negative correlation is clearly represented for the entire lexicon, in the case of Italian the line shows to be very negatively correlated. The higher the log(frequency) the lower the simulated cycle time, indicating the strong influence of the frequencies on the activation in the model.

Justifying the correlations

The correlations achieved by both simulations and their respective frequencies indicate similar behavior as described in the curve of the figure above. A strong negative correlation was found being above -0.90. As can be seen, there is a slight difference in the correlation strength, but only to such a small degree that the difference can be viewed as insignificant. It can be explained however, as it must be noted that English and Italian have a different kind of source for their frequency values. While the Italian values are gathered through CoLFIS, the English values have been gathered through the English lexicon project. Both are obviously based on literature present in their respective languages, and therefore incorporate cultural differences and differences in vocabulary usage. However, the main difference lies in the collected literature types, where CoLFIS focusses on newspapers and books while the frequencies gathered for the ELP were achieved by taking the subtitles created for BBC programs. Keeping in mind this difference of frequency gathering and how the frequency effects the cycle times for Multilink, it can be explained how the cycle times differ in actual values, but not in behavior as the shape of the figure 5 depict.

Chapter 4: Simulating the Italian empirical data

After laying the foundation for the simulation of the Italian lexicon with Multilink, the empirical data introduced earlier on can now be examined, in particular the recent study by Cangelosi and Dijkstra. As previously stated, this study examined the effect of varying degrees of cross-linguistic English-Italian overlap on bilingual word recognition, by varying Levenshtein distances instead of collected ratings. This study involved Italian-English participants performing an Italian-English lexical decision task.

The authors' prediction (based on Dijkstra et al., 2010) was that more cognate facilitation would occur relative to non-cognates the more similar the cognate members would be cross-linguistically. It was also expected that identical cognates, like pizza and radio, would show even more facilitation than cross-linguistically very similar but not identical cognates, like numero and angelo.

The original experiment was an Italian-English lexical decision task in which the participant was faced with 428 different stimuli, of which 50% were non-words and 50% actual English words. The participants were asked to press the correct button 'a' (for a nonword) or 'l' (for a word), based on each presented stimulus.

The mean reaction time for each word over 17 participants was used for the simulation with Multilink. First, all word data from participants with an error rate of above 20% and all nonword responses were excluded from the simulation. The reason for this is that, taking into account the small number of current participants, the effect of one participant failing to correctly respond could be either representative for a large group of participants or a simple outlier. Furthermore, the empirical data were cleaned before the simulations were run. First, items with an incorrect response rate of 50% or higher were excluded from further processing. For these items, it can be noted that multiple participants did not know the word. Next, the reaction times that fell below 200 ms and above 2000 ms were excluded. These reaction times can be considered as nonvalid; they imply either a guess on the part of the participant or were simply too long to be taken as meaningful reflections of underlying processing.

Simulating the lexical decision task

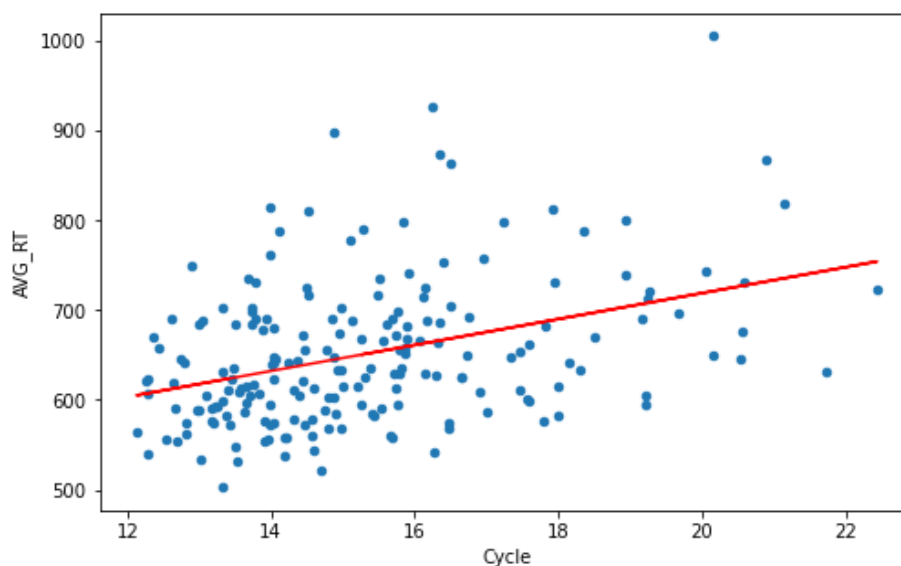


Figure 6. Empirical results compared to the simulated results by multilink. Note. The red line indicates the best fit line between the experiment 1 results and the lexical decision task performed by multilink. N=196.

After simulating the processing of each of the 196 words in the stimulus list, the correlation plot in figure 6 was achieved, showing the average reaction times of the experiment against the cycle times provided by Multilink. The comparison resulted in a correlation of 0.38 (N=196), which can be considered positive to a moderate degree (relative to earlier studies with Dutch and English). However, further interest of the empirical study lies with the cognates and the performance differences between the various cognate types.

Considering cognate pairs varying in Levenshtein distance

Next, the average cycle time and the average reaction time were determined for each of the Levenshtein distance categories separately. For each cognates type in the stimulus list (varying in Levenshtein distance between 0 and 6), the Pearson correlation was calculated between the average reaction times and average cycle times.

Levenshtein distance	0	1	2	3	4	5	6
Number of items	31	28	26	25	34	29	23
Correlation	0.429	0.342	0.161	0.583	0.329	0.378	0.566

The figures 7a and 7b below represent the behavior of both the participants and Multilink when confronted with the stimuli.

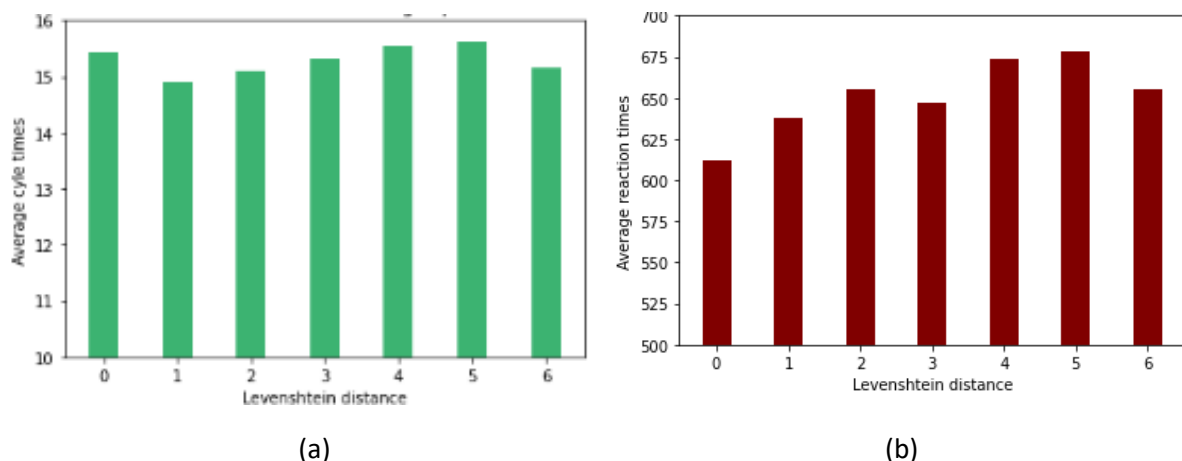


Figure 7. a) Simulated results from multilink using the lexical decision task. B) Empirical results from Experiment 1, showing the fluctuation between average reaction times when compared between levenshtein distance. Note. The fluctuation depicted in figure a is based on the average results of the participants in experiment 1 that responded correctly in the lexical decision task. N=196.

The simulated results from the lexical decision task show some differences compared to the empirical reaction times. The first noticeable difference is that the identical cognate group displays the highest cycle time of any group, while the average reaction times of the participants indicate this should be the lowest. One potential reason for this remarkable finding is that the subjective frequencies in the lexicon for identical cognates were lower than for the other cognate types. Remember that we previously established that Multilink fundamentally follows word usage frequencies when simulating cycle times. However, it might underestimate the contribution of some factors in the lexical decision task. In the following section, we will see, however, that there is another more likely cause for the deviating behavior of the identical cognates.

Simulations with a generalized lexical decision task

The presence of identical cognates in the experiment might induce participants to respond to the Italian counterpart of these items, rather than the English one. They might, for instance, respond to the Italian word AGENDA rather than the English one. In all cases, they would still give the correct answer ('yes'). For this reason, we decided to conduct our simulations again, but now using generalized lexical decision as a task. In generalized lexical decision, participants respond to any word they detect, irrespective of the language to which it belongs. The figures 8a and 8b below represent the performance of both the participants and that of Multilink assuming a generalized Italian-English lexical decision task. The motivation underlying this comparison is hereby named the "recognition is recognition" effect. This effect basically entails that irrespective of the language in which the participant recognizes the word, they will be more eager to press 'yes'. When the activation threshold of any language's orthographic representation can lead to a yes response, the average cycle times become shorter, because a recognition of the stimuli as an Italian word (instead of an English target word) is also accepted as valid response.

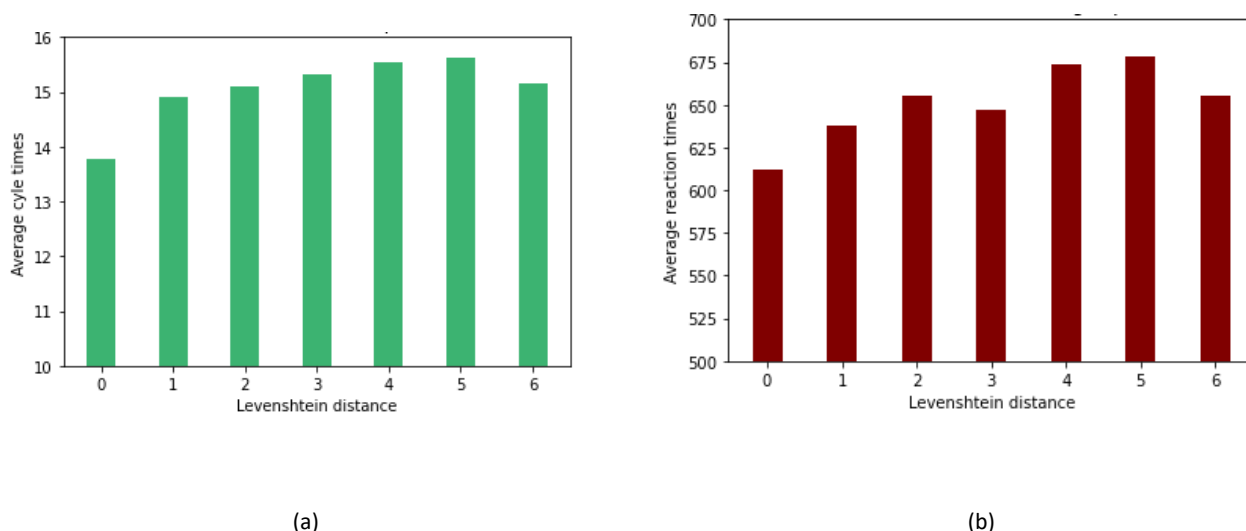


Figure 8. a) Simulated results from multilink using the generalized lexical decision task. b) Empirical results from Experiment 1, showing the fluctuation between average reaction times when compared between levenshtein distance. Note. The decrease in average cycle times for identical cognates, levenshtein distance 0, compared to figure 7a. is very clear, showing a lot more similarity between figures a and b.

As will be noted by comparing the new simulation figure to the older one, there is a clear difference in response to the identical cognates, having a Levenshtein distance of 0. All other cognate and control types are responded to on the basis of their Italian reading, making their cycle times identical to the earlier simulation. Nevertheless, the overall between the reaction times and cycle times across the seven categories increased to 0.43 (N=196). The overall correlation between the average reaction times and average cycle times, combined over Levenshtein distances, when simulating the generalized lexical decision task is 0.93 (N=7). Although only based on N=7 this correlation is very strong, reflecting the strong similarity in model and human behavior.

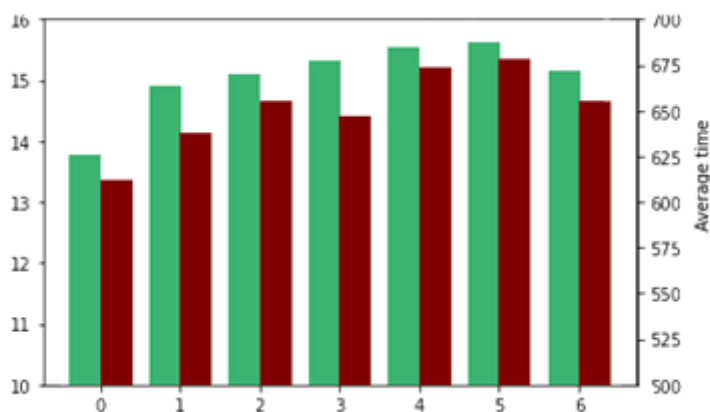


Figure 9. The results of the generalized lexical decision task simulation of the stimulus composition list against the empirical results for the stimulus composition list of experiment 1. Note. The clear depiction of the similar behavior can be viewed as described in the text.

Figure 9 puts these plots next to each other. Here it becomes even more clear how similar empirical and simulation performance is. These data suggest that the generalized lexical decision task is the more appropriate approach to simulate the participants's performance in the experiment conducted by Cangelosi and Dijkstra. As this suggests that participants did not only say 'yes' when they encountered the Italian reading of identical cognates, but also to the earlier available English reading.

As explained above, the Italian-English experiment by Cangelosi and Dijkstra intended to replicate a similar experiment done previously by Dijkstra et al. (2010) as a basis for further research. As in the lexical decision task by Dijkstra et al., an increasingly strong orthographic similarity led to a shorter reaction time in our study. It is possible that the Dutch-English participants in Dijkstra et al. had a higher L2 proficiency than the Italian-English participants from Cangelosi and Dijkstra, and therefore (as instructed) performed their lexical decision task in a language specific way (e.g., responding only when an English reading of the cognate was recognized). However, here we will further explore the simulations based on the generalized lexical decision task, which allow the strategy of responding to Italian readings of cognates (although it was probably only applied in the case of identical cognates).

Based on Figures 7 to 9 we can make a clear distinction of three groups in the stimuli: identical cognates with a Levenshtein distance of 0, non-identical cognates with a Levenshtein distance of 1, 2, and 3, and a final item group of other words having a Levenshtein distance of 4 or more to competitors. In the following sections, we will examine the three groups of items one by one.

Identical cognates

The set of identical cognates sharing the same orthographic and meaning representation as their counterpart in English, consisted of 31 items. The correlation observed between the average reaction time and cycle times for this specific group is 0.566 (N=31), which is a considerable to high correlation. Note that it would increase even further when some outlier items having contradicting reaction and cycle times would be removed.

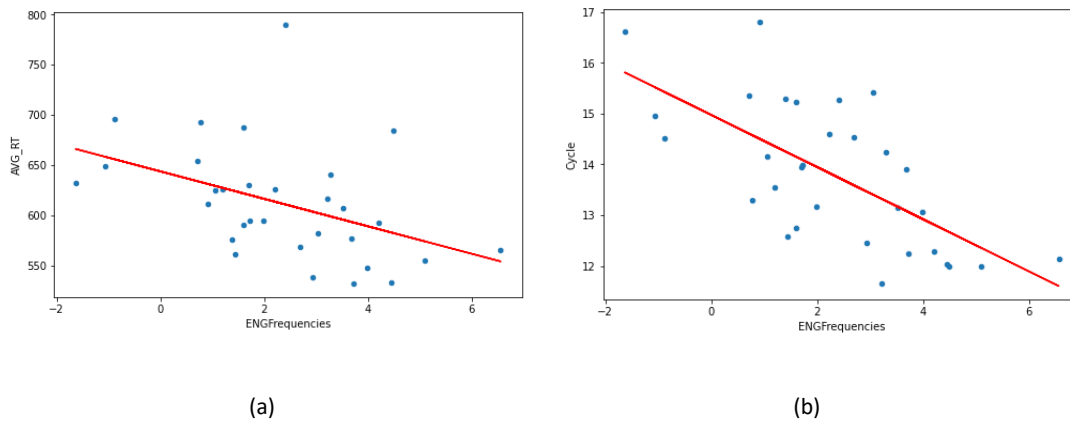


Figure 10. a) Empirical results correlated to the log(English frequency) of the identical cognates from the stimulus list. b) Simulated results correlated to the log(English frequency) of the identical cognates from the stimulus list. Note. The figure depicts the correlation between those variables, indicating a strong negative correlation for the identical cognates.

The figures above present the average reaction times and cycle times against the log frequencies of the English representations. Reaction times had a moderate negative correlation of -0.436 ($N = 31$) with log frequency. Cycle times had an even stronger negative correlation of -0.666 ($N = 31$). In other words, the model is very focused on frequency as a variable to represent the language usage of the participant.

Non-identical cognates

For the non-identical cognates, words with a Levenshtein distance of 1, 2, or 3, the following figures comparing response and cycle times were created. The simulations indicate that for non-identical cognates and control words, the model response is always based on the English reading of a target word. From the figures 11a and 11b below it is already apparent, the behavior is not entirely the same. As for the Levenshtein distance of 2, the empirical results seemed to be higher compared to the others. While Multilink simulates the expected behavior of increasing cycle times the more orthographic dissimilar the stimuli get, it is interesting to see that the empirical data does not support this.

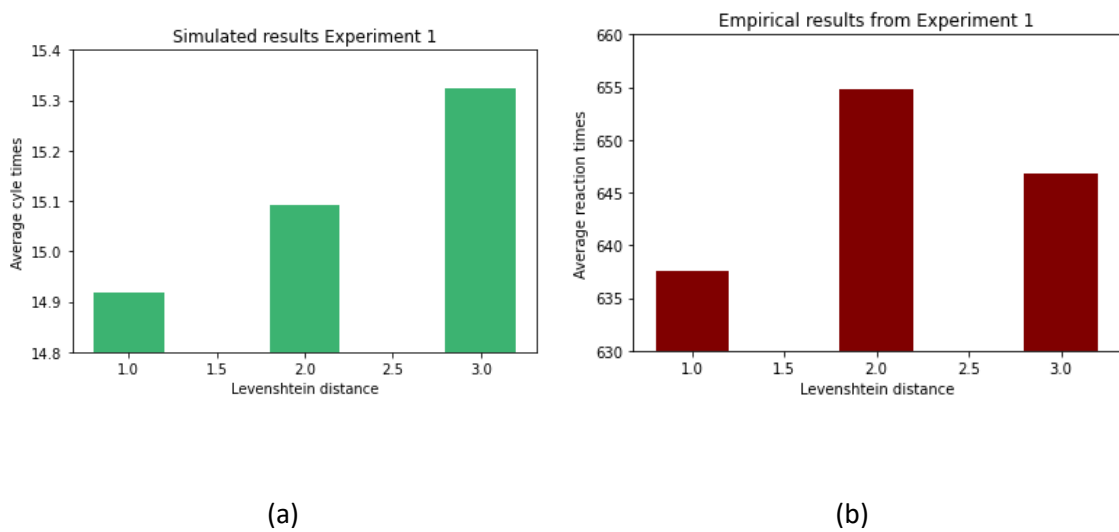


Figure 11. a) Simulated results of experiment 1 for the generalized lexical decision task by multilink. b) The empirical results of experiment 1, the average reactions times. Note. These are the results of the stimuli from the stimulus list with levenshtein distances 1 through 3, the non-identical words.

Looking at the figures displaying the cycle and average reaction times compared to the $\log(\text{frequencies})$ it can be seen how Multilink showcases the clear frequency effect, gaining a remarkable -0.928 correlation ($N=3$), while also achieving -0.49 for $N=79$. Comparing that to the -0.321 correlation ($N=79$) of the average reaction times and the frequencies it can be understood how a difference in behavior could have been expected.

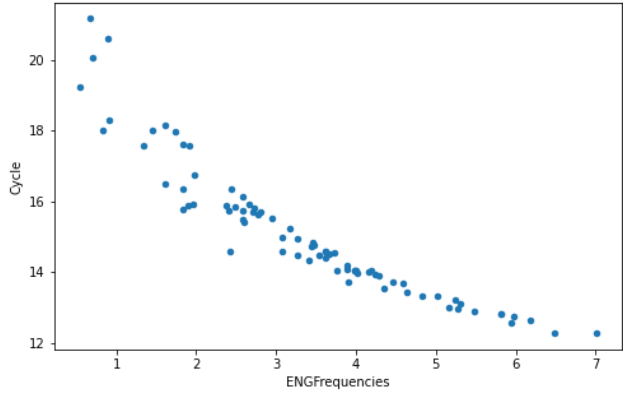


Figure 12. Simulated results compared to the $\log(\text{English frequency})$ of the non-identical cognates by multilink. Note. The spread depicted shows a strong negative correlation as previously established with the entire lexicon.

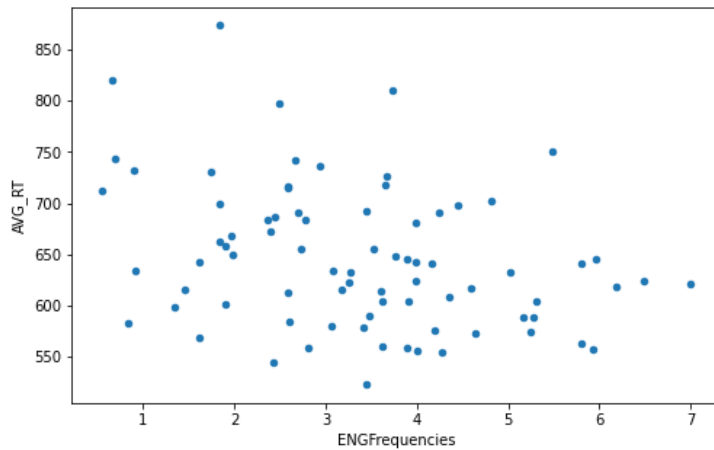


Figure 13. Empirical results compared to the log(English frequency) of the non-identical cognates. Note. The spread depicted shows no real correlation to a very strong degree, this therefore explains the difference in fluctuation between the previous figures 11a and 11b.

Control words

The final group of stimuli consists of words with Levenshtein distances above 3, the ‘other word’ category. These words are orthographically dissimilar, and can be considered as English control words in the sense that they will be least affected by items from the non-target language, Italian. The Pearson correlation between Multilink’s average cycle times and the obtained response times for these items was 0.99997 (N=3). A correlation as this is significant as the p value is smaller than 0.001, while also achieving a correlation of -0.50 for N=86.

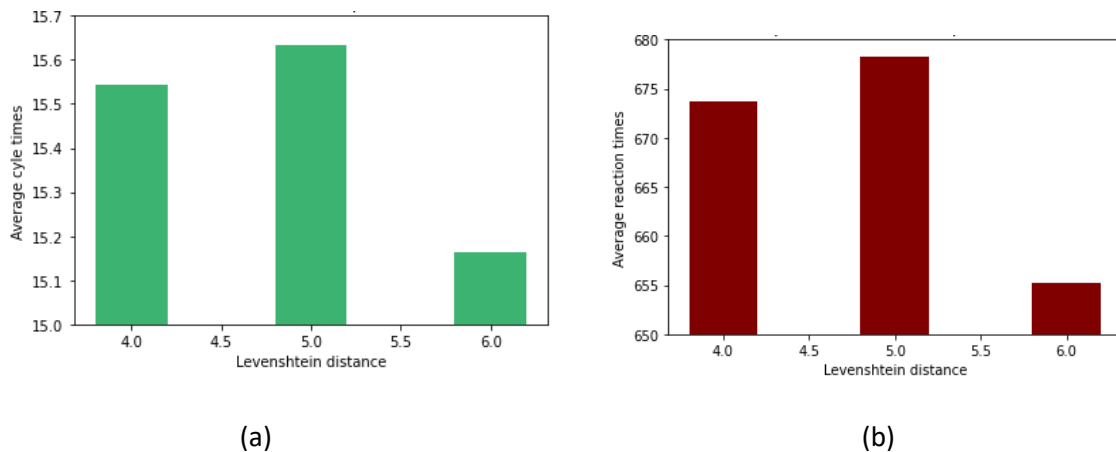


Figure 15. a) The simulated results for the control words by multilink. b) Empirical results for the control words. Note. The same fluctuation can be seen between the levenshtein distances.

As can be seen by comparing the simulated cycle times to the log of the English frequency, where the same curve can be seen as well as, the simulation result closely resembling the RT pattern, which itself is similar to that in Dijkstra et al. (2010, Figure 3, p.12).

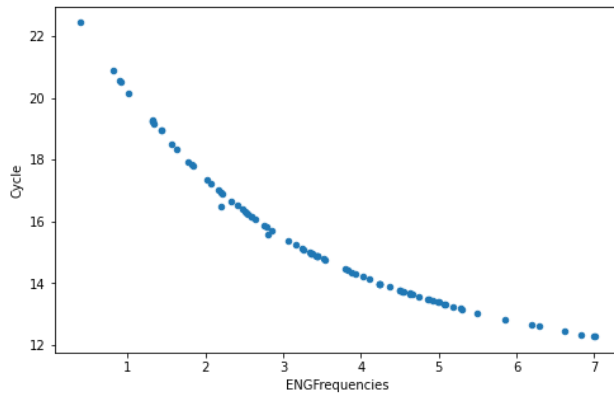


Figure 16. Simulated results correlated with the log(English frequency). Note. The figure depicts a strong negative correlation between the two variables, combined with the fact that the fluctuation show similar patterns confirms Multilink’s capabilities of simulating control words. And also showing the expected results as earlier in 2010.

Conclusion

In this chapter, we described and simulated a recent experiment by Cangelosi and Dijkstra. In an English lexical decision task, Italian-English bilinguals responded more quickly to identical cognates than to non-identical cognates, which were responded to more quickly than purely English control words.

Simulations with Multilink, a computational model for monolingual and bilingual word retrieval, were able to replicate this pattern of results. However, the model’s simulations strongly indicate that participants responded to the Italian (L1) reading of identical cognates, rather than to their English (L2) reading. Although they were instructed to respond to words only when they were also English, no error would not result in the case of identical cognates when their response was based on Italian.

In sum, Multilink represents the participants’ performance in the experiment rather well, as is indicated by a strong positive Pearson correlation between cycle times and reaction times. The argument was also strengthened by a consideration of the relationship between word frequency, cycle times, and reaction times of the target items in the experiment. Results produced by Multilink were also similar to those achieved by Dijkstra et al. (2010). In all, we conclude that the extension of Multilink to Italian was successful. At present, Multilink is the only computational model capable of correctly simulating human performance in an English-Italian lexical decision task.

Chapter 5: General discussion

The first goal of this research was to construct the making Italian lexicon – By extending the Dutch/English lexicon, the integrated lexicon for this study was created. Expanding the lexical data with more words and IPA notation for the phonological representations. Using colfis the frequencies were acquired leading to a complete and useable integrated lexicon. The lexicon used for the simulations contained a total of 1459 English-Italian word pairs. These words range from 3 to 8 letters in length.

Second, we wanted to perform a simulation for the Italian lexicon as a whole – After completing the integrated lexicon, the next step was the simulation of the entirety of it. The purpose of simulating the entire Italian lexicon is to establish the capabilities of Multilink when faced with Italian words compared to English. Of all 1459 words in the stimulus set, all 1459 returned a correct response, which means all words had attained an activation of 0.72 within 40 time cycles. And as expected, a negative correlation was found for both English as Italian frequencies. For Italian a correlation of -0.937 (N=1459) between the achieved cycle times and the frequencies was achieved. However for English a correlation of -0.919 (N=1453) was achieved, therefore it was justified to use this lexicon.

Third, we aimed at simulating an Italian-English lexical decision task – Cangelosi & Dijkstra – Now that the lexicon was ready, incorporating all the items required for simulating, the next step was to focus on the Italian-English lexical decision task performed by Cangelosi & Dijkstra. The goal of the experiment was to examine the effect of cross-linguistic English-Italian overlap with varying levenshtein distances. The empirical data involved Italian-English bilinguals performing an Italian-English lexical decision task. It was hypothesized, the more cognate facilitation, the stronger the activation compared to non-cognates. As the entire stimulus composition list was incorporated into the lexicon the simulations were done. Before analyzing the results, these were thoroughly filtered, leading to 196 final items to be analyzed, see Appendix B.

The figures incorporated in chapter 4 show the similarities in behavior between Multilink and the empirical data. However when the first correlation was calculated a moderate correlation of 0.38 (N=196) was achieved. Therefore after looking at the simulated results it was hypothesized another effect was at play. The recognition is recognition effect, causing the participants to likely respond yes even though they recognized the word only as Italian but figured it might as well be English. Therefore the results were simulated again as a general lexical decision task, allowing for only the identical cognates to also be recognized as Italian words. This increased the correlation to 0.42 (N=196) and 0.93 (N=7) when taking the averages of each levenshtein distance category.

This way of analyzing the results lead to a clear distinction of 3 different groups within the stimulus list. The first group being the identical cognates N = 32, with a levenshtein distance of 0. Where the empirical results had a moderate negative correlation of -0.436 (N = 31) with the log frequency. And the simulated times had an even stronger negative correlation of -0.666 (N = 31).

The second group of items, the non-identical cognates N =79, with a levenshtein distance range of 1 to 3. For this group a difference in behavior was noted, but as expected as the log frequencies were also not strongly correlated, -0.32, with the reaction times. While a correlation of -0.49 was achieved between the log frequencies and the cycle times.

Finally, the third group of items of 86 control words, N=86, with a levenshtein distance above 3. For this group a correlation between the log frequencies and the cycle times of -0.5 was achieved.

The results represent the expected behavior of how more the cognate facilitate effect, how faster the reaction and cycle times. The simulations done by multilink show exactly that, and even similar to that of the participants in the study. The results of those simulations are similar to those of Dijkstra et al 2010, further supporting the possibility of the hypothesis proposed by Cangelosi and Dijkstra.

The simulation of an Italian lexical decision task had never been done before, and that did therefore bring some issues along, as the lexical data available are rather scarce. This led to time issues and extensions on deadlines to still be able to achieve these findings. However in the end still only 17 different participants results were used, and these results were then also thoroughly filtered, leading to rather small groups of items for each levenshtein distance. Meaning the results gained from this research can become more definitive, as with more participants the analysis of the simulated results can become more trustworthy.

While stated in chapter 4 that results showing similar behavior as that of Dijkstra et al. 2010, is found the results from Cangelosi and Dijkstra's 2022 research using the empirical data of experiment 1 is yet to be released. However the hypothesis has been supported by this research, therefore the expectations to find the expected cognate facilitate effect is quite likely.

For future research the interesting possibilities would be either to look at other languages or words of longer length. For the first idea, comparing the cognate processing between Spanish and Italian speakers would lead to interesting insights. Currently results for both Spanish and Italian cognate processing simulated by Multilink is present. While the languages share similar roots, throughout history different influences have altered them in their respective ways. It would therefore be very interesting to research whether similar results as for Italian can be achieved with simulations for Spanish. As the results of this research and other such as Dijkstra et al., 2010, have shown that Multilink is capable of correctly simulating the behavior for lexical decision tasks. Achieving results similar to those researches when simulating Spanish, would therefore be the goal.

While for the second option, increasing word lengths would challenge Multilinks capabilities, but also the way of indicating degree of similarity should perhaps change. As the current stimuli used varied between lengths of 3 and 8, however if one would use larger words, the orthographic similarity of a word with length 12 and levenshtein distance 3 could be considered as similar as a word of length 4 and levenshtein distance 1, both could be seen as similar when using percentages.

But looking back we can say a Italian-English lexicon has successfully been created. Which led to promising results when simulated using Multilink, as it clearly showed the frequency-effect. Using these results as a foundation showing the capabilities of Multilink, we then succeeded in simulating an actual Empirical research. Simulating an Italian-English lexical decision task, which led to fitting results. These results have indicated that the word recognition process of Multilink clearly works for more than just English and Dutch as previously established.

Appendix

A: Multilink parameters

```
"description": "Multilink standard parameters",
"MIN_ACT": -0.2,
"MAX_ACT": 1.0,
"DECAY_RATE": 0.07,
"MIN_REST": -0.2,
"MAX_REST": 0.0,
"MAX_OPB": 0.6402259325203161,
"I_rest": 1.0,
"O_rest": "NOTE: now defined in params/Activation.java due to deprecation of Nashorn Engine",
"L_rest": -0.2,
"S_rest": "NOTE: now defined in params/Activation.java due to deprecation of Nashorn Engine",
"P_rest": "NOTE: now defined in params/Activation.java due to deprecation of Nashorn Engine",
"IO_alpha": "NOTE: now defined in params/Activation.java due to deprecation of Nashorn Engine",
"IO_multiplier": 0.2,
"connectionParameters": {
  "OL_alpha": 0.0,
  "LP_gamma": 0.0,
  "SP_alpha": 0.3,
  "PO_alpha": 0.03,
  "LO_gamma": 0.0,
  "OS_alpha": 0.03,
  "OP_alpha": 0.03,
  "PS_alpha": 0.3,
  "PP_gamma": 0.0,
  "LP_alpha": 0.0,
  "SO_alpha": 0.03,
  "OL_gamma": 0.0,
  "LO_alpha": 0.0,
  "PL_gamma": 0.0,
```

```

"PL_alpha": 0.0,
"OO_gamma": 0.0,
"SS_gamma": -0.05
}

```

B: Stimulus composition list experiment 1

Stimulus with its respective Levenshtein distance.

YOGA 0	TRAM 0	SODA 0	IDEA 0
TEST 0	PIZZA 0	RADIO 0	TENNIS 0
BANANA 0	SPORT 0	GHETTO 0	GAS 0
PASTA 0	AGENDA 0	EXTRA 0	BAR 0
PIANO 0	BONUS 0	VIDEO 0	CACTUS 0
AREA 0	ZEBRA 0	FILM 0	SAGA 0
STOP 0	PONY 0	MANGO 0	ZERO 0
BOOM 0	POKER 0	AUDIO 0	ROSE 1
PURE 1	VALLEY 1	LOGIC 1	EAST 1
MELON 1	ANGEL 1	ACID 1	POEM 1
ECHO 1	VOICE 1	FATAL 1	MASS 1
DRAMA 1	SALT 1	CAUSE 1	POET 1
RUMOR 1	MUSIC 1	PEACE 1	LARGE 1
FINAL 1	NATURE 1	POSE 1	OCEAN 1
PART 1	PAUSE 1	NOTE 1	NUMBER 2
SEED 2	LION 2	BREEZE 2	THRONE 2
NOSE 2	COLOUR 2	RICH 2	TYPE 2
CIRCUS 2	PEARL 2	CODE 2	STORY 2
BALL 2	PROOF 2	MEMBER 2	DEBT 2
DUKE 2	CARROT 2	ANGLE 2	STUDY 2
SIGN 2	NORTH 2	HUMAN 2	VALUE 2
POWER 2	FAIRY 3	BEARD 3	TRAIN 3
STREET 3	CHOIR 3	TOWER 3	SCHOOL 3
FACE 3	SURE 3	PLATE 3	ORPHAN 3
RAIL 3	SILK 3	MIND 3	POOR 3
FEVER 3	SPACE 3	PAGE 3	ACTION 3
JUICE 3	ORDER 3	ISLAND 3	BEER 3
END 3	SOIL 3	EMPTY 4	GRASS 4
RHYTHM 4	GRASS 4	MOUSE 4	SNOW 4
SWORD 4	HEAD 4	SOAP 4	HAMMER 4
SHARK 4	LIGHT 4	BOOK 4	WIND 4
YEAR 4	SOUP 4	SHIP 4	MOON 4
HAMMER 4	SHARK 4	CHURCH 4	BODY 4
DOUBT 4	CHAIN 4	ROAD 4	FIRE 4
FEAR 4	HOUSE 4	WING 4	WOMAN 4

BULL 4	SCREW 4	WALL 4	CHAIR 4
SHARP 5	DUTY 5	HUNGER 5	THIEF 5
HONEY 5	SUMMER 5	BREAST 5	WARMTH 5
MILK 5	WATER 5	THROAT 5	ENGINE 5
EAGLE 5	WIFE 5	DIRT 5	CROWD 5
SHAPE 5	MAIL 5	GUILT 5	NOISE 5
TREE 5	MOVIE 5	LAZY 5	HUGE 5
CAGE 5	MONEY 5	SKIRT 5	UNCLE 5
ROOF 5	SWAMP 6	ROOM 6	STONE 6
SMOOTH 6	CHOICE 6	DONKEY 6	FAST 6
SLEEVE 6	TOOL 6	WITCH 6	DUCK 6
POCKET 6	PEEL 6	DAY 6	SON 6
PRISON 6	DROP 6	ELBOW 6	BRIDGE 6
SHEEP 6	FOG 6	PENCIL 6	LEAD 6

References

- Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3), 445–459. <https://doi.org/10.3758/bf03193014>
- Brysbaert, M., & Duyck, W. (2010). 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), 359–371. <https://doi.org/10.1017/S1366728909990344>
- Comesaña, M., Ferré, P., Romero, J., Guasch, M., Soares, A. P., & García-Chico, T. (2015). Facilitative effect of cognate words vanishes when reducing the orthographic overlap: The role of stimuli list composition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 614–635. <https://doi.org/10.1037/xlm0000065>
- Dijkstra, T., & 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* (pp. 189–225). Lawrence Erlbaum Associates Publishers. <https://psycnet.apa.org/record/1998-06129-006>
- Dijkstra, T., & van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5(3), 175–197. <https://doi.org/10.1017/s1366728902003012>
- Dijkstra, T., & 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* (pp. 189–225). Lawrence Erlbaum Associates Publishers. <https://psycnet.apa.org/record/1998-06129-006>
- Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62(3), 284–301. <https://doi.org/10.1016/j.jml.2009.12.003>
- Dijkstra, T., & Rekké, S. (2010). Towards a localist-connectionist model of word translation. *The Mental Lexicon*, 5(3), 401–420. <https://doi.org/10.1075/ml.5.3.08dij>
- Dijkstra, T., Wahl, A., Buytenhuijs, F., van Halem, N., Al-jibouri, Z., de Korte, M., & Rekké, S. (2018). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 22(04), 657–679. <https://doi.org/10.1017/s1366728918000287>

- Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change*, 16(03).
<https://doi.org/10.1017/s0954394504163023>
- Kessler, B. (2005). Phonetic comparison algorithms. *Transactions of the Philological Society*, 103(2), 243–260.
<https://doi.org/10.1111/j.1467-968x.2005.00153.x>
- Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice Effects in Large-Scale Visual Word Recognition Studies: A Lexical Decision Study on 14,000 Dutch Mono- and Disyllabic Words and Nonwords. *Frontiers in Psychology*, 1. <https://doi.org/10.3389/fpsyg.2010.00174>
- Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2011). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, 44(1), 287–304.
<https://doi.org/10.3758/s13428-011-0118-4>
- Kroll, J.F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition*, 13(3), 373–381.
<https://doi.org/10.1017/s136672891000009x>
- Kroll, J., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language*, 33(2), 149–174. <https://doi.org/10.1006/jmla.1994.1008>
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375–407.
<https://doi.org/10.1037/0033-295X.88.5.375>
- McClelland, J. L., & Rumelhart, D. E. (1988). Explorations in parallel distributed processing: A handbook of models, programs, and exercises. The MIT Press.
- Meade, G., Midgley, K. J., Dijkstra, T., & Holcomb, P. J. (2018). Cross-language Neighborhood Effects in Learners Indicative of an Integrated Lexicon. *Journal of Cognitive Neuroscience*, 30(1), 70–85.
https://doi.org/10.1162/jocn_a_01184
- Mulder, K., Schreuder, R., & Dijkstra, T. (2013). Morphological Family Size effects in L1 and L2 processing: An electrophysiological study. *Language and Cognitive Processes*, 28(7), 1004–1035.
<https://doi.org/10.1080/01690965.2012.733013>
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. <http://www.usf.edu/FreeAssociation/>.
- Paradis, M. (1981). A neurolinguistics theory of bilingualism.
<https://books.google.nl/books?id=Unksj1qgeeMC&pg=PA110&lpg=PA110&dq=paradis+1981&source=bl&ots=QVsjMvOJYM&sig=ACfU3U3DPMgGTtK6p6IDzrM8s1LKTQxn6Q&hl=en&sa=X&ved=2ahUKEwjott3Sot74AhXe8LsIHWBFBs4Q6AF6BAGCEAM#v=onepage&q=paradis%201981&f=false>
- Potter, M. C., So, K. F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23(1), 23–38.
[https://doi.org/10.1016/s0022-5371\(84\)90489-4](https://doi.org/10.1016/s0022-5371(84)90489-4)
- Roelofs, A. (2008). Dynamics of the attentional control of word retrieval: Analyses of response time distributions. *Journal of Experimental Psychology: General*, 137(2), 303–323. <https://doi.org/10.1037/0096-3445.137.2.303>
- Roma Università di Salerno. (1995). *Corpus e Lessico di Frequenza dell'Italiano Scritto*. Istituto di Scienze e Tecnologie della Cognizione, CNR.

Vanlangendonck, F. (2012). Conflict in the bilingual brain: The case of cognates and false friends. *Proceedings of the Master's programme Cognitive Neuroscience*, 7, 61–78.

Vladimir I. Levenshtein, Binary codes capable of correcting deletions, insertions, and reversals, *Doklady Akademii Nauk SSSR*, 163(4):845-848, 1965 (Russian). English translation in *Soviet Physics Doklady*, 10(8):707-710, 1966.