# Top-down language-to-word inhibition effects in bilingual speakers

Victor van de Gevel

S4630807

Lynx Group

Prof. Ton Dijkstra

Radboud University

February 13, 2020

Created by Icons Producer
from Noun Project

# Abstract

Multilink is a computational connectionist model inspired by the Bilingual Interactive Activation + (BIA+) model and the Revised Hierarchical Model (RHM) Dijkstra, Wahl, Buytenhuijs, Van Halem, Al-Jibouri, De Korte & Rekké (2019). It is able to simulate recognition, translation, and production of Dutch and English words varying in length and usage frequency for monolinguals and bilinguals. The activation time-course of orthographic, semantic, and phonological word representations in multiple tasks can be mimicked. In this project, we will assess the role of top-down inhibition from language nodes to words by varying relevant parameters in Multilink. By doing so we can investigate whether or not word recognition can be accelerated when we are aware of language information, knowing which language a target word belongs to. This amounts, in fact, to testing BIA+ against BIA, because the existence of language-to-word inhibition is assumed by BIA, but not by BIA+. Cycle times obtained from Multilink simulations are correlated with reaction times (RTs) from multiple datasets: the English Lexicon Project (ELP), Dutch Lexicon Project (DLP), and a dataset for English lexical decision collected by Mulder et al. (2018). Simulations allow us to assess whether there is any evidence for top-down effects from one language on the words of another language. For a broader understanding of processing mechanisms, lateral inhibition between word nodes, bottom-up facilitation from word-to-language nodes, second language proficiency, and neighborhood effects will also be investigated. To obtain well-fitting simulations, the contributions of model architecture, parameters, and lexical content are analyzed and adapted. No significant evidence had been found for top-down language-to-word inhibition. These results are in accordance with the BIA+ model.

# 1. Introduction

Bilingualism and even multilingualism are prominent in the world. In fact, according to a European inventory, the Eurobarometer, in 2012, the percentage of bilingual speakers was 75% of the Dutch population, and 34% being multilingual, speaking at least three languages (Special Eurobarometer 386, 2012). Speakers are called ‘bilingual’ if they are able to use two languages fluently. Bilingualism is, in fact, a common phenomenon across the world and is becoming more and more frequent: At least one in three people are assessed to be bilingual or even multilingual (Wei, 2000). Becoming proficient in multiple languages yields important advantages, both financially, culturally, and cognitively. Bilingual speakers have more job opportunities available to them, and knowing a second language allows for an immersion into foreign cultures and unfiltered access to their history, arts, science, and literature. A wide range of economic and social advantages, or ‘communicative benefits’, is also covered (Selten & Pool, 1991).

Not only from a societal but also from a personal perspective, bilingualism opens many doors. It not only enables us to obtain much more knowledge than when restricted to only the native language but it also facilitates the communication between people of different nationalities. One interesting bilingual interaction that I was involved in happened when at one time I was ordering coffee in Dutch. I was busy texting in English while the coffee was being prepared and served. This resulted in me saying ‘thank you’ in English, instead of the Dutch ‘dankjewel’ when being handed the coffee. This took me by surprise, causing me to be lost for words. This interaction stimulated me to find out more about language switching and its underlying mechanisms, bringing about the following research question. Does the use of one language suppress the use of another? In order to answer this question, we will first be looking at the different models for word recognition and the underlying processes like word retrieval and cognitive control. Word retrieval is the ability to retrieve a word from a personal lexicon. Cognitive control is the process by which goals or plans influence our behavior, allowing us to choose between languages. With the knowledge of these concepts, we can continue our investigation with simulation.

## 1.1 Models for word retrieval

There are many approaches to describing word retrieval, for instance, by means of locally or distributed connectionist models. A connectionist model consists of a network of interconnected nodes. The information given to the nodes and the strength between their connections comprise the knowledge of a network (Rast, 2008). In localist connectionist models, information is represented in terms of separate pieces, for example, visual information as letters in a word. In distributed models, information is presented as a pattern of activation spread over multiple processing units all contributing to different representations. Localist and distributed models differ in the way input is processed at later stages of representation, not in the way the input patterns are being represented. However, localist models frequently use both localist and distributed representations. Layers in a model can show both local and distributed ways of representation of words or letters. Thus, it makes no sense to say that this localist model does not have distributed representations (Page, 2000). It is better to define a model on the presence or absence of localist representations.

A model can be non-implemented or implemented. For an implemented model, called a ‘computational model’, a software program has been made that captures the mechanisms proposed to underlie the cognitive function at hand (in our case, word recognition in monolinguals and bilinguals).

An example of an implemented model of the localist connectionist type is the Interactive Activation (IA) model (McClelland & Rumelhart, 1981, 1982). The IA model distinguishes successive processing layers of representations and processing. In the first layer, visual features are detected in the input by the determination of line segments in various orientations. In the second layer, nodes respond to letters in various word positions. In the third layer, nodes respond to individual familiar words. Nodes in the same layer inhibit each other, while nodes in different layers may activate each other through bottom-up activation and top-down feedback.

Three implemented bilingual models of the localist connectionist type are the Bilingual Interactive Activation (BIA) model (Dijkstra & van Heuven, 1998), its successor the BIA+ model (Dijkstra & van Heuven, 2002) and Multilink (Dijkstra, Wahl, Buytenhuijs, Van Halem, Al-Jibouri, De Korte, Rekké, 2019). These models will be discussed in the following sections.

## (1) **The Bilingual Interactive Activation (BIA) model**

The BIA model globally has the same architecture and parameter settings as the monolingual Interactive Activation model (McClelland & Rumelhart, 1981, 1982). However, apart from introducing an integrated Dutch and English lexicon, a major extension is the addition of a representational layer containing two language nodes that are connected to all the word nodes in both lexicons. Language nodes accumulate the activation of all words from one lexicon and suppress all words in the other. This implements a top-down language-to-word inhibition mechanism. To obtain a full picture of the effects and the interaction of the language level and the word level of BIA, we will also investigate bottom-up word-to-language facilitation and lateral word-to-word inhibition in comparison to language-to-word inhibition.

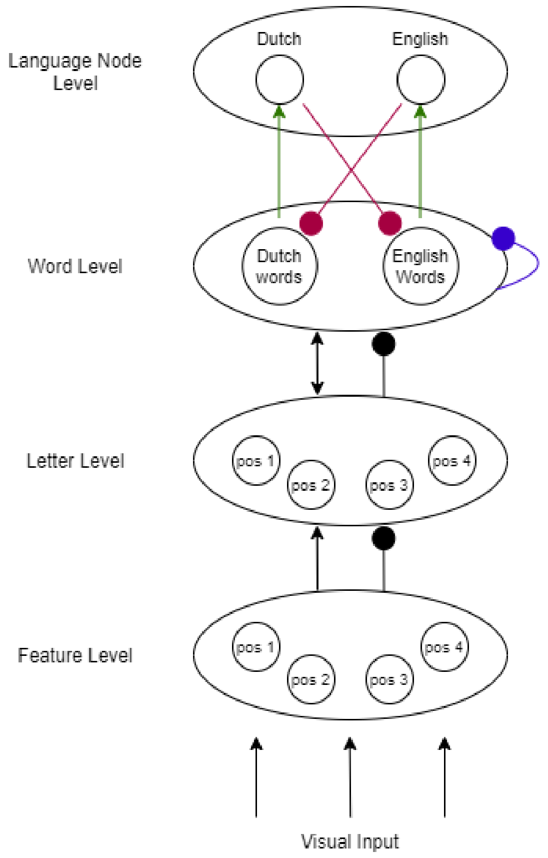

*Figure 1: The Bilingual Interactive Activation model. Facilitatory connections as arrows, inhibitory connections as dots. We are interested in the word and language node levels. Top-down language to word inhibition is indicated in red, bottom-up word to language facilitation in green, and lateral word to word inhibition in purple.*

## (2) **The Bilingual Interactive Activation Plus model**

BIA+ extends BIA by adding phonological and semantic lexical representations to the available orthographic ones. It also assigns a different role to the language nodes, rejecting the idea of top-down language-to-word inhibition. Furthermore, it makes a distinction between the effects of non-linguistic and linguistic context, based on the difference between the word system and a task/decision system regulating control. A major difference between the two models lies in their view on the role of the language nodes. BIA accepts the top-down language-to-word inhibition, while BIA+ rejects it. This can be seen when comparing the BIA (Figure 1) and BIA+ (Figure 2) architecture; while the models have similarities, top-down inhibition from the language nodes towards the word nodes is only present in BIA, not in BIA+.

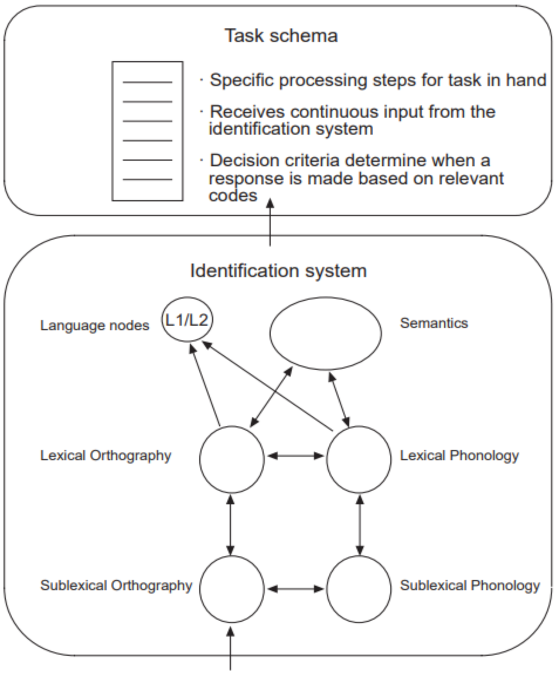

*Figure 2: BIA+ model for bilingual word recognition. Facilitatory connections as arrows, represent activation between pools, Inhibitory connections are omitted.*

## (3) **The Multilink model**

Multilink is a localist-connectionist model derived from BIA+ and RHM (Kroll & Stewart, 1994). It integrates bilingual word recognition as seen in the framework of the computational BIA+ model and bilingual word production and translation as seen by the verbal RHM. It is able to simulate the recognition, translation, and production of cognates and non-cognates of different lengths and frequencies for seminal aspects of bilingual visual word recognition, lexical-semantic processing, word naming, and word translation. It considers effects of lexical similarity, cognate status, relative L2-proficiency, and translation direction. Model-to-model comparisons show that Multilink provides higher correlations with empirical data than both the IA and BIA+ models do. Multilink shows a clear relation with the BIA and BIA+ models: nodes representing words are linked to the language nodes. In BIA+ and Multilink these representations are split into phonology, orthography, and semantics. The activation of these representations can be checked by the task/decision system used by Multilink. This system is responsible for selecting output representations, setting parameters, specifying responses, and checking language membership based on the presented stimuli and the task at hand. The task/decision system operates like cognitive control, being able to influence the behavior based on current plans or goals.

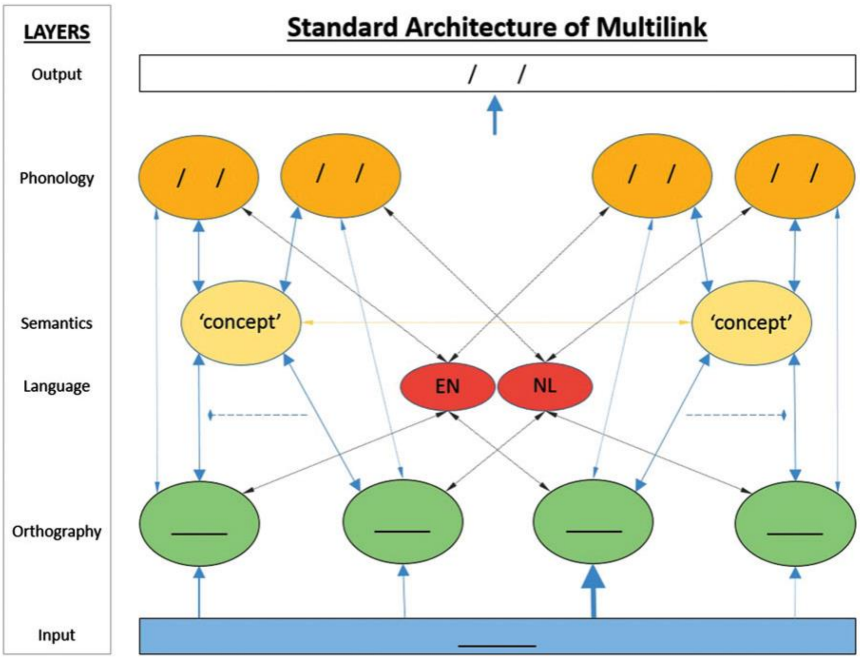

*Figure 3: Standard Architecture of Multilink. The input is shown as blue underscore. Orthographic representations as green underscore. Phonological representations as slashes. English is indicated by EN, Dutch by NL. Output is task-dependent, slashes for phonology, underscore for orthography.*

## 1.2 Cognitive control and language selection

Bilingual word processing is not just about finding a word in a lexicon containing more than one language; it is also about handling complications that arise in bilinguals because they have multiple languages at their disposal. Important aspects of such cognitive control are, for instance, how they should decide between activated representations from different languages, how they can shift from one language to another, and how they should suppress information from an irrelevant language (Miyake, 2000). In this respect, the linguistic experience of bilingualism has consequences for nonverbal cognitive performance, advantages for executive control contribute to cognitive reserve, allowing for a better coping mechanism for Alzheimer’s disease and postpone its symptoms (Bialystok, 2011). These advantages are ascribed to better cognitive control in bilinguals. However, bilinguals show a disadvantage when performing lexical retrieval tasks, experiencing more tip-of-the-tongue states when compared to monolinguals (Pelham & Abrams, 2014). Because of the prevalence of bilingualism and its positive and negative effects, further research in bilingualism is of great importance.

Bilinguals have the ability to switch between their acquired languages. This is called language switching or code-switching. This topic has become a major focus of attention in linguistics for multiple reasons (Gardner-Chloros, 2009). Language switching can further our understanding of language comprehension and production. By analyzing speech containing code-switching, we can find out to what extent and in what way words or morphemes from different languages can be combined. The manner in which bilinguals of a particular community combine their languages is a way of expressing their social identity. Their languages and sociolinguistic environment are influencing what patterns will emerge. By comparing code-switching across different communities and language combinations, the relative role of linguistic and sociolinguistic factors can be revealed, which is an important matter in Linguistics. Characteristic code-switching patterns can be used to identify subgroups within a society. Therefore, code-switching can help us understand identity formation and expression in bilinguals (Gardner-Chloros, 2009). It can take place intentionally, but also by accident, as in my case when ordering coffee. Language switching is a form of cognitive control, in linguistic models it is often prescribed to the language nodes.

## 1.3 Cognitive control and language nodes

The concept of the language node enables us to represent multiple aspects of bilingual word recognition, structurally and dynamically (Dijkstra & Van Heuven, 1998). In this way, the language node is similar to the idea of a language tag or language label we can retrieve when a word is recognized. The concept of the language switch, as a possibility to account for language-specific access, also has similarities to the language nodes. However, the language node has a more continuous, dynamic and context-driven way of representing the prominence of multiple lexicons relative to each other. Within BIA, the language node represents all factors influencing bilingual word recognition not yet specified. Context factors can influence top-down inhibition towards another lexicon. This enables us to see the effects of priming in orthographically similar words, or words of the same language. Finally, it is also assumed that the language nodes are involved in experimental task performance.

The existence and nature of language nodes is much debated. Dijkstra and Van Heuven (2002) argue that representational and functional aspects are confounded in BIA's complex use of the language nodes. In the BIA+ model, these aspects are therefore separated. The language nodes are considered as *tags* and can pick up global lexical activation, but the *functional* application of language information and (non-linguistic and linguistic) context is allocated to a task-decision system.

Language nodes are also incorporated in the Multilingual Interactive Activation (MIA) model (Dijkstra, 2003), a multilingual version of BIA. Word candidates are activated by the input feed activation from their particular language node. In this model the language node sends top-down inhibition to all activated lexical candidates from the other language. This happens when one language node becomes more activated than another. In accordance with this idea, recognition of interlingual homographs could be faster when using this activation. Dijkstra, Timmermans, and Schriefers (2000) provide evidence against this idea. Their first experiment involved a language decision task in which the participant had to press a button when an English word was presented, and another whenever a Dutch word was presented. Participants in the second and third experiment only reacted when one target language was presented in a go/no-go task, they did not respond to a non-target language. In all three experiments homographs elicited inhibition effects, these findings were similar to those found in Dijkstra et al. (1998). In the Dutch go/no-go task, the bilingual Dutch-English participants showed that even when performing the task in their native language, the non-target language still had an effect on interlingual homograph identification. Thus, the participants were not aided in the word selection process by the language membership information from the target word. They also found that these words could be overlooked, especially when the activation of the other language was high. The recognition of Dutch words was delayed when the opposite high-frequency English word was detected a recognized first, implying that language membership information only comes available after recognition of a word.

The Multilingual Processing Model (de Bot, 2004), a model for word production, assumes three stores with information: a store with conceptual features, one with syntactic procedures and another for form elements. Each of these stores has subsets of words specific to a language with overlap reflecting cognates. The processing components are controlled by the associated language node. Whenever a specific language is activated, the language node connected to this language will activate other components. This higher activation causes the element from the right language to be selected. Partial activation of a language activates elements from that particular language at other relevant levels. The language node transmits language selection information from the conceptual level to the lower level and also between lower-level components. By accumulating information about the activation of languages it can compare the intended language with the used language.

Language membership information can be used by cognitive processing. The BIA+ model assumes that language membership information is processed too late to affect word identification. However, this has been disputed by recent research. Hoversten, Brothers, Swaab, and Traxler (2015) found evidence that language membership information can adjust the processing depth for top-down tasks. Casaponsa and Duñabeitia (2014, 2015, and 2016) found that orthographic markedness from one language can be constrained by the activation of another. In another study, Hoversten, Brothers, Swaab, and Traxler (2017) found that orthographic language membership information can be used early in word recognition processes. In their experiment, the time course of information regarding the activation of orthographic language membership was investigated by means of ERPs. The authors found that the information was available early, prior to lexical access. This enables early language membership identification, language change detection, and cross-language activation modulation. These studies demonstrate that early activation of the language node is possible during orthographic decoding.

Grainger, Declerck, and Marzouki (2017) performed two generalized lexical decision experiments with French-English bilinguals with mixed sets of French and English words and pseudo-words. In the first experiment, they superimposed each word or pseudo-word on a picture of a French or UK flag, while also manipulating flag-word congruency. Lexical decisions to word stimuli were faster when following the congruent flag relative to following the incongruent flag. However, the flag was not informative with respect to the lexical decision response or the language of the word. The effect was only seen for the French words, French being the primary language. In their second experiment, they replicated the flag-language congruency effect in a priming paradigm. The word and pseudo-words targets were primed by a brief representation of the flag. In this experiment, the effects were seen in both languages. They take these findings as evidence for a mechanism that processes linguistic and nonlinguistic information regarding the presence or absence of a given language automatically. From this follows that language membership information can modulate lexical processing, which is in line with BIA, but not BIA+.

Dijkstra and van Heuven (2019) argue that the findings in this behavioral context do not provide strong evidence against BIA+ or Multilink. They argue that all that the blue-white-red flag might be activating is the concept of French, but they are not convinced that this concept is identical to the French language membership representation of a word. Also, if beneficial to the performance of the task, the flag information or language membership itself might be used as a response cue by the task/decision system. They do not favor the assumption made in BIA about the language membership label of a word being linked to all other words in all other languages to make inhibition possible. Because of the lack of empirical evidence, they argue that future work should investigate the facilitatory or inhibitory roles of language membership representations in word recognition in isolation and in sentences.

Language nodes enable us to represent multiple aspects of bilingual word recognition, structurally and dynamically. Their existence and nature are much debated, many models have arisen with different integrations of the language nodes. Recent research by Hoversten et al. (2015, 2017) shows language membership information can be used by cognitive processing, disputing the BIA+ model assumption that language membership information is processed too late to affect word identification. Grainger et al. found evidence for a mechanism that processes linguistic and nonlinguistic information regarding the presence or absence of a given language automatically; Dijkstra and van Heuven argue against these findings. All in all, there are a lot of different views on the language nodes, language membership, and language-to-word inhibition. In the next section, simulations will be discussed, subjugating the varying opinions to further analysis.

## 2. Simulations with Multilink

Within Multilink we are able to simulate a lexical decision task. In a lexical decision task, a participant needs to decide whether or not the presented combination of letters is a real word. For example, the word ‘lynx’ can be presented, to which we respond with a positive answer, it is a real word. If a non-word like ‘xnyl’ is presented, we respond negatively. We want to know if this task can be influenced by language information, knowing which language a target word belongs to.

Our investigation will be embodied by a series of simulations. The first will be about lateral word-to-word inhibition. This inhibition can be within and across languages. A monolingual speaker is able to experience only word-to-word inhibition within his native language. Additionally, a bilingual speaker also experiences this inhibition across its known languages. In the second series of simulations we will be looking at effects of top-down language-to-word inhibition for bilinguals of two different levels of secondary language proficiency, that is, balanced and unbalanced bilinguals. Top-down language-to-word inhibition is an across-languages phenomenon. One language suppressing the use of another, from the language level towards the word level. In the third simulation, neighborhood effects will be investigated in regards to language-to-word inhibition, words are neighbors when they only differ in one letter. We can use this data to investigate the effect of language-to-word inhibition of different neighborhood subgroups, determined by the presence or absence of neighbors for a particular word. Next, we will look at the effect of language-to-word inhibition on bottom-up word-to-language facilitation, a within language process in which the use of words of a language facilitate the use of that same language. Lastly, the effect of language-to-word inhibition on the aforementioned word-to-word inhibition will be taken into consideration.

The facilitatory and inhibitory processes discussed above interact and are part of the language level and word level of the BIA model as seen in Figure 1. Investigation towards these processes and their mutual interaction is of importance for understanding word recognition in bilinguals. These manipulations are possible within Multilink by adapting certain parameters. These parameter settings can correspond to BIA or BIA+; by adjusting them we can simulate both models within Multilink.

## 2.1 Specific parameter settings

By manipulating the top-down language-to-word parameters, the LO$_{\gamma}$ (Language-Orthography $\gamma$) and LP$_{\gamma}$ (Language-Phonology $\gamma$), we can effectively test the BIA and BIA+ architectures by means of Multilink. When using a LO$_{\gamma}$ and LP$_{\gamma}$ value of 0.0, we are simulating BIA+. In contrast, if we choose a value between 0 and -1, we are working with a BIA architecture. We choose to investigate the following negative values of language-to-word inhibition: -0.001, -0.005, -0.01, -0.015, and -0.02. We chose these values because preliminary testing shows a reliable effect can only be seen at low values. The value -0.001 being very low, and -0.02 being the last usable value, which after the model starts to rapidly decline in functionality. Before our investigation towards top-down language-to-word inhibition, we will establish a baseline by looking at lateral word-to-word inhibition for all datasets. We will differentiate between word-to-word inhibition, as OO$_{\gamma}$ (Orthography-Orthography $\gamma$), values of 0.0 (turned off), high inhibition of -0.001 and low inhibition of -0.0001.

Multiple other parameters are of interest for the investigation of top-down language-to-word inhibition, bottom-up word-to-language facilitation, and lateral word-to-word inhibition. OL$_{\alpha}$ (Orthography-Language $\alpha$) and PL$_{\alpha}$ (Phonology-Language $\alpha$) are the parameters causing bottom-up word-to-language facilitation. OO$_{\gamma}$ causes lateral word-to-word inhibition. During the language-to-word trials we manipulate LO$_{\gamma}$ and LP$_{\gamma}$, we set the OL$_{\alpha}$ and

$PL_\alpha$ value to 0.6 and $OO_\gamma$ to -0.0001. The obtained data will also be used for the calculation of the cycle times and the neighborhood effects. During word-to-language trials, we manipulate $LO_\gamma$, $LP_\gamma$ as previously explained, $OO_\gamma$ is set to -0.0001. $OL_\alpha$ and $PL_\alpha$ will take the values of 0.2, 0.4, 0.6, 0.8, and 1.0, giving a wide range of parameter space. During the word-to-word trials, we manipulate $LO_\gamma$ and $LP_\gamma$ the same way, and $OO_\gamma$ will take the same values as tested before. $OL_\alpha$ and $PL_\alpha$ are set to the optimal value found in the word-to-language facilitation trials, 0.2 for the DLP, 1.0 for Mulder et al. For clarity, see the table below.

| | **$LO_\gamma$** | **$LP_\gamma$** | **$OL_\alpha$** | **$PL_\alpha$** | **$OO_\gamma$** |
|---|---|---|---|---|---|
| **Word-to-word inhibition** | | | | | 0.0,-0.001,-0.0001 |
| **Language-to-word inhibition** | 0.0,…,-0,02 | 0.0,…,-0,02 | | | |
| **Proficiency effects** | 0.0,…,-0,02 | 0.0,…,-0,02 | | | |
| **Neighborhood effects** | 0.0,…,-0,02 | 0.0,…,-0,02 | | | |
| **Word-to-language facilitation** | 0.0,…,-0,02 | 0.0,…,-0,02 | 0.2,…,1.0 | 0.2,…,1.0 | |
| **Word-to-word inhibition** | 0.0,…,-0,02 | 0.0,…,-0,02 | | | 0.0,-0.1,-0.0001 |

*Table 1: The relevant parameters (horizontally) for each trial (vertically). $LO_\gamma$ & $LP_\gamma$ are responsible for language-to-word inhibition, $OL_\alpha$ & $PL_\alpha$ apply word-to-language facilitation, and $OO_\gamma$ represents word-to-word inhibition.*

## 2.2 Relevant datasets

To answer our research question, we can correlate the obtained cycle times from the Multilink simulations to actual reaction times from previously done lexical decision task experiments. If the correlations for BIA with language inhibition are higher than for BIA+, which has no language-to-word inhibition, the assumption made by BIA is more likely to be true. If the correlation is lower, the assumption of BIA+ is more likely to be true. For the Multilink simulations, we will use a lexicon representing Dutch bilinguals, with their primary language being Dutch and their secondary English. To simulate the lower subjective second language frequency of English in Dutch-English speakers, frequency is adapted to be a fourth of a native English speaker. We will be using the dataset including words and reaction times from the Dutch Lexicon Project (DLP) (Keuleers, Diependaele & Brysbaert, 2010), Mulder, van Heuven, & Dijkstra (2018), and the English Lexicon Project (ELP) (Balota, 2007). The DLP and Mulder et al. dataset are chosen because they represent Dutch-English bilinguals. The ELP represents English monolinguals. It allows us to create a baseline for inhibitory processes by comparing lateral inhibition effects between monolinguals and bilinguals.

The English Lexicon Project provides a standardized behavioral and descriptive dataset of 40481 words and 40481 non-words (Balota, 2007). Across six universities, data from 816 native English participants were collected in a lexical decision task and data from 444 other native speaking English speakers were collected in a speeded naming task. Because the ELP is a monolingual dataset, we can use it to set up a monolingual English baseline in word-to-word inhibition trials.

The Dutch Lexicon Project consists of a database of Dutch words and reaction times delivered by Dutch-English bilingual speakers (Keuleers, Diependaele & Brysbaert, 2010). A group of 39 participants made lexical decisions to 14,000 words and 14,000 non-words. The DLP aims at making a broad range of word recognition data available, to allow researchers to run regression analyses over the entire range of a variable. It also allows them to run virtual experiments to be able to test hypotheses quickly. Different from other large word recognition studies using many participants who responded to only a small part of the stimuli, the DLP had the participants respond to all stimuli. This enabled new ways of analysis, for instance, being able to compare individual differences, while also increasing the strength of the analysis. Their data represents an important collection of very long series of individual reaction times that is of interest for our own research.

Mulder et al. (2018) conducted neighborhood experiments with bilinguals speaking Dutch and English testing the effects of secondary language (L2) proficiency and neighborhood characteristics within and between languages. The L2 proficiency of Dutch bilinguals has increased over the last 20 years. To consider this impact, Mulder et al.

conducted a replication of the ELD task by van Heuven et al. (1998). In line with this hypothesis, they found that English (and not Dutch) word characteristics dominated the word and non-word responses, only the non-words revealed an interaction of English and Dutch neighborhood size. Next, they performed another ELD and a progressive demasking task, while introducing a stronger neighborhood manipulation. English words without orthographic neighbors in either language, called hermits, were contrasted with words having neighbors in either Dutch or English or having neighbors in both. For both the ELD and demasking task, target processing was strongly affected by the presence of English neighbors and only weakly by Dutch neighbors. These effects were interpreted to be a consequence of an interaction between language-specific global lexical activation and lexical competition. Language-specific global lexical activation being the activation within one particular language. The higher the global lexical activation, the more neighbors are activated. Lexical competition is a result of inhibitory links between words in the lexicon as assumed in BIA and BIA+.

# 3 Simulations

The use of inhibitory processes like lateral word-to-word inhibition or top-down language-to-word inhibition in word recognition models is under discussion; this creates the necessity for further investigation. Therefore, we will start our investigation with simulation A, looking at the effects of lateral word-to-word inhibition. In simulation series B we will look at effects of top-down language-to-word inhibition and proficiency effects. This is divided into B1 and B2, looking at the effects for unbalanced bilinguals and balanced bilinguals, respectively. These simulations specifically concerned with language-to-word inhibition are fundamental towards answering our research question regarding the existence and effect of language suppression. In simulation C, language-to-word inhibition and neighborhood effects will be investigated; by investigating these effects we can gain insight to the effect language-to-word inhibition in different neighborhood classes. In simulation series D we will look at the effect that top-down language-to-word inhibition has on other parameters. In D1 we will look at bottom-up facilitation and in D2 at lateral word-to-word inhibition. The processes of interest in simulations series D apply to the language and word levels. The investigation towards these processes and their interaction are of importance for a broad picture of both levels and our understanding for word recognition in bilinguals.

## Simulation A - Lateral inhibition between words

Lateral inhibition is an important process in the production (Sadat, Martin, Costa & Alario, 2014) and comprehension of language (BijeljacBabic, Biardeau & Grainger, 1997; Meade, Grainger, Midgley, Emmorey & Holcomb, 2018). The process entails activation and selection of word representations. Evidence for the existence of lateral inhibition between word representations has been found by BijeljacBabic, Biardeau, and Grainger (1997). They found that masked neighbor word primes can interfere with target processing and slow lexical decision tasks, this effect occurs across languages and can be seen quite early in secondary language acquisition in bilinguals (Meade, Grainger, Midgley, Emmorey & Holcomb, 2018). The use of inhibitory processes like lateral word-to-word inhibition is under discussion. Within Multilink there have been different versions including or excluding lateral inhibition. Because of the previously mentioned evidence this process deserves further investigation.

Lateral word-to-word inhibition involves inhibitory connections between activated words. It can both affect the activation of words within and across languages. For monolinguals, words are only inhibited by other words from the same language. In bilingual speakers, words are inhibited by words from the same language but also words from other known languages. The ELP dataset, being monolingual, is only affected by the within-language inhibition. Whereas the DLP and Mulder et al. datasets are also affected by cross-language inhibition.

Word-to-word inhibition effects vary for the monolingual and bilingual case. The ELP shows a general decrease in correlation for any increase in word-to-word inhibition. The DLP shows an increase in correlation for inhibition values of -0.0001 and -0.001, the Mulder et al. dataset shows an increase for -0.0001. The difference in the patterns between the monolingual and bilingual datasets suggests that lateral word-to-word inhibition is present in bilinguals. This is in accordance with research done by Meade et al. (2018), wherein it was found that competition occurs across languages. We do not see the same effect in correlation for the monolingual ELP, a possible explanation could be that word-to-word inhibition is a solely across language phenomenon.

Both lateral word-to-word and top-down language-to-word inhibition could reduce the activation of the translation between languages. However, they do so in different ways. Where word-to-word inhibition goes from word-to-word representation, language-to-word inhibition performs from the language membership representation to word representation. The latter form of inhibition will be discussed next.

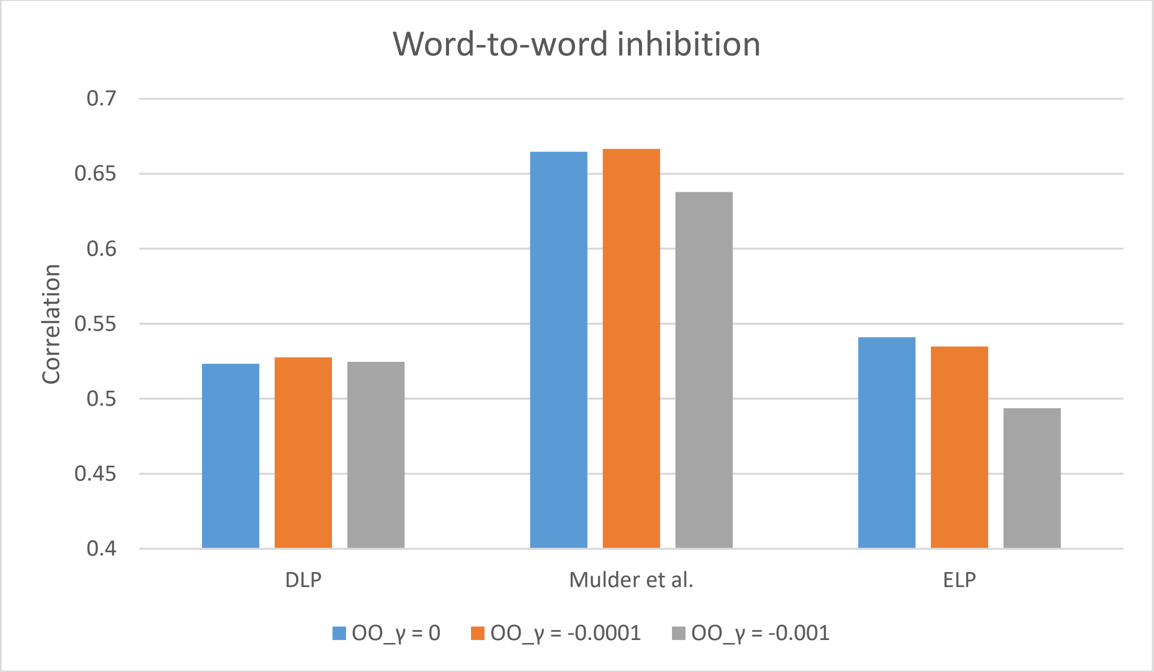

*Figure 4: Word-to-word inhibition. The ELP shows a decrease in correlation for any increase in word to word inhibition. The DLP shows an increase in correlation for inhibition values of -0.0001 and -0.001, the Mulder et al. dataset shows an increase for -0.0001. Other inhibition values show a decrease in correlation.*

## Simulation series B: Top-down language-to-word inhibition and L2 proficiency effects

The main goal of this simulation series is to look at the effects of top-down inhibition, in addition to that we can check the effects of proficiency by comparing two levels of proficiency. Top- down inhibition acts like a language filter with cross-language feedback from the language level. This filter modules lexical activity based on the activity of the language nodes. Top-down language-to-word inhibition is an across languages phenomenon, one language suppressing the use of another, from the language level towards the word level. For instance, when reading in English, a bilingual will have more difficulty reading a Dutch word compared to a bilingual reading the same Dutch word while only reading in Dutch. Following the assumption of the BIA+ model we have the null hypothesis: top-down language-to-word inhibition does not occur in bilingual word recognition. The alternative hypothesis is: top-down language-to-word occurs in bilingual word recognition.

In the framework of the present Dutch-English version of Multilink, L2 Proficiency refers to the secondary language proficiency level of a bilingual speaker, for Dutch-English bilinguals this is the English language. We can measure medium to high second language proficiency, for instance with the ‘Lexical Test for Advanced Learners of English’ also called ‘LexTALE’ developed Lemhöfer & Broersma (2012). This test estimates the receptive English world knowledge of bilingual speakers with English as a second language by having participants give a yes or no response to real English words and pseudo words.

Proficiency effects can help us find out what the effect of language-to-word inhibition is on second language proficiency. Proficiency is determined by frequency of a word, this influences the word activation which is influenced by language-to-word inhibition. Therefore, language-to-word inhibition might have a different impact on the L2 of more proficient bilinguals in which the frequencies of words are higher compared to less proficient bilinguals. In simulation B1 we will first be looking at the language-to-word inhibition effects for unbalanced bilinguals, then we will look at the effects for balanced bilinguals, comparing the unbalanced and balanced bilinguals to investigate the effects of proficiency in B2.

### Simulation B1: language-to-word inhibition for unbalanced bilinguals

The DLP and Mulder et al. studies obtained their lexical decision reaction times from Dutch-English bilinguals. These bilinguals speak their secondary language fluently, but not at the same level as their primary language. This unbalanced bilingualism has been mimicked in Multilink by using a lexicon in which English word frequency obtained from the ELP study is divided by four while Dutch words remain at the same frequency as presented in the DLP, creating an unbalanced lexicon.

Language-to-word inhibition as determined by $LO_\gamma$ and $LP_\gamma$ parameters as displayed in Figure 5 show positive and negative effects in correlation to the DLP and Mulder et al. dataset. The correlation of the DLP without gamma inhibition is 0.552. There is an increase in correlation for $\gamma = -0.005$ to 0.562 (p-value = 0.280, N = 1478) and for $\gamma = -0.01$ to 0.556 (p-value = 0.412, N = 1478), all other gamma values show a decrease in correlation. The correlation of the Mulder et al. dataset without gamma inhibition is 0.666. There is an increase in correlation for γ = -0.001 to 0.667 (p-value = 0.494, N = 101) and for γ = -0.005 to 0.667 (p-value = 0.493, N = 101), all other gamma values show a decrease in correlation.

The increase in correlation found for simulations with gamma inhibition compared to simulations without gamma inhibition are not significant. We fail to reject the null hypothesis, indicating that no significant evidence for top-down language-to-word inhibition has been found. This is in line with the BIA+ model and the results of the Dutch go/no-go task performed by Dijkstra et al. (1998). Word selection processes are not aided by language membership information from the target word.

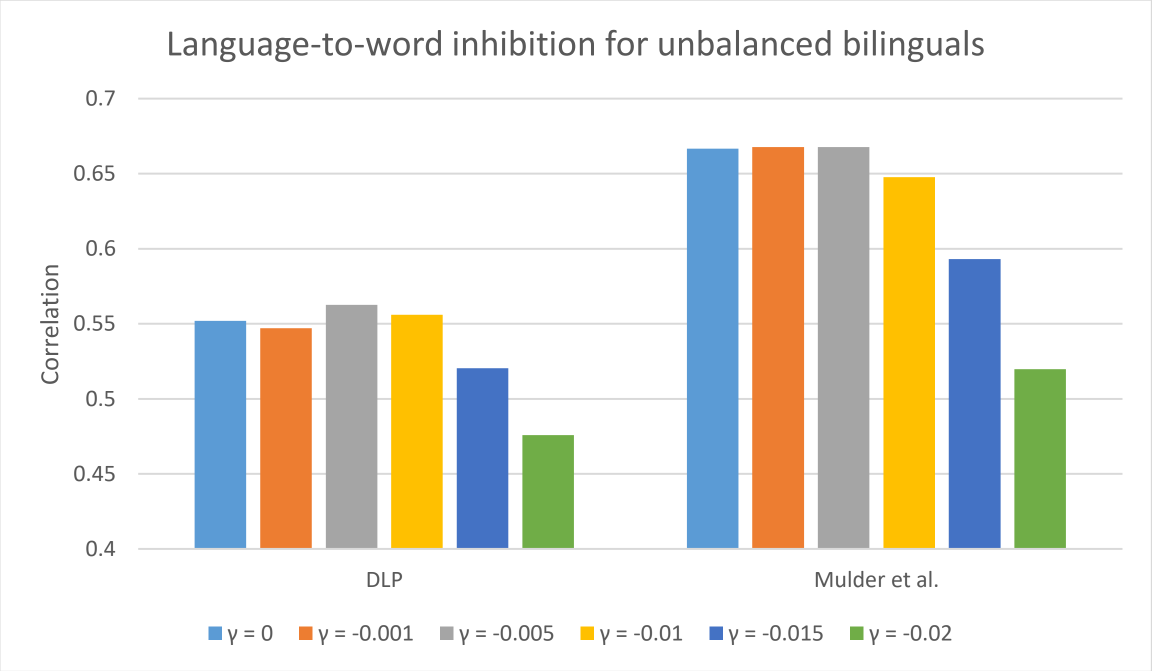

*Figure 5. Language-to-word inhibition for unbalanced bilinguals. LO_gamma & LP_gamma as γ. Correlation of the DLP without gamma inhibition is 0.552. Increase in correlation for γ = -0.001 (0.562), for γ = -0.005 (0.556) and for γ = -0.01 (0.556), all other gamma values show a decrease in correlation. Correlation of the Mulder et al. dataset without gamma inhibition is 0.666. Increase in correlation for γ = -0.001 (0.667) and for γ = -0.005 (0.667), all other gamma values show a decrease in correlation.*

### Simulation B2: language-to-word inhibition for balanced bilinguals

In this simulation we are working with a lexicon mimicking the lexicon of a balanced bilingual. A balanced bilingual has the same level of proficiency for both his primary and secondary language. In this experiment we are using the original word frequencies from the ELP and DLP, resulting in a balanced lexicon. By comparing the results from unbalanced and balanced bilinguals we can find out what effect proficiency has on language-to-word inhibition.

The results show a similar pattern as the previous experiments, but effects are reduced. Figure 6 shows a minor increase in correlation for one gamma value in the DLP and Mulder et al. dataset, the other cases show negative effects. The correlation of the DLP without gamma inhibition is 0.551. There is an increase in correlation for $\gamma = -0.005$ to 0.553 (p-value = 0.4765, N = 1478). All other gamma values show a decrease in correlation. The correlation of the Mulder et al. dataset without gamma inhibition is 0.657. All gamma values show a decrease in correlation.

These results of balanced bilinguals are similar to unbalanced bilinguals. The increase in correlation found for simulations with gamma inhibition compared to simulations without gamma inhibition are not significant. Therefore we fail to reject the null hypothesis, indicating that no significant evidence for top-down language-to-word inhibition has been found. Opting for exclusion of language-to-word inhibition in accordance with the BIA+ model.

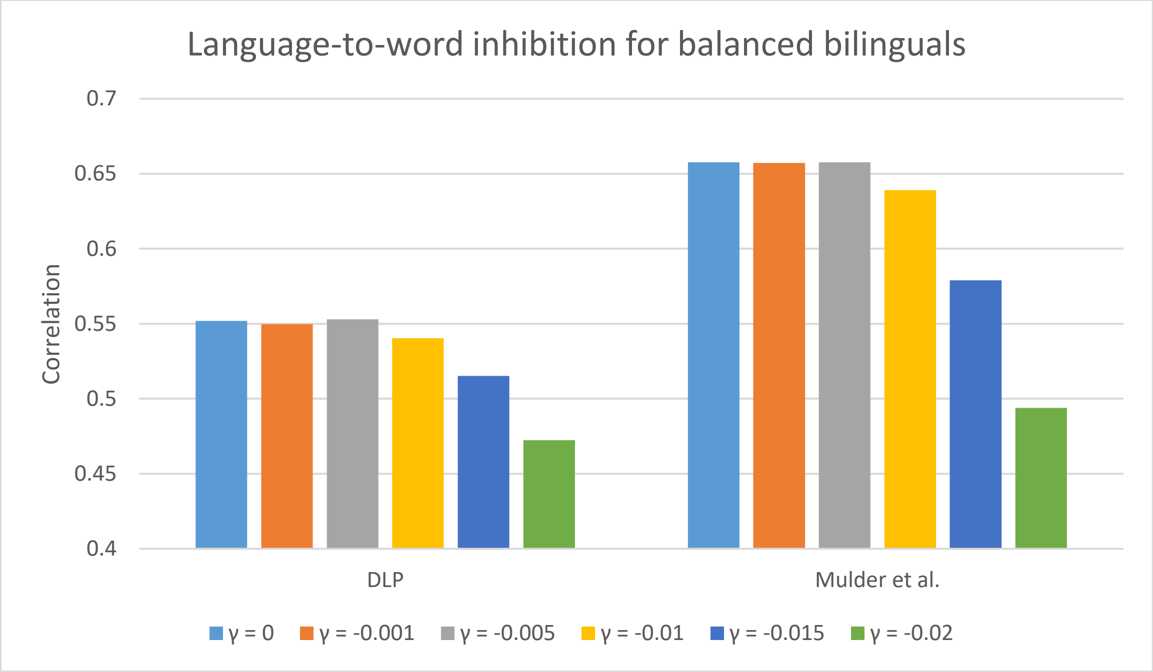

*Figure 6. Language-to-word inhibition for balanced bilinguals. LO_gamma & LP_gamma as γ. Correlation of the DLP without gamma inhibition is 0.551. Increase in correlation for γ = -0.005 (0.553) all other gamma values show a decrease in correlation. Correlation of the Mulder et al. dataset without gamma inhibition is 0.657. Increase in correlation for γ = -0.005 (0.657), all other gamma values show a decrease in correlation.*

### Simulation C: language-to-word inhibition and neighborhood effects

In the following simulation we will investigate the neighborhood effects relative to language-to-word inhibition. We can use the resulting data to investigate the effect of language-to-word inhibition in different neighborhood subgroups. The activation of words causes them to suppress other words in their neighborhood, as they are in competition. By looking at neighborhood effects we can measure the strength of competition between words. Knowing which language a particular word belongs to could influence its recognition. Using this language information pertains to the activation of language-to-word inhibition. Whenever we know which language we are using we are able to apply inhibition to the opposing language.

Words are neighbors if they are only one letter apart. A word can have neighbors within their own language, for example the English words ‘goat’ and ‘boat’. A word can also have neighbors with words from other languages, across language neighbors, for instance the Dutch word ‘snor’ and the English word ‘snow’. We can use this data to investigate different classes of neighborhood, determined by the presence or absence of neighbors for a particular word. The datasets can be divided into four subgroups. The first category being words having only Dutch neighbors like the word ‘fluid’ having a Dutch neighbor ‘fluit’ but having no neighbors in the English language. Secondly having only English neighbors, for instance ‘cloud’ having the neighbor ‘loud’, but no Dutch neighbors. Words having neighbors in both languages, like ‘bird’ having an English neighbor ‘bard’ and a Dutch neighbor ‘bord’. And lastly, hermits, words without neighbors in either language like the word ‘lynx’. Dividing the datasets shows inhibition is favored by different subgroups for each dataset.

If we assume words of one language can influence the activation of words in another language, according to the language non-selective access account, neighbors from one language can influence processing of a word in a different language (Mulder et al. 2018). This means that co-activation of lexical representations can occur within and across languages when the words are neighbors. Every word in a neighborhood from a language has a particular frequency and representational strength. The total activation of all the words in this neighborhood represents the strength of their language. By manipulating neighborhoods within and across languages we can manipulate the strength of each of the languages for particular words.

The simulations with the DLP and Mulder et al. dataset showed an increase of correlation between the cycle times and reaction times for both datasets when increasing language-to-word inhibition. This increase is likely to be caused by neighborhood effects within the datasets. As seen in figure 7, within the DLP the group with Dutch and English neighbors show an increase for $LO_\gamma$ and $LP_\gamma$ values of -0.005 and -0.01. Because this pattern of increase is similar to the increase in correlation for the whole dataset, it is likely that this subgroup is the cause for this increase. Figure 8 shows favorability towards the hermit case for the Mulder et al. dataset. $LO_\gamma$ and $LP_\gamma$ show an increase at values of -0.001, -0.005 and -0.01. The increase in correlation for the hermit case explains the increase in correlation for the overall dataset.

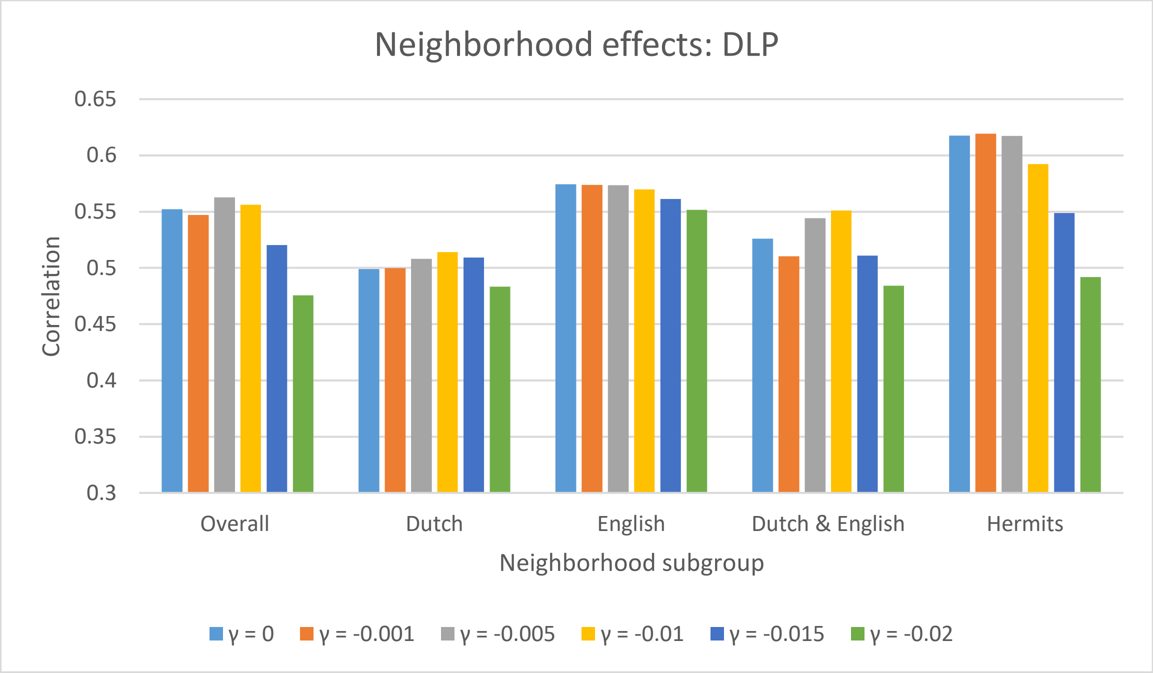

*Figure 7. Neighborhood effects: DLP. LO_gamma & LP_gamma as γ. Overall correlation of neighborhood effects and specific neighborhood effects. Words with Dutch and English neighbors show an increase in correlation for gamma values of -0.005 and -0.01.*

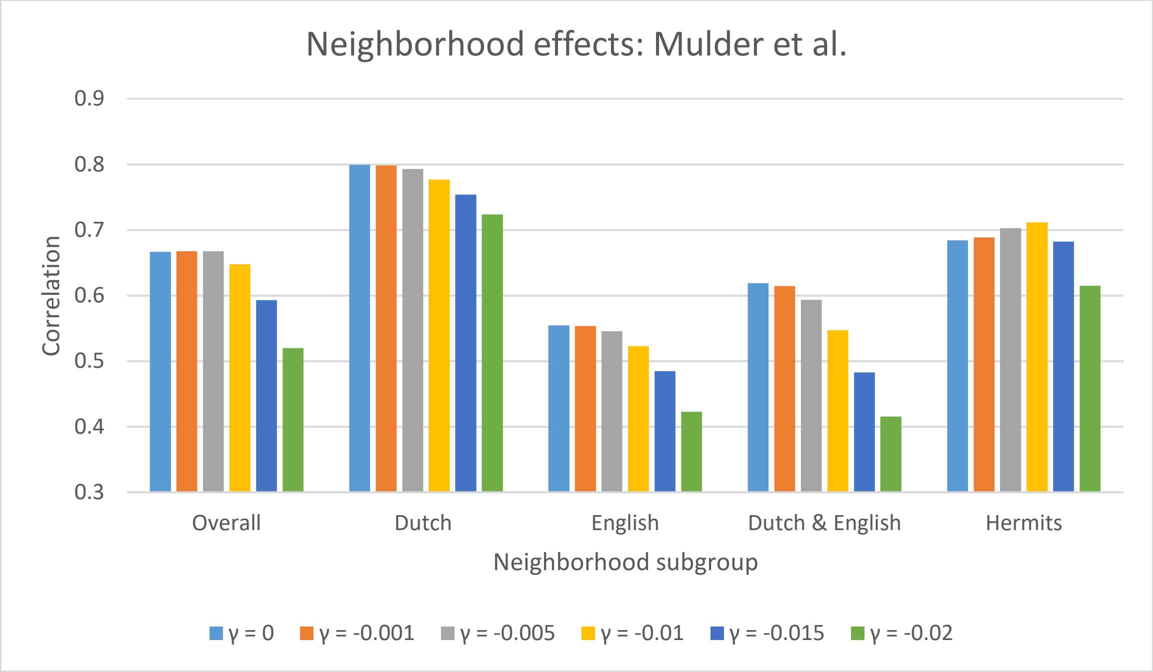

*Figure 8. Neighborhood effects: Mulder et al. LO_gamma & LP_gamma as γ. Overall correlation of neighborhood effects and specific neighborhood effects. Hermits show an increase in correlation for gamma values of -0.001, -0.005 and -0.01.*

## Simulation series D: language-to-word inhibition effects on word-to-language facilitation and word-to-word inhibition

In simulation series D we will look at the effect of language-to-word inhibition on other facilitatory and inhibitory processes in the language and word levels of word recognition. First we will look at bottom-up word-to-language facilitation in D1. Lastly, in D2, the effect of language-to-word inhibition on word-to-word inhibition will be taken into consideration. This will advance our understanding of the language and word level concerned with these processes, and the interaction of these processes themselves.

### Simulation D1: language-to-word inhibition effects on word-to-language facilitation

Bottom-up word-to-language facilitation is a within language parameter, using words in one language facilitate the use of the overall language. For example, whenever bilinguals are having a conversation in English, they expect the conversation to continue in English.

Words will be interpreted as being English more quickly. When a bilingual is switching between languages within a conversation, the Dutch ‘dat is oke’ could be interpreted as the English ‘it is’. Because they are phonologically and semantically similar, bilinguals could interpret them either way without it causing problems when integrating them in a sentence. Because of word-to-language facilitation it is more likely to be interpreted as ‘it is oke’ when a bilingual is speaking in English. We will investigate whether this facilitation influences the language-to-word inhibition concerning the same language and word nodes.

Word-to-language facilitation shows different results for the DLP and Mulder et al. dataset. Overall, language-to-word inhibition has no interesting impact on word-to-language facilitation. The DLP shows a decrease in correlation as word-to-language facilitation as $OL_\alpha$ and $PL_\alpha$ increases. It favors a low alpha value of 0.2. The Mulder et al. dataset increases in correlation as alpha increases, it favors a high alpha value of 1.0. These favorable values will be used for the final lateral word-to-word inhibition trials.

| | **$\alpha = 0.2$** | **$\alpha = 0.4$** | **$\alpha = 0.6$** | **$\alpha = 0.8$** | **$\alpha = 1.00$** |
|---|---|---|---|---|---|
| **$\gamma = 0.0$** | 0.5520496 | 0.5520496 | 0.5520496 | 0.55205 | 0.55205 |
| **$\gamma = -0.001$** | 0.5479731 | 0.5475051 | 0.5471285 | 0.546915 | 0.546868 |
| **$\gamma = -0.005$** | 0.5627836 | 0.5631699 | 0.5626791 | 0.562339 | 0.561925 |
| **$\gamma = -0.01$** | 0.5579444 | 0.5568885 | 0.5561027 | 0.555904 | 0.555709 |
| **$\gamma = -0.015$** | 0.531013 | 0.5208823 | 0.5202847 | 0.520154 | 0.520302 |
| **$\gamma = -0.02$** | 0.4898644 | 0.4775307 | 0.4757317 | 0.475844 | 0.476264 |

*Table 2. Word-to-language facilitation: DLP. $LO_\gamma$ & $LP_\gamma$ as $\gamma$, $OL_\alpha$ & $PL_\alpha$ as $\alpha$. Decrease in correlation as alpha increases.*

| | **α = 0.2** | **α = 0.4** | **α = 0.6** | **α = 0.8** | **α = 1.00** |
|---|---|---|---|---|---|
| **γ = 0.0** | 0.6666736 | 0.6666736 | 0.6666736 | 0.666674 | 0.666674 |
| **γ = -0.001** | 0.6670876 | 0.6674813 | 0.6676682 | 0.667684 | 0.667687 |
| **γ = -0.005** | 0.663482 | 0.6664355 | 0.6677894 | 0.668431 | 0.668977 |
| **γ = -0.01** | 0.6388821 | 0.6433144 | 0.6477011 | 0.65138 | 0.653162 |
| **γ = -0.015** | 0.5852365 | 0.5872322 | 0.5930139 | 0.598803 | 0.603218 |
| **γ = -0.02** | 0.5196944 | 0.5149803 | 0.5198473 | 0.527409 | 0.534371 |

*Table 3. Word-to-language facilitation: Mulder et al. $LO_\gamma$ & $LP_\gamma$ as γ, $OL_\alpha$ & $PL_\alpha$ as α. Increase in correlation as alpha increases.*

## Simulation D2: language-to-word inhibition effects on word-to-word inhibition

In the BIA model, the two mechanisms influence the word level. Top-down inhibition acts like a language filter with cross-language feedback from the language level. This filter modulates lexical activity based on the activity of the language nodes. Simultaneously, lateral inhibition suppresses items other than the target within and across languages speeding up the recognition process. The two combinations might be accountable for complex experimental results for interlingual homographs and neighbors (Dijkstra & Van Heuven, 1998). It is stated that including one or both of these inhibitory processes could be beneficial to multilink (Declerck, Meade, & Grainger, 2019). This experiment will investigate this claim.

In our experiments the activation of language-to-word inhibition and word-to-word inhibition show a covariate effect, the presence of one of the inhibitory processes influencing the other. Having no or low language-to-word inhibition as depicted by $LO_\gamma$ and $LP_\gamma$ (γ) show the highest correlation with the low value of word-to-word inhibition, $OO_\gamma$ (OO_γ). Larger language-to-word inhibition values show the highest correlation when not using word-to-word inhibition. In both the DLP (figure 10) and Mulder et al. dataset (figure 11), language-to-word inhibition values of γ = 0.0 and γ = -0.001 show an increase in correlation for word-to-word inhibition value of OO_γ = -0.000. Larger values of γ show the highest correlation with OO_ γ = 0.0. This suggests that using both inhibitory processes is not beneficial to multilink. The model seems to perform best with either one or the other and underperforms with too much inhibition from either parameter.

| | ***OO_γ = 0.0*** | ***OO_γ = -0.0001*** | ***OO_γ = -0.001*** |
|---|---|---|---|
| ***γ = 0.0*** | *0.5488683* | *0.5520496* | *0.545352* |
| ***γ = -0.001*** | *0.544444* | *0.5479731* | *0.5429954* |
| ***γ = -0.005*** | *0.5652249* | *0.5627836* | *0.5441904* |
| ***γ = -0.01*** | *0.5609112* | *0.5579444* | *0.5203177* |
| ***γ = -0.015*** | *0.5344706* | *0.531013* | *0.5077285* |
| ***γ = -0.02*** | *0.4928525* | *0.4898644* | *0.4694624* |

*Table 4. Word-to-word inhibition: DLP. $LO_\gamma$ & $LP_\gamma$ as γ, $OO_\gamma$ as OO_ γ. For γ = 0.0 and γ = -0.001, increase in correlation for OO_ γ = -0.0001. Larger values of γ prefer OO_ γ = 0.0.*

| | ***OO_γ = 0*** | ***OO_γ = -0.0001*** | ***OO_γ = -0.001*** |
|---|---|---|---|
| ***γ = 0.0*** | *0.6646146* | *0.6666736* | *0.6378482* |
| ***γ = -0.001*** | *0.666154* | *0.6676868* | *0.6366109* |
| ***γ = -0.005*** | *0.6710634* | *0.6689772* | *0.6269765* |
| ***γ = -0.01*** | *0.6593872* | *0.6531618* | *0.6032416* |
| ***γ = -0.015*** | *0.6098114* | *0.6032176* | *0.5608311* |
| ***γ = -0.02*** | *0.5395444* | *0.5343709* | *0.5000847* |

*Table 5. Word-to-word inhibition: Mulder et al. $LO_\gamma$ & $LP_\gamma$ as γ, $OO_\gamma$ as OO_ γ. For γ = 0.0 and γ = -0.001, increase in correlation for OO_ γ = -0.0001. Larger values of γ prefer OO_ γ = 0.0.*

## 4. General Discussion

In the above section we have done extensive research regarding the language and word levels of lexical processing in the Multilink model of word retrieval. We have investigated the multiple models while taking into consideration the theoretical elements relevant to these models. In particular we have looked at the BIA, BIA+ and Multilink models and their difference regarding language nodes and top-down language-to-word inhibition.

By investigating these concepts, we effectively investigated whether or not word recognition can be influenced when we are aware of language information, knowing which language a target word belongs to. After a brief summary of the simulations we are ready to answer our research question. Does the use of one language suppress the use of another?

### 4.1 Simulations

In simulation A, investigating lateral inhibition between words suggests that lateral word-to-word inhibition is present in bilinguals. This is in accordance with research done by Meade et al. (2018), which finds that competition occurs across languages. The same results were not found for the monolingual ELP; a possible explanation could be that word-to-word inhibition is a solely across-language phenomenon, but this is open for further investigation.

Simulations B1 and B2, we looked at top-down language-to-word inhibition and L2 proficiency effects. In simulation B1 an increase in correlation for simulation with gamma inhibition compared to simulations without gamma inhibition was found. Similar results were found in B2. However, these findings are not statistically significant. Therefore, we fail to reject the null hypothesis: top-down language-to-word inhibition does not occur in bilingual word recognition. Indicating that no significant evidence for top-down language-to-word inhibition has been found. Thus, according to these results language information does not influence word recognition. This is in line with the BIA+ model and in accordance with the

Dutch go/no-go task by Dijkstra et al. (1998) in which the participants were not aided in the process of word selection by the language information from the target word.

In simulation C we investigated language-to-word inhibition and neighborhood effects. Looking at different subgroups classified by a word having neighbors in either the Dutch or English language, having neighbors in both, or having no neighbors at all, called hermits. We found that the increase in correlation found in simulations B1 and B2 could be explained by looking at a particular neighborhood subgroup. The pattern of increase for a subgroup is very similar to the overall increase. In the case of the DLP this was the subgroup regarding words with neighbors in both languages, for the Mulder et al. dataset the hermit subgroup showed this pattern. As to why the two datasets prefer different subgroups is up for further research.

Simulations series D provided further insight into the processes within the language and word levels and their interaction. In D1 we looked at language-to-word inhibition effects on word-to-language facilitation; the only finding of interest was the difference in word-to-language favorability by the two databases. The DLP favors low facilitation while the Mulder et al. dataset favors high facilitation. This difference is of interest in future research. In D2 we examined language-to-word inhibition effects on word-to-word inhibition. The results suggest that using both inhibitory processes is not beneficial to multilink, as was hypothesized by Declerck, Meade, & Grainger (2019). The model seems to prefer either one or the other and underperforms with too much inhibition from either parameter. This finding is very interesting for further research, because the consensus about inhibitory processes is still divided.

## 4.2 Conclusion

Returning to our example that prompted our research question; ordering coffee in Dutch, while texting in English seemed to have had an effect on my language ability. By using one language, I was facilitated to keep using it; texting in English caused me to speak in English. And having used this language showed inhibition towards using the other language; speaking English caused me to be lost for words and not able to respond in Dutch. However, it is more likely to have been caused by the surprise of speaking English. In this investigation no significant evidence for top-down language-to-word inhibition was found. Thus, having language information, knowing which language is being spoken, does not enhance word recognition. These findings are in accordance with the BIA+ model, which assumes there is no top-down language-to-word inhibition. This holds for both balanced and unbalanced bilinguals. Furthermore, the increase in correlation for language-to-word inhibition could be explained by favorability of a dataset towards a neighborhood subgroup. Investigation toward the different processes regarding the language and word levels revealed that using both lateral and top-down inhibition is not beneficial to Multilink.

## References

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. Behavior Research Methods, 39, 445-459.

Bialystok, E. (2011). Reshaping the mind: the benefits of bilingualism. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *65*(4), 229.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition*, *25*(4), 447-457.

Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Frontiers in psychology*, *5*, 424.

Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2015). How do bilinguals identify the language of the words they read?. *Brain research*, *1624*, 153-166.

Casaponsa, A., & Duñabeitia, J. A. (2016). Lexical organization of language-ambiguous and language-specific words in bilinguals. *The Quarterly Journal of Experimental Psychology*, *69*(3), 589-604.

De Bot, K. (2004). The multilingual lexicon: Modelling selection and control. *International Journal of Multilingualism*, *1*(1), 17-32.

Declerck, M., Meade, G., & Grainger, J. (2019). On keeping cool: The role of inhibition in bilingual word processing. *Bilingualism: Language and Cognition*, *22*(4), 699-700.

Dijkstra, A. (Ton) & 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. Mahwah, NJ: Lawrence Erlbaum Associates

Dijkstra, T., Timmermans, M. and Schriefers, H. (2000) Cross-language effects on bilingual homograph recognition. Journal of Memory and Language 42, 445–464.

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.

Dijkstra, T., & Van Heuven, W. J. (2002). Modeling Bilingual word recognition: Past, present and future. *Bilingualism: Language and Cognition*, *5*(3), 219-224.

Dijkstra, T. (2003). Lexical processing in bilinguals and multilinguals: The word selection problem. In *The multilingual lexicon* (pp. 11-26). Springer, Dordrecht.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657-679.

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Modelling bilingual lexical processing: A research agenda and desiderabilia. *Bilingualism: Language and Cognition*, *22*(4), 703-713.

Gardner-Chloros, P. (2009). *Code-switching*. Cambridge University Press.

Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1158.

Grainger, J., Declerck, M., & Marzouki, Y. (2017). On national flags and language tags: Effects of flag-language congruency in bilingual word recognition. *Acta psychologica*, *178*, 12-17.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language?. *Bilingualism: language and cognition*, *1*(3), 213-229.

Hoversten, L. J., Brothers, T., Swaab, T. Y., & Traxler, M. J. (2015). Language membership identification precedes semantic access: Suppression during bilingual word recognition. *Journal of cognitive neuroscience, 27*(11), 2108-2116.

Hoversten, L. J., Brothers, T., Swaab, T. Y., & Traxler, M. J. (2017). Early processing of orthographic language membership information in bilingual visual word recognition: Evidence from ERPs. *Neuropsychologia*, *103*, 183-190.

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*, 174.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior research methods, 44*(2), 325-343.

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

Meade, G., Grainger, J., Midgley, K. J., Emmorey, K., & Holcomb, P. J. (2018). From sublexical facilitation to lexical competition: ERP effects of masked neighbor priming. *Brain research, 1685,* 29-41.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100.

Mulder, K., van Heuven, W. J., & Dijkstra, T. (2018). Revisiting the neighborhood: How L2 proficiency and neighborhood manipulation affect bilingual processing. *Frontiers in Psychology, 9*.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences*, *23*(4), 443-467.

Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 313.

Rast, R. (2008). *Foreign language input: Initial processing*. Multilingual Matters.

Rumelhart, D. E., & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological review, 89*(1), 60.

Sadat, J., Martin, C. D., Costa, A., & Alario, F. X. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Cognitive psychology*, *68*, 33-58.

Selten, R., & Pool, J. (1991). The distribution of foreign language skills as a game equilibrium. In *Game equilibrium models IV* (pp. 64-87). Springer, Berlin, Heidelberg.

Wei, L. (2000). Dimensions of bilingualism. *The bilingualism reader, 3*.

# Appendix

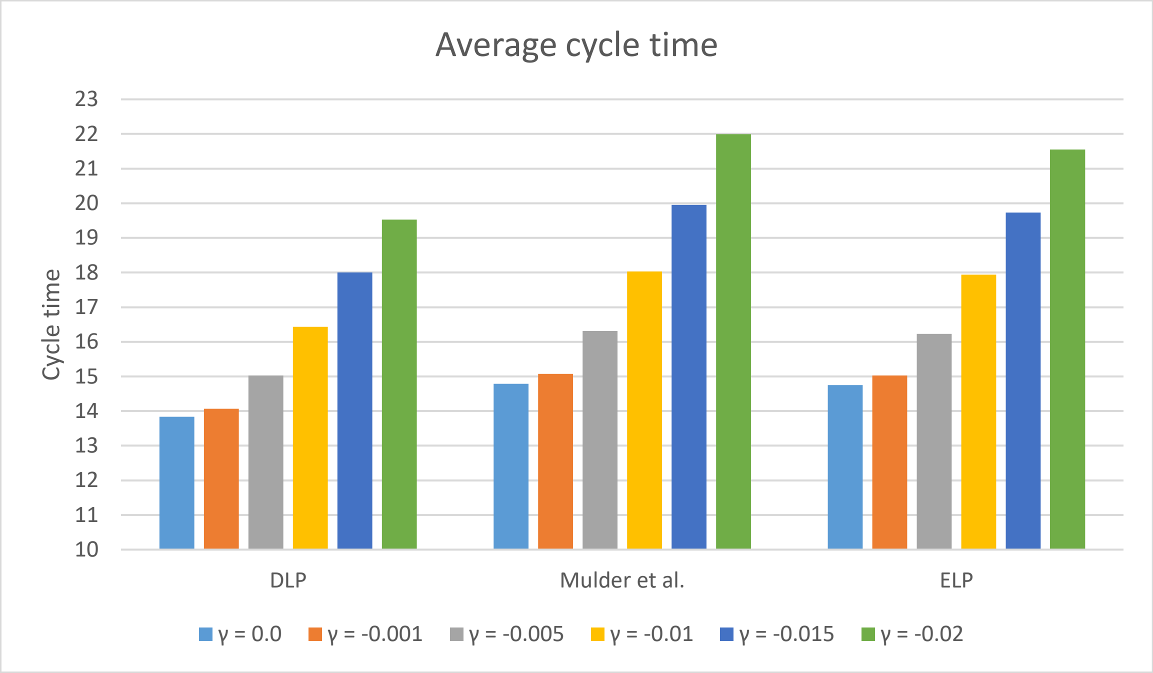

*Figure 9. Average cycle time. LO_gamma & LP_gamma as γ. Average cycle time increases for all datasets when increasing language to word inhibition.*

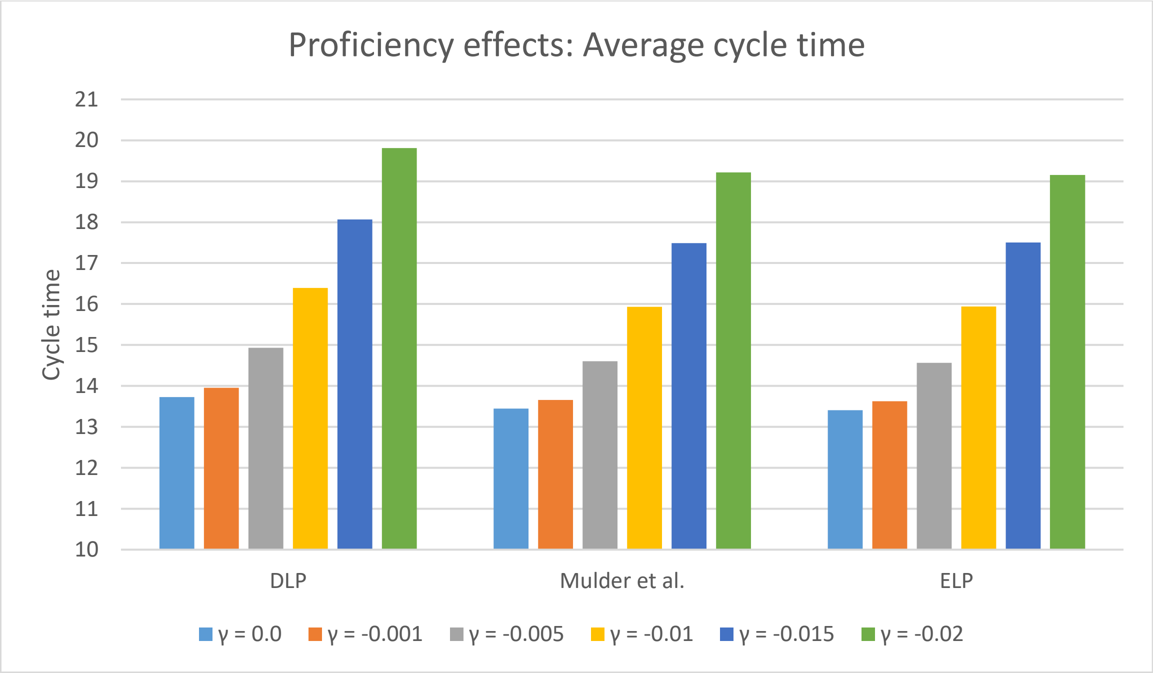

*Figure 10. Average cycle time. LO_gamma & LP_gamma as γ. Average cycle time increases for all datasets when increasing language to word inhibition.*

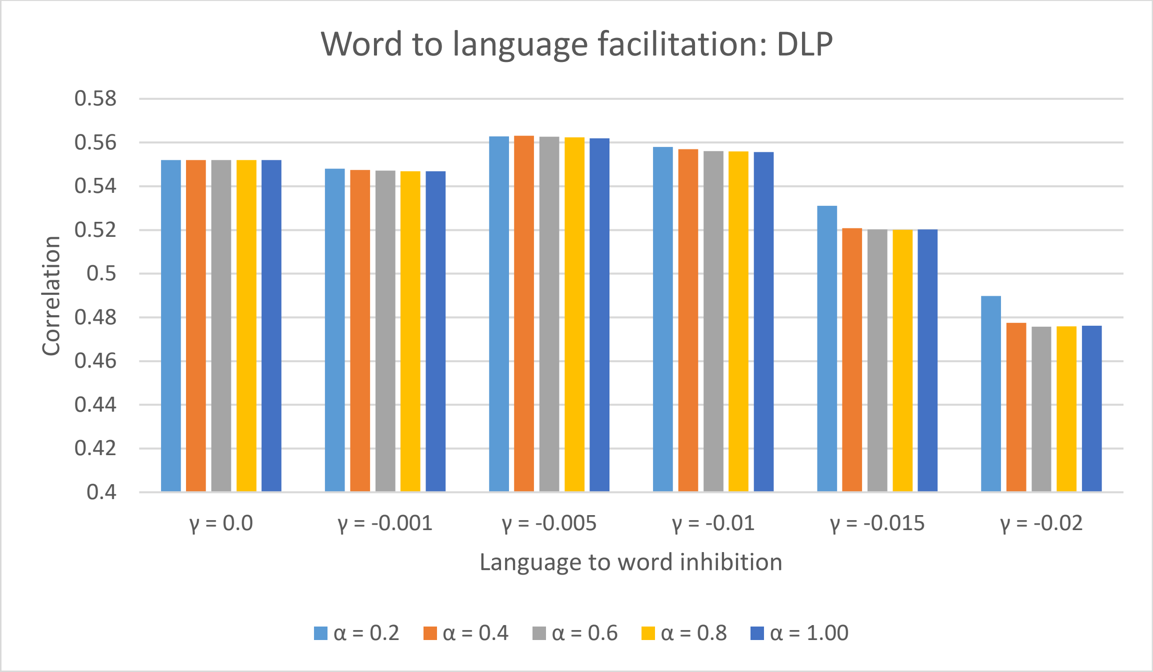

*Figure 11. Word to language facilitation: DLP. LO_gamma & LP_gamma as γ, OL_alpha & PL_alpha as α. Decrease in correlation as alpha increases.*

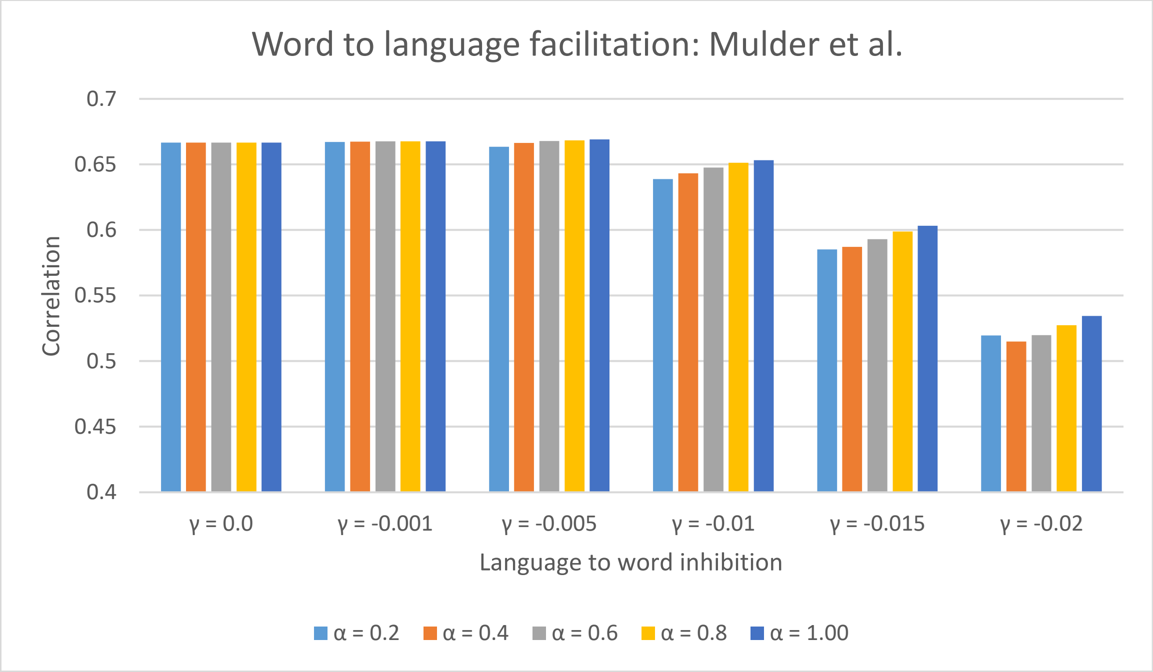

*Figure 12. Word to language facilitation: Mulder et al. LO_gamma & LP_gamma as γ, OL_alpha & PL_alpha as α. Increase in correlation as alpha increases.*

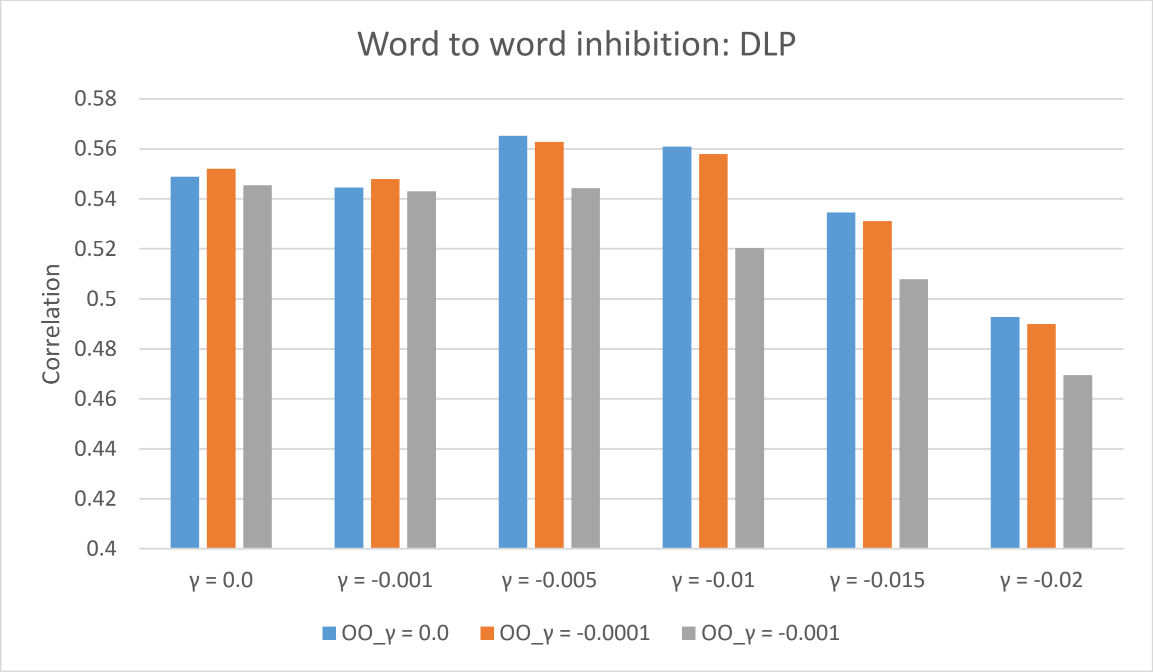

*Figure 13. Word to word inhibition: DLP. LO_gamma & LP_gamma as γ, OO_gamma as OO_ γ. For γ = 0.0 and γ = -0.001, increase in correlation for OO_ γ = -0.0001. Larger values of γ prefer OO_ γ = 0.0.*

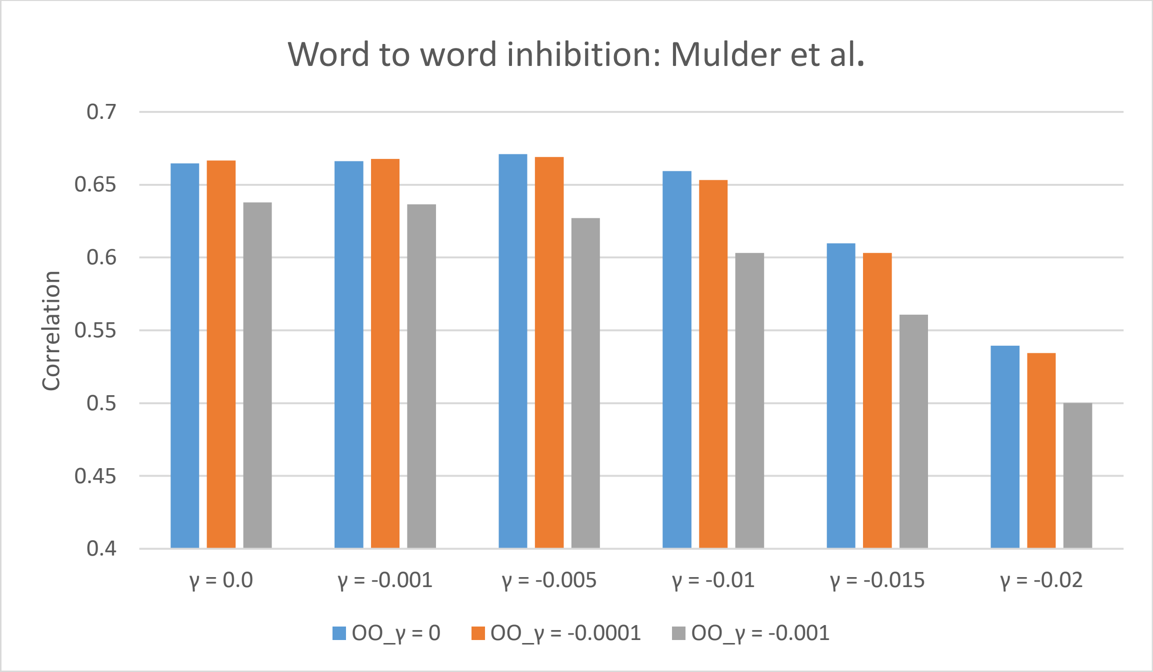

*Figure 14. Word to word inhibition: Mulder et al. LO_gamma & LP_gamma as γ, OO_gamma as OO_ γ. For γ = 0.0 and γ = -0.001, increase in correlation for OO_ γ = -0.0001. Larger values of γ prefer OO_ γ = 0.0.*