Bilingualism: Language and Cognition 22 (4), 2019, 703–713 © Cambridge University Press 2018 doi:10.1017/S1366728918000986

AUTHORS RESPONSE

Modelling bilingual lexical processing: A research agenda and desiderabilia*

TON DIJKSTRA
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
ALEXANDER WAHL
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
FRANKA BUYTENHUIJS
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
NINO VAN HALEM
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
ZINA AL-JIBOURI
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
MARCEL DE KORTE
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen
STEVEN REKKÉ
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen

(Received: August 13, 2018; accepted: August 23, 2018)

Like the BIA model (Dijkstra & van Heuven, 1998; van Heuven, Dijkstra & Grainger, 1998) and the BIA+ model (Dijkstra & van Heuven, 2002), the Multilink model is a symbolic, localist-connectionist, interactive model for lexical processing in the visual domain. In our view, the SYMBOLIC nature of Multilink makes it attractive and easily interpretable, even in relation to brain activity (Page, 2000, p. 501; 2017). Symbolic localist-connectionist models have a long tradition and have been applied to many different areas of cognitive research (e.g., Grainger & Jacobs, 1998). As a consequence, a lot is known about their properties and limitations (e.g., Bowers, 2009). These models can also easily be organized hierarchically in a cognitive functional way, and they have a reasonable degree of flexibility while still being falsifiable. Thus, despite the availability of other sophisticated frameworks for modeling language processes, a lot can still be gained from localist models.

In the following, we respond to comments on Multilink by researchers from many disciplines, including linguists, psychologists, researchers in AI and education, and others. We first position Multilink relative to BIA and BIA+, and present our plans for model development. In the core of our reply, we zoom in on reviewer comments regarding the nature of the reading bilingual or L2 learner, the scripts they have, task demands, and the organization of the bilingual lexicon. Together, these comments constitute a research agenda going considerably beyond the original intentions of the Multilink model.

* The authors thank Aaron van Geffen, Randi Goertz, Erik Lormans, and Ardi Roelofs for valuable theoretical discussions and/or help in simulations. They also thank Eva van Assche, Nicolas Dirix, Wouter Duyck, and Annette De Groot for kindly making available empirical data for simulation purposes.

Positioning Multilink in the research domain relative to BIA and BIA+

Although Multilink is also applicable to aspects of monolingual word retrieval, it puts a special focus on bilingual processing. As Goral (2018) states, there is much to be said for "the view taken in Libben, Goral, and Libben (2017) that the bilingual lexicon – not the monolingual – should be considered as the default." And as Li and Grant (2018) point out, Multilink has a considerably sized lexicon that extends the 4-letter-word lexicons of BIA / BIA+ to include 3-to-8 letter words in both English and Dutch. To allow maximal contrast with BIA and BIA+ for the evaluation of underlying mechanisms, we differentiated the present version of Multilink with respect to the earlier models in several respects. For instance, sublexical segmental representations were skipped to

Address for correspondence:
Ton Dijkstra, Donders Institute for Brain, Cognition, and Behaviour, Donders Centre for Cognition, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands
t.dijkstra@donders.ru.nl

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

allow recognition of words of different lengths, word frequency was implemented without using a rank ordering procedure (cf. Duyck, Vanderelst, Desmet & Hartsuiker, 2008), lateral inhibition was switched off, and only lexical-phonological and -semantic representations were incorporated. As such, Multilink has no 'nesting' relation to (i.e., is not embedded in) the BIA or BIA+ models (Grainger & Jacobs, 1998). However, the three models are nevertheless closely related and can be considered as members of the same general class of models.

An advantage of this cross-model similarity is that, for instance, verbal accounts of context effects and task/decision level effects formulated for BIA / BIA+ can be easily transferred to Multilink. Thus, in contrast to Van Hell (2018), we argue that Multilink does not "leave behind an important theoretical advancement that was touched upon in extending BIA to BIA+ (Dijkstra & van Heuven, 2002): how linguistic context influences word processing." In fact, the interactive mechanisms operating between semantic sentence context and lexical retrieval proposed by Dijkstra, Van Hell, and Brenders (2015), and supported by the meta-review of Lauro and Schwartz (2017), can be fully combined with Multilink. We propose that Multilink inherits these important insights from BIA+. In addition, in our paper, we demonstrated that the similarity of the three localist connectionist models allows for close model-to-model comparisons relative to shared datasets (like DLP, ELP, and diverse empirical studies). Nevertheless, as van Heuven and Wen (2018) say, it is also important to contrast different variants of Multilink itself, covering multidimensional parameter variations, in order to better understand the consequences of architectural, processing, and parameter assumptions in the various models.

Model development strategy

Reviewer suggestions for model extension and differentiation are in line with the construction and development strategy we have in mind for Multilink. We began the project by roughly following advice on modelling by Jacobs and Grainger (1994) and Grainger and Jacobs (1996, 1998). Specifically, we have implemented a network architecture based on considerations in the empirical literature, subsequently tested its processing with baseline parameter settings, and applied it to existing empirical studies. In the next development phase, we will extend the model and test its predictions in new research. In terms of parameter settings and special provisions, we are moving from a simple baseline model to more complex variants. For instance, we began by exploring only one setting of all parameters across various empirical studies, without any lateral inhibition or word association, and we simplified lexical-semantic, phonological, and language membership representations. Thus, we implemented our computational model not as an attempt in achieving a tight approximation of reality, but also (or rather) as a tool for better understanding of it. We aimed for necessary simplifications yet sufficient adequacy in generating explanatory value and new hypotheses.

In this light, we do not disagree with many points the reviewers make. For instance, we are not so foolish as to deny the importance of lateral inhibition or learning aspects, having investigated these phenomena in our own research. However, with a 'null model' for bilingual lexical processing as a baseline, anchor, or reference, we are now in a position to extend it step by step to more complex model variants.

Multilink: theoretical issues and desired extensions

Overall, the reviewers point out many attractive aspects of Multilink. As we hoped, at the current model development stage, they appear to consider RESEARCH GENERATIVITY as a more useful model evaluation criterion than GOODNESS OF FIT. Indeed, the reviewers themselves generate several new hypotheses based on the presented theoretical framework. They also deepen our insight into many theoretical issues by suggesting adaptations and extensions of the current model. Thus, they formulate many DESIDERABILIA that constitute important points of a research agenda for the bilingual domain AT LARGE.

We will discuss these in the following sections by focussing on model dimensions that stretch from more general to more specific. We will discuss how simulations could be set up for various (1) populations of bilinguals and L2 learners, (2) script types, and (3) task demands, cognitive control, and dependent variables; and (4) how the lexicon should be organized with respect to (4a) orthography, (4b) phonology, (4c) semantics, and (4d) language membership information.

(1) Populations of bilinguals and L2 learners

We built Multilink to simulate word reading and translation performance of balanced and relatively L2 proficient unbalanced adults. Frankly speaking, with respect to unbalanced bilinguals, we had a population of participants in mind that is rather standard in this vein of research: university students in psychology or linguistics. Our reviewers ask us to consider how performance might take place in heritage speakers, populations with limited literacy, and young L2 learners.

As a starting point, it seems obvious that, in order to capture individual differences in performance in simulations, we must precisely characterize the mental lexicon of each individual or group of individuals involved (but see Tokowicz, 2018, for some suggestions that can already be explored). Only when such empirical evidence

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

is available, can it be properly modelled. Li and Grant (2018) rightly point out that language proficiency cannot and should not be reduced to just differences in the size and frequency of words in the lexicon, and for L2 many aspects such as instruction method and social context, working memory (Tokowicz, 2018), and learning strategies should be taken into account (see Sianipar, Middelburg & Dijkstra, 2015; Xiang, Dediu, Roberts, Van Oort, Norris & Hagoor, 2012; Xiang, Van Leeuwen, Dediu, Roberts, Norris & Hagoort, 2015, for studies on the classroom L2 learner). We also agree with Li and Grant (2018) that an important step ahead is to characterize the L1/L2 lexicon of learners without resorting to the lexicons of native speakers of the L1 and L2 (e.g., by using L2 learners' subjective ratings). The use of bilingual rather than native norms to define the bilingual lexicon is indispensable especially when the lexicons of bilinguals are not independent.

Nevertheless, in relation to the Revised Hierarchical Model (Kroll & Stewart, 1994), Kroll, van Hell, Tokowicz, and Green (2010, p. 374) wrote that '... the assumption of functional separation of the bilingual's lexicons [...] may in fact better accommodate the wider range of evidence for parallel access in same-script, different-script, and sign-speech bilinguals. At best, the evidence is equivocal on this issue.' Although we are in favor of an integrated lexicon, we initially set the lateral inhibition parameter in Multilink to zero (see below) to see how far this notion could bring us.

As Costa and Pickering (2018) point out, when we are learning a second language, we do not ignore the first language, but use available information from the first language in a sensible interactive way to facilitate learning (Meade & Dijkstra, 2017). In the case of cognates, for instance, their definition as 'translation equivalents with form overlap' already points at L1 information sources that will affect L2 learning: their cross-linguistic overlap in semantics, orthography, and phonology (see Dijkstra, Haga, Bijsterveld & Sprinkhuizen-Kuyper, 2011, for demonstrations in a distributed connectionist framework). Multilink whole-heartedly subscribes to the assumption of mutual influences between languages during learning and processing, even when these are as yet limited in the current base-line model. At the same time, the activation of a particular word does not affect ALL other words in the same or different lexicon, so the limits of cross-talk need to be established (also see Lemhoefer, Dijkstra, Schriefers, Baayen, Grainger & Zwitserlood, 2008).

As suggested by Li and Grant (2018), and Jiang (2018), simulating lexical processing in heritage speakers also requires a close consideration of their mental lexicon architecture and processing aspects. In our own work (e.g., van Rijswijk, 2016, with Muntendam and Dijkstra), we saw some evidence that in heritage speakers the very first language (learned only one or two years earlier than the second) dominated during the formation of the L1/L2 lexicon architecture, while processing aspects were sensitive to the word frequency in the later dominant L2. Because Multilink makes a clear distinction between structure and process, this aspect could be implemented in terms of representations and parameter settings.

Even so, the important point remains that heritage speakers use some vocabulary only in some situations and not in others (e.g., Turkish in the mosque, Dutch in the street). Perhaps a context-dependent adjustment of relative language activation could be arranged via the task/decision system or via language membership nodes (e.g., Hoversten, Brothers, Swaab & Traxler, 2015).

There are some other groups of language users for whom we know even less about their lexicon architecture. Mishra (2018) points out that the mental lexicon of speakers with limited literacy has been developed without the presence of orthographic codes. In the absence of orthographic-phonological interactions, the phonological system might develop differently from that in literate bilinguals for one or both languages. Furthermore, how to characterize the lexicon of bilinguals using sign language in combination with another language is also not fully clear. Evidently, until precise information about these aspects is available, simulations are not possible. We will come back to this point later.

(2) Bilinguals reading in alphabetic and non-alphabetic scripts

Multilink can currently only perform bilingual simulations for language combinations that share a Latin alphabetic script. Goral (2018), Jiang (2018), and van Heuven and Wen (2018) raise the question of whether bilingual activation is fundamentally different between same-script and different-script bilinguals.

There is evidence that perceptual similarity between scripts is an important factor underlying co-activation in reading bilinguals. For instance, when two scripts partially overlap, words in the languages involved may be co-activated to some extent by the input (Iakovleva, Piasecki & Dijkstra, 2015; Van Kesteren, Dijkstra & De Smedt, 2012). For instance, Iakovleva et al. (2015) showed that in English word processing, Russian–English bilinguals activate lexical candidates that are similar to the input word in both Russian and English (language non-selective lexical access). English–Russian cognates were recognized more quickly than English control words (cognate facilitation effect; Dijkstra, Miwa, Brummelhuis, Sappelli & Baayen, 2010). Interestingly, cognates with ambiguous orthography, i.e., shared letters mapping onto different phonemes in the two languages, were processed more slowly than cognates with mismatching orthography, due to decreased facilitation deriving from the other cognate member.

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

There is also evidence that even for reading bilinguals with non-overlapping scripts, phonological and semantic representations of both associated languages can still be co-activated at some time during lexical access (e.g., Gollan, Forster & Frost, 1997, for Hebrew–English; Kim & Davis, 2003, for Korean–English; Nakayama, Sears, Hino & Lupker, 2012, for Japanese–English; Wen & van Heuven, 2018, for Chinese–English). For instance, in line with the study by Kim (2017) that Jiang (2018) refers to, Miwa, Dijkstra, Bolger, and Baayen (2014) discussed an extension of the BIA+ model that explains cognate effects in Japanese–English bilinguals processing English words. The model proposes that English letters do not only activate English phonemes, but also similar-sounding Japanese counterparts. In this way, the English word *coffee* /khɔfi/ can sublexically activate part of its Japanese equivalent /koohii/ (with the double vowels representing a moraic long vowel). In other words, cognate effects still arise, but their locus may be different from that in alphabetic scripts (see Wen & van Heuven, 2018, for a symbolic implementation of repetition priming effects between Chinese and English). Thus, in contrast to Mishra (2018), we argue that Multilink (as well as BIA+) is clearly relevant "for those researchers who are trying to understand the psycholinguistics and cognitive processes of bilinguals whose languages often belong to different language families."

Evidence that even largely separate modalities converge at least partially at more central levels of processing comes from bimodal bilinguals using sign language and spoken language (Morford, Occhino-Kehoe, Pinar, Wilkinson & Kroll, 2015; Ormel, Hermans, Knoors & Verhoeven, 2012). In our view, an interpretation of underlying activation processes may benefit from assuming symbolic representations even for these bilinguals.

Our conclusion is that, in order to include a non-alphabetic input representation in a Multilink-type of computational model, a precise specification of the orthographic input for the script at hand is needed, as well as a similarity index for items within and across languages. Currently, there is a lack of both for most language and script combinations in bilinguals.

(3) Task demands, cognitive control, and dependent variables

In our keynote paper, we applied Multilink to a limited number of tasks, in particular lexical decision, word naming, and word translation. Recently, we conducted new simulations on bilingual lexical decision data kindly provided by our colleagues. First, we tested Multilink with log frequency resting level activations but no lateral inhibition on data provided by De Groot (30 September 2016; see keynote Simulation Series 3 for word naming simulations using the same set of 440 words). For English lexical decision (LDE-A), we obtained a correlation of 0.59 (N=419). For Dutch lexical decision (LDD-A), this correlation was 0.55 (N=429). Next, we simulated the Dutch (L1) lexical decision task by Van Assche, Duyck, Hartsuiker, and Diependaele (2009). This led to $r = 0.30$ for identical and non-identical cognates (N=34) and $r = 0.60$ for control words (N=24). Similar simulations of the English (L2) lexical decision task by Van Assche, Drieghe, Duyck, Welvaert, and Hartsuiker (2011) led to $r = 0.61$ for identical and non-identical cognates (N=24) and $r = 0.68$ for controls (N=26). Finally, we simulated other unpublished lexical decision data by Dirix, resulting in $r = 0.63$ (N = 77) for Dutch and $r = 0.55$ (N = 66) for English.

Although simulations are not yet available, Multilink is also equipped to mimic orthographic and semantic priming tasks. This is possible because Multilink allows read-out of orthographic, phonological, semantic, and language membership codes. Furthermore, we have considered different response criteria, applying recognition thresholds, Luce's choice rule, and probability-based lexical selection. Nevertheless, Multilink's task/decision system needs to be extended in terms of how it handles different task demands and cognitive control issues. This is certainly also necessary to account for mechanisms like semantic checking and cognitive control in a trilingual system (Biloushchenko, 2017), as mentioned by Ivanova and Kleinman (2018).

We were happy that Mishra (2018) referred to the aspect of cognitive control. At present, practically all available models for word retrieval lack a specification of how processing depends on task demands, stimulus list composition, and stimulus properties. However, as has been argued by several colleagues and ourselves in the past (Grainger & Jacobs, 1996; Green, 1998; Dijkstra & van Heuven, 1998), we believe that lexical selection and read-out are critically dependent on such factors.

Two examples of the importance of task demands and stimulus list composition are concerned with cognates and interlingual homographs. In our keynote paper, we simulated lexical decision results by Vanlangendonck (Vanlangendonck, 2012, Experiment 1; Vanlangendonck, Peeters, Rueschemeyer & Dijkstra, submitted) on cognate facilitation. In a second experiment, she was the first to show that such effects can be turned into mirrored inhibition effects when Dutch words replace half of the pseudowords. It was argued that this can be explained by an effect of competition by the other cognate member on response selection. As a second example, consider the processing of interlingual homographs in the word translation task. Suppose a French–English bilingual encounters the French word COIN (cf. Beauvillain & Grainger, 1987). How is this word translated into the English word CORNER instead of being repeated as

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

the English word COIN? The input COIN will activate orthographic, semantic, and phonological representations in both languages. Without a task/decision system that makes smart use of language membership information with respect to input and output, the problem cannot be solved.

We are currently implementing a task/decision account in Multilink based on newly collected data that can simulate the performance of cognates and interlingual homographs (Goertz, Wahl & Dijkstra, in preparation; also see Christoffels, Ganuschchak & Koester, 2013). Our basic assumption is that the task/decision system can use complex task-sensitive criteria for read-out based on several representations, e.g., semantic, phonological, and language membership representations. As in BIA+, the task/decision system monitors and uses activation in the lexical processing system, but does not change it.

At present, models like BIA, BIA+, and Multilink, mostly apply a simple decision rule: when the activation of a word surpasses a recognition threshold, the item is considered to be recognized. Simulations are thus focussed on the dependent variable of reaction time. However, Tokowicz (2018) points out that important information about representation and processing can also be gained from considering errors and error rates. Especially because we also intend for Multilink to have some relevance for individual differences and learning, the model should take into account the errors that bilinguals make during processing. As the proverb says: "To err is human". It is interesting to note that Multilink can (and does) already produce errors. An error arises if, after a certain number of time cycles, a competitor word rather than the target word reaches its recognition threshold, or when no threshold is reached. It is possible to simulate error performance also by using a stochastic recognition threshold (see Jacobs & Grainger, 1992) or variations in temporal deadline. Thus, there are various decision criteria relating to error rates that can be explored in future simulations with Multilink. Furthermore, the nature of the errors Multilink makes is revealing. For instance, for certain parameter settings, the model translated English ANT (Dutch: 'mier') into Dutch TANTE ('aunt') because ANT is embedded in both AUNT and in TANTE.

(4) Issues with respect to lexical organization

Orthography and lateral inhibition

Several reviewers consider the absence of lateral inhibition in Multilink as problematic. They hold that inhibition between lexical neighbors and by language membership representations (e.g., language nodes and tags) down to lexical representations is a necessary feature for models of word recognition, both monolingual and bilingual.

Van Heuven and Wen (2018) point out that inhibition effects in the masked (orthographic) priming task were simulated quite well by BIA due to the presence of lateral inhibition. With lateral inhibition switched off, Multilink essentially treats orthographic primes as if they were nonwords, whereas, for instance, Dijkstra, Hilberink-Schulpen, and van Heuven (2010) observed clear inhibitory effects of orthographic overlap for word primes in a masked priming task.

Furthermore, as argued by Declerck, Meade, and Grainger (2018), reducing the activation if two nodes converge on the same semantic representation in the current model might be superfluous with lateral inhibition at work.

We do not disagree with the usefulness of lateral inhibition, which we have indeed exploited in earlier work. As Declerck et al. (2018) indicate, "lateral inhibition is already implemented in Multilink, but was deactivated in the current version". The present baseline model without lateral inhibition allows a maximally contrastive comparison to models with lateral inhibition, such as BIA and BIA+. Nevertheless, we still obtained surprisingly high correlations between simulations and empirical data for just one parameter setting across different tasks. Van Heuven and Wen (2018) point out that different model variants should be compared against each other, and contrasting Multilink without and with lateral inhibition is one important possibility (see below).

Although it is evident that inhibitory mechanisms must be built into the model, how to do this and in which parts of the model is not clear at all. For instance, Ivanova and Kleinman (2018) state that "inhibition of any kind seems problematic for the facilitatory effect of translation equivalent distractors in picture-word interference studies". In addition, there are many options that require the modellers' consideration before they can be implemented. First, should we include only lateral inhibition (within the identification system) as a competitive mechanism or also inhibition in terms of response level competition (in the task/decision system)? And, within the word identification system, should it be lateral inhibition, inhibition in terms of the Luce choice rule, or bottom-up mismatch? Bard (1990) has already argued that it may be difficult to distinguish lateral inhibition from other competitive mechanisms; it might be best to consider these in their interaction with high and low frequency words. Sorting out this interaction requires considerable simulation work.

Second, following the localist connectionist paradigm, lateral inhibition should probably be implemented for both orthographic and phonological representations, but what about semantic representations? Should we forge connections "only between translation equivalents, between synonyms as well, or between all words in a speaker's lexicon" (Ivanova & Kleinman, 2018)? If we

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

move to distributed semantic representations (or smaller grain size representations), lateral inhibition might be unwanted or unnecessary.

Third, should lateral inhibition be implemented in both comprehension and production, using the same parameter settings? As Mishra (2018) wonders: "Does inhibition work similarly for non-selective activation in naming, translation, and spoken word recognition?" On the basis of production studies, it could be argued that not just resting level activation, but also connection strengths between representations should be adapted (cf. the "weaker link" hypothesis by Gollan, Montoya, Cera & Sandoval, 2008).

Finally, there is complex evidence from behavioral and brain studies that a distinction must be made between several different types of inhibition effects, for instance, domain-general or specific to language (de Bruin, Roelofs, Dijkstra & FitzPatrick, 2014), sustained or transitory, and global or local (Peeters & Dijkstra, 2017). The consequences for modelling are as yet unspecified.

We propose to study these questions within a systematic program of simulation studies. Currently, at the beginning of the program, many parameters, including lateral inhibition, are set at a baseline value of zero.

Simulations including lateral inhibition

The three panels in Figure 1 illustrate how strength differences in lateral inhibition affect the activation of translations of the input word CURSE and its competitors COURSE and NURSE. The panels clearly show a difficulty in finding proper parameter settings for orthographic (called OO_gamma) and phonological (PP_gamma) lateral inhibition: its effects are non-linear. Very small changes in parameter value result in strongly increasing orthographic recognition times (in Figure 1: from about 15 to over 30 time cycles with a threshold set at 0.72) in interaction with word frequency. Cycle times for phonological output are increased even more. When lateral inhibition is set at high values, non-target candidates are knocked out and simulation results approach those reported without lateral inhibition. For the data by Pruijn, Peacock, and Dijkstra (in preparation) (Simulation Series 3), we obtained correlations of 0.57 for Dutch to English and 0.68 for English to Dutch translation, similar to those in the keynote, for small values of lateral inhibition (using a large lexicon of 1367 word pairs). The consequences of these promising results for fitting neighborhood effects and priming tasks remain to be assessed. In contrast to what Declerck et al. (2018) suggest, the 'cognate patch' could not (yet) be replaced by lateral inhibition: identical cognates suffered disproportionally from lateral inhibition relative to non-identical cognates and control words. Only after suitable settings of the lateral inhibition parameter have been determined, can we start to simulate the masked translation priming studies mentioned by van Heuven and Wen (2018; e.g., Wen & van Heuven, 2017), as well as orthographic priming studies.

Phonology and sublexical representations

Several reviewers consider the mapping relation between orthography and phonology, and the implementation of the latter. As Jiang (2018) rightfully points out, a restricted mapping from orthographic to phonological representations at the lexical level will not suffice for simulating the many tasks in which sublexical representations play an important role. We did neither expect this nor intend it to do so. Implementing sublexical orthographic and phonological representations (like in SOPHIA, see Thomas & van Heuven, 2005) will be unavoidable in the long run. To do so, informed decisions must be made on how to represent auditory input in terms of sublexical units of different lengths.

Mishra (2018) points out that spoken words also influence orthographic processing (e.g., Dijkstra, Roelofs & Fieuws, 1995). Although the Multilink model does not accept spoken input, the implied interactivity between phonological and orthographic representations is implemented.

Jiang (2018) further argues that the current version of Multilink assumes language-specific phonological representations, whereas "there are good reasons for postulating a partially integrated (rather than a language-dependent) system of phonological representations." For instance, L2 segments may be parasitic on L1 segments and interact with their L1 counterparts (also see Post da Silveira, 2016). Because we specified the phonological representations in Multilink by using those for L1 speakers in CELEX, each language is currently indeed characterized in terms of its native phonological distinctions, without any accommodation for bilinguals. In a future version of Multilink, this must surely be changed. We therefore call upon researchers in the domain of bilingual speech recognition to provide precise characterisations of particular L1/L2 phonetic-phonological systems and their phonemic inventories. Important insights into these matters have been gained in the context of, for instance, BIMOLA (Léwy & Grosjean, 2008), BLINCS (Shook & Marian, 2013), and SOMBIP and DEVLEX (Zhao & Li, 2010, 2013). As a simple temporary modelling solution, we might try to capture cross-linguistic similarities or differences in phonological segments in terms of Levenshtein distance or ratings (Levenshtein, 1966). Less overlap between phonological representations across languages should result in less co-activation (e.g., /f/ and /v/ might be considered to lead to partial activation depending on feature overlap).

The nature of orthographic and phonological representations also has consequences for the processing of translation equivalents with form overlap between languages (cognates). Similarly sounding cognates

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

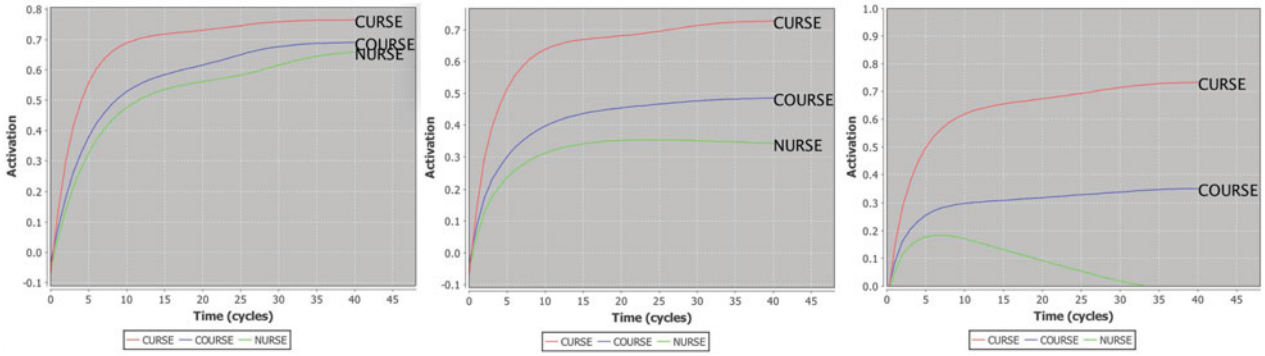


Figure 1. (Colour online) Multilink's activation of the input word CURSE and neighbors COURSE and NURSE when no lateral inhibition is present (left panel) and when lateral inhibition set at −0.05 (middle panel) or at −0.10 (right panel).

occur in language pairs all over the world, and, in contrast to Mishra (2018), we believe they are important to clarify effects of simultaneous (auditory) form and meaning similarity on bilingual processing. Whereas orthographic similarities are restricted to overlapping scripts, phonological similarities can be found everywhere. We do agree on the importance of studying the consequences of an absent orthographic system in (bilingual) speakers with limited literacy, e.g., to assess the stability and 'fuzziness' of the phonological system – this is a problem for most if not all existing models. Again, empirical studies providing an adequate representation of the mental lexicons of these individuals must open the door to meaningful simulations.

Semantics and word associations, features, and vectors

The present version of Multilink incorporates an admittedly oversimplified representation of semantics: lexical concepts are represented as holistic units. In line with the suggestions by Li and Grant (2018), Multilink has already been prepared for simulating spreading activation via semantic associations (De Deijne & Storms, 2008; see Dijkstra & Rekké, 2010, for a picture). Our bilingual lexicons include English items that were selected on the basis of their presence in an association database (Nelson, McEvoy & Schreiber, 1998). Thus, simulations can easily be prepared to study language-independent and -dependent effects of association at form and meaning levels. Implementation of multiple mappings (in relation to context sensitivity) would also be desirable (Tokowicz & Kroll, 2007). In line with symbolic model extensions, the Distributed Feature model proposed by Van Hell and De Groot (1998) could be implemented. By means of distributed features, language-specific connotations or concepts could be incorporated in the model.

More thorough model adaptation would involve replacing the current semantic representations by subsymbolic representations or vector representations, even if this might mean the model is no longer purely localist connectionist. Such work is currently being done for SOPHIA (collaboration van Heuven / Dijkstra).

Language membership: tag, functional representation, or lexical property

Multilink's architecture incorporates two language membership representations, each of which can be activated by orthographic and phonological representations of the words of that language (as well as via sentence context information). There are several interesting issues linked to the notion of 'language membership', as also pointed out by Johns and Putnam (2018).

Current symbolic models mostly assume that language membership can be seen as a specific, separate type of linguistic representation (in addition to word form, lexical concept, and lemma). However, even if language information has its own representation, it is unclear how this interacts with other types of representations (cf. Hoversten et al., 2015). With respect to language production, should language membership be considered as just a label (tag) or as a more functional node, such as a separate semantic representation (Ivanova & Kleinman, 2018) or a lemma? Or is the intention to speak in a particular language an integral aspect of the message? With respect to recognition, to what extent does non-linguistic context activate language membership representations either directly or indirectly?

Consider an experiment in which French and English flags are used as background context for French and English words (Grainger, Declerck & Marzouki, 2017). A blue-white-red flag might activate the concept "French", but is this concept identical to the French language membership representation of a word? Furthermore, the task/decision system might decide to use the flag information as a RESPONSE CUE depending on whether this information is beneficial to task performance or not (see Casaponsa, Carreira & Duñabeitia, 2015; Van Kesteren et al., 2012). It is also conceivable that language membership itself is used as a cue by the task/decision

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

system depending on its usefulness. Thus, obtaining behavioral context effects in this situation might not constitute direct evidence against BIA+ or Multilink; perhaps EEG or fMRI-data could help to disentangle effects of interactive activation during identification or decision-making. Given what we see as a lack of convincing empirical evidence, we agree with Declerck et al. (2018) that future work should investigate the facilitatory or inhibitory role of language membership representations in the recognition of words in isolation and in sentences. We currently do not favor the BIA-model's assumption that the language membership label of a target word is linked up to each and every word of ALL OTHER LANGUAGES of the bilingual to make inhibition possible.

As an intriguing alternative, Johns and Putnam (2018) propose to represent language membership as a property of the lexicon resulting from gradient emergent principles in cross-situational learning. The learner might come to associate words probabilistically with particular contexts and feature sets. An abstract representation of language membership could arise 'grounded in overlapping sets of features as opposed to singular concepts'. Implementation of word association and semantic priming would then be done in quite a different way than following a symbolic view of language membership. At present, this alternative is clearly out of Multilink's scope.

Conclusions

The interest in implemented models such as BIA, BIA+, and Multilink shows that researchers in the field of bilingualism like to have explicit frameworks for theoretical interpretation and embedding of their studies and as inspiration for empirical research. Even when such models are relatively underspecified (e.g., BIA), have coding problems (e.g., position-coding in BIA / BIA+), or missing components (e.g., sublexical representations in Multilink), thought experiments and simulations often have heuristic value for future research by clarifying research questions, quantifying results, and generating research ideas. Working in a unifying framework like the symbolic, localist-connectionist paradigm is useful for bringing different approaches and viewpoints into contact (we note that human communication itself is symbol-oriented). The multilingual network of representations in Multilink allows for the capture of quite diverse bilingual phenomena. It has clearly delineated representations, is context-sensitive, and relatively flexible.

On the basis of the various reviews, we conclude that the null-model version of Multilink underestimates the degree of cross-language interaction: bilingual processing must be made sensitive to phenomena like lateral inhibition effects, merged phoneme representations, and L1-dependent learning strategies. Nevertheless, given that the present version of Multilink led to high correlations between simulation and empirical data, it already appears to capture some fundamental characteristics of the bilingual's lexicon architecture. In addition, several of the problems indicated by the reviewers can be solved easily in future Multilink variants.

In all, we argue that the present version of Multilink is a promising first step towards a more complex model that even better accounts for the intriguing complexity, context sensitivity, interactivity, and variability of languages. In a modest vein, it seems appropriate to end this paper with a variation on some famous lines from *Rigoletto* by Giuseppe Verdi: "La lingua è mobile, qual piuma al vento, muta d'accento, e di pensiero" ('Language is flighty. As a feather in the wind, it changes in voice and in thought.')

References

Bard, E. G. (1990). Competition, lateral inhibition, and frequency: Comments on the chapters by Frauenfelder and Peeters, Marslen-Wilson, and others. In G.T.M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge MA: MIT Press (pp. 185–210).

Beauvillain, C., & Grainger, J. (1987). Accessing interlexical homographs: Some limitations of a language selective access. *Journal of Memory and Language, 672,* 658–672.

Biloushchenko, I. (2017). *How trilinguals process cognates and interlingual homographs: the effect of activation, decision, and cognitive control*. Doctoral thesis, University of Antwerp.

Bowers, J. S. (2009). On the biological plausibility of grandmother cells: Implications for neural network theories in psychology and neuroscience. *Psychological Review, 116* (*1*), 220–251. doi:10.1037/a0014462

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

Christoffels, I. K., Ganuschchak, L., & Koester, D. (2013). Language conflict in translation: An ERP study of translation production. *Journal of Cognitive Psychology, 25,* 646–664.

Costa, A., & Pickering, M. (2018). The role of learning on bilinguals' lexical architecture: Beyond separated vs. integrated lexicons. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000809

Declerck, M., Meade, G., & Grainger, J. (2018). On keeping cool: The role of inhibition in bilingual word recognition. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000676

De Bruin, A. M. T. de, Roelofs, A. P. A., Dijkstra, T., & FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: fMRI evidence from trilingual speakers. *NeuroImage, 90,* 348–359.

De Deijne, S., & Storms, G. (2008). Word associations: Norms for 1,424 words in a continuous task. *Behavior Research Methods, 40,* 198–205.

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

Dijkstra, T., Haga, F., Bijsterveld, A., & Sprinkhuizen-Kuyper, I. (2011). Assessing mechanisms of inhibition in localist and distributed connectionist models by simulating L2 word acquisition. In J. Altarriba & L. Isurin (Eds.), *Memory and Language: Theoretical and Applied Approaches to Bilingualism*. John Benjamins.

Dijkstra, T., Hilberink-Schulpen, B., & van Heuven, W. J. B. (2010). Repetition and masked form priming within and between languages using word and nonword neighbors. *Bilingualism: Language and Cognition, 13,* 341–357.

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity affects cognate recognition. *Journal of Memory and Language, 62,* 284–301.

Dijkstra, T., & Rekké, S. (2010). Towards a localist-connectionist model for word translation. *The Mental Lexicon, 5*(*3*)*,* 403–422. *Special issue on Methodological and analytic frontiers in lexical research*, edited by G. Jarema, G. Libben, & Ch. Westbury.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology, 49* (*2*)*,* 264–271.

Dijkstra, T., Van Hell, J. G., & Brenders, P. (2015). Sentence context effects in bilingual word recognition: cognate processing, sentence language, and semantic constraint. *Bilingualism: Language and Cognition, 18,* 597–613.

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA-model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). Hillsdale, NJ: Lawrence Erlbaum Associates.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Duyck, W., Vanderelst, D., Desmet, T., & Hartsuiker, R. (2008). The frequency effect in second-language visual word recognition: evidence against a language-selective rank hypothesis. *Psychonomic Bulletin and Review, 15*(*4*)*,* 850–855.

Goertz, R., Wahl, A., & Dijkstra, A. (in preparation). Translating interlingual homographs: Empirical and simulation data. [working title]

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1122–1139.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always a means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58,* 787–814.

Goral, M. (2018). The bilingual mental lexicon beyond Dutch-English written words. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000743

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.

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

Grainger, J., & Jacobs, A. M. (Eds.). (1998). *Localist Connectionist Approaches to Human Cognition*. Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81.

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,* pp. 2108–2116. doi:10.1162/jocn_a_00844

Iakovleva, T., Piasecki, A. E., & Dijkstra, A. F. J. (2015). Are you reading what I am reading? The impact of contrasting alphabetic scripts. In V. Pirelli, C. Marzi, & M. Ferro (Eds.), *NetWordS 2015. Word Knowledge and Word Usage. Representations and Processes in the Mental Lexicon* (pp. 112–116).

Ivanova, I., & Kleinman, D. (2018). Multilink for bilingual language production. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000718

Jacobs, A. M., & Grainger, J. (1992). Testing a semistochastic variant of the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 1174–1188.

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

Jiang, N. (2018). Phonology-based bilingual activation among different-script bilinguals? *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000664

Johns, M. A., & Putnam, M. T. (2018). Language membership as a gradient emergent feature. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S136672891800072X

Kim, J., & Davis, C. (2003). Task effects in masked cross-script translation and phonological priming. *Journal of Memory and Language, 49,* 484–499.

Kim, S. (2017). *The cross-language activation of Korean and English in a monolingual setting*. Qualifying Paper, The University of Maryland, March 27, 2017.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

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,* 373–381.

Lauro, J. G., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *The Journal of Memory and Language, 92,* 217–233.

Lemhoefer, K., Dijkstra, A., Schriefers, H., Baayen, H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language:

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

A mega-study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 12–31.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Cybernetics and Control Theory, 10,* 707–10.

Léwy, N., & Grosjean, F. (2008). The Léwy and Grosjean BIMOLA model. In F. Grosjean (ed.), *Studying bilinguals*, pp. 201–210. Oxford: Oxford University Press.

Libben, M., Goral, M., & Libben, G. (2017). (eds.), *Bilingualism: A Framework for Understanding the Mental Lexicon*. John Benjamins.

Li, P., & Grant, A. (2018). Scaling up: How computational models can propel bilingualism research forward. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000755

Meade, G., & Dijkstra, A. (2017). Mechanisms underlying word learning in second language acquisition. In M. Libben, M. Goral, & G. Libben (eds.), *Bilingualism: A Framework for Understanding the Mental Lexicon*. John Benjamins (pp. 49–71).

Mishra, R. K. (2018). A few suggestions on broadening the cross-linguistic relevance of the Multilink model. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000834

Miwa, K., Dijkstra, T., Bolger, P., & Baayen, R. H. (2014). Reading English with Japanese in mind: Effects of frequency, phonology, and meaning in different-script bilinguals. *Bilingualism: Language and Cognition, 17,* 445–463. doi: 10.1017/S1366728913000576

Morford, J. P., Occhino-Kehoe, C., Pinar, P., Wilkinson, E., & Kroll, J. F. (2015). The time course of cross-language activation in deaf ASL–English bilinguals. *Bilingualism: Language & Cognition*, 337–350.

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: Evidence for integrated phonological representations. *Language and Cognitive Processes, 27,* 1563–1583. DOI: 10.1080/01690965.2011.606669

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/.

Ormel, E., Hermans, D., Knoors, H., & Verhoeven, L. (2012). Cross-language effects in visual word recognition: The case of bilingual deaf children. *Bilingualism: Language and Cognition, 15,* 288–303.

Page, M. (2000). Connectionist modelling in psychology: A localist manifesto. *Behavioral and Brain Sciences, 23,* 443–512.

Page, M. (2017). Localist models are compatible with information measures, sparseness indices, and complementary-learning systems in the brain. *Language, Cognition, and Neuroscience, 32,* 366–379.

Peeters, D., & Dijkstra, T. (2017). Sustained inhibition of the native language in bilingual language production: A virtual reality approach. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728917000396

Post da Silveira, A. (2016). *Word stress in second language word recognition and production*. Doctoral thesis Radboud University, Nijmegen, the Netherlands.

Pruijn, L., Peacock, J., & Dijkstra, T. (in preparation). Mechanisms of word translation: Empirical and simulation data. [working title]

Sianipar, A., Middelburg, R., & Dijkstra, T. (2015). When feelings arise with meanings: How emotion and meaning of a native language affect second language processing in adult learners. *PLoS ONE, 10*(*12*): e0144576. doi:10.1371/journal.pone.0144576

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition, 16,* 304–324. doi:10.1017/S1366728912000466.

Thomas, M. S. C., & van Heuven, W. J. B. (2005). Computational models of bilingual comprehension. In J.F. Kroll & A.M.B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 202–225). Oxford University Press.

Tokowicz, N. (2018). The role of errors and individual differences in bilingual translation. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S136672891800069X

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes, 22,* 727–779.

van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., and Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64,* 88–107.

van Assche, E., Duyck, W., Hartsuiker, R. J., and Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science, 20,* 923–927.

Van Hell, J. (2018). Words go only so far: Linguistic context affects bilingual word processing. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000706

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193–211.

van Heuven, W. J. B., & Wen, Y. (2018). The need for a universal computational model of bilingual word recognition and word translation. *Bilingualism: Language and Cognition*. https://doi.org/10.1017/S1366728918000688

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

Van Kesteren, R., Dijkstra, T., & De Smedt, K. (2012). Markedness effects in Norwegian-English bilinguals: Task-dependent use of language-specific letters and bigrams. *The Quarterly Journal of Experimental Psychology, 65,* 2129–2154. DOI:10.1080/17470218.2012.679946

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.

Vanlangendonck, F., Peeters, D., Rueschemeyer, S.A., & Dijkstra, T. (submitted). Stimulus list composition can turn cognate facilitation effects into mirrored inhibition effects.

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press

van Rijswijk, R. (2016). *The strength of a weaker language: Language production and comprehension by Turkish heritage speakers in the Netherlands*. Doctoral thesis Radboud University, Nijmegen, the Netherlands.

Wen, Y., & van Heuven, W. J. B. (2017). Non-cognate translation priming in masked priming lexical decision experiments: A meta-analysis. *Psychonomic Bulletin & Review, 24* (*3*), 879–886.

Wen, Y., & van Heuven, W. J. B. (2018). Limitations of translation activation in masked priming: Behavioural evidence from Chinese-English bilinguals and computational modelling. *Journal of Memory and Language, 101,* 84–96. doi: 10.1016/j.jml.2018.03.004

Xiang, H., Dediu, D., Roberts, L., Van Oort, E., Norris, D., & Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory and IQ in the perisylvian language network. *Language Learning,* 62 (Supplement S2), 110–130. doi:10.1111/j.1467-9922.2012.00708.x

Xiang, H., Van Leeuwen, T. M., Dediu, D., Roberts, L., Norris, D. G., & Hagoort, P. (2015). L2-proficiency-dependent laterality shift in structural connectivity of brain language pathways. *Brain Connectivity, 5*(*6*), 349–361. doi:10.1089/brain.2013.0199.

Zhao, X., & Li, P. (2010). Bilingual lexical interactions in an unsupervised neural network model. *International Journal of Bilingual Education and Bilingualism, 13,* 505–524.

Zhao, X., & Li, P. (2013). Simulating cross-language priming with a dynamic computational model of the lexicon. *Bilingualism: Language and Cognition, 16,* 288–303.

https://doi.org/10.1017/S1366728918000986 Published online by Cambridge University Press