

Exploring early language detection in balanced bilingual children: The impact of language-specificity on cross-linguistic nonword recognition

International Journal of Bilingualism

2018, Vol. 22(3) 305–315

© The Author(s) 2016

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1367006916672751

journals.sagepub.com/home/ijb**Pauline Schröter**

Max Planck Institute for Human Development, Germany

Sascha Schroeder

Max Planck Institute for Human Development, Germany

Abstract

Aims and objectives: Recent findings on the mechanisms of lexical access suggest that bilinguals are sensitive to the orthographic structure of their languages. Several studies have demonstrated that if presented with language-specific sub-lexical information, bilingual adults use this information to speed up word recognition, which provides evidence for language-selective lexical access. In the present study, we investigated the presence of such an early language detection mechanism in children.

Methodology: Forty-six balanced bilingual third-graders performed two seemingly monolingual lexical decision tasks, one in English and one in German, including nonwords with different degrees of word-likeness in each language.

Data and analysis: Accuracy scores and reaction times were analyzed for nonwords using mixed-effects models with the statistical software *R*.

Findings: Results show no impact of language-specific sub-lexical information on children's performance in either task. We argue that bilingual lexical access is initially language-nonspecific, and that sensitivity to language-specific orthographic structures first emerges over time. In contrast to bilingual adults, language detection in bilingual children is exclusively based on lexical information.

Originality: The present study provides first data on the detection mechanism for language membership at the early stages of bilingual reading development. We are the first to demonstrate an important difference in the architecture of the bilingual lexicon between children and adults.

Implications: Findings contribute to knowledge on the development of lexical access in bilinguals and pose limitations to the generalizability of the Bilingual Interactive Activation Plus (BIA+) extended model.

Keywords

Bilingual children, language detection, nonword recognition, lexical access, BIA+ extended

Corresponding author:

Pauline Schröter, Max Planck Institute for Human Development, MPRG Reading Education and Development (REaD), Lentzeallee 94, 14195 Berlin, Germany.

Email: pauline.schroeter@iqb.hu-berlin.de

Within research on bilingualism, there is ample evidence that same-script bilinguals activate both of their languages when reading in one of them. Data collected over the past two decades have shown that in individuals with a certain level of second language proficiency, visually presented words are simultaneously accessed in both of their languages (e.g. Duyck, 2005; van Assche, Duyck, Hartsuiker, & Diependaele, 2009; van Hell & Dijkstra, 2002; van Heuven, Dijkstra, & Grainger, 1998). Widely cited evidence for cross-linguistic activation is the cognate facilitation effect, which refers to the processing advantage for words that are orthographically and semantically similar in both of a bilingual's two languages. Within the frame of interactive activation (IA) models (McClelland & Rumelhart, 1981), the effect is commonly attributed to the fact that cognates share their orthographic, semantic, as well as phonological representations in the mental lexicon and thus reach their activation threshold sooner than matched non-cognates (e.g. Lemhöfer & Dijkstra, 2004). In contrast, inter-lingual homographs, which share their form but not their meaning between languages, have been found to cause null or even inhibitory effects (e.g. Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Timmermans, & Schriefers, 2000; Dijkstra, van Jaarsveld, & Ten Brinke, 1998). Based on these findings, the current model of bilingual word recognition, the Bilingual Interactive Activation Plus model (BIA+) (Dijkstra & van Heuven, 2002), postulates that bilingual lexical access is language-nonspecific and based on an integrated lexicon. It states that upon the presentation of a visual letter string, sub-lexical orthographic representations are activated, which subsequently activate sub-lexical phonological representations as well as orthographic and phonological entries on the lexical level. These lexical entries, in turn, activate semantic representations and initiate the process of language detection through so-called language nodes. However, the mechanism by which a word is associated with a respective language is still unclear.

Supporting the view of language detection postulated by the BIA+, studies with same-script bilinguals have demonstrated that when the language context is ambiguous, language information is accessed through the lexical representations of words (Chauncey, Grainger, & Holcomb, 2008; Dijkstra et al., 1999; Midgley, Holcomb, & Grainger, 2009a, 2009b; von Studnitz & Green, 2002). Language detection, therefore, has been assumed to be the result of top-down modulations from the language nodes feeding information back to the lexical units (Casaponsa, Carreiras, & Duñabeitia, 2015). Latest findings, however, suggest that balanced bilinguals are sensitive to the orthographic structure of their languages prior to word recognition. Recent studies have shown that if presented with language-specific cues – such as unique graphemes, more frequent bigrams, or larger orthographic neighborhood size – bilingual participants show reduced parallel language activation (e.g. Casaponsa, Carreiras, & Duñabeitia, 2014; Casaponsa & Duñabeitia, 2016; Lemhöfer & Dijkstra, 2004; Lemhöfer & Radach, 2009; van Kesteren, Dijkstra, & de Smedt, 2012). For instance, investigating the impact of language-specific versus language-nonspecific sub-lexical information, Casaponsa and Duñabeitia (2016) demonstrated that the absence of such cues promoted some degree of language-nonspecific lexical access, whereas their presence reduced interference from the non-target language. The authors concluded that bilinguals develop fine-grained sensitivity to language-specific sub-lexical information, which leads to a different organization of lexical representations depending on the degree of language-specificity of the words. They hypothesized that mechanisms of lexical access might be shaped by sub-lexical distributional probabilities within and between languages. While in the absence of sub-lexical cues lexical access is language-nonspecific, in their presence language-selective access is enabled. This, however, poses a challenge to the BIA+ as postulated by Dijkstra and van Heuven (2002). Addressing this challenge, van Kesteren et al. (2012) proposed to extend the model by adding sub-lexical language nodes. Accordingly, in addition to lexical nodes that are connected to the lexical level, there are sub-lexical nodes which can be directly accessed through excitatory connections from the sub-lexical level. This way, depending on the presence or absence of language-specific

sub-lexical cues, language detection can also happen prior to lexical access. Oganian, Conrad, Aryani, Heekeren, and Spalek (2016) further proposed the alternative view of a unique set of language nodes that might accumulate lexical and sub-lexical information in parallel.

Notwithstanding ambiguities on the specific locus of language nodes, there is consensus on the fact that these nodes enable bilinguals to use sub-lexical information in order to detect the language membership of a letter string. Evidence for this account has been provided by reaction time studies using a range of different paradigms, including lexical decision (Lemhöfer & Dijkstra, 2004; Lemhöfer & Radach, 2009; van Kesteren et al., 2012), masked priming (Casaponsa & Duñabeitia, 2016), progressive demasking (Casaponsa et al., 2014), and naming (Oganian et al., 2016), as well as by experiments using event-related potentials (Casaponsa et al., 2015). Likewise, different markers for language membership have been studied. Exploring the nature of word-likeness, Bailey and Hahn (2001) compared measures of sequence probability and neighborhood size in their ability to explain empirical word-likeness judgments in English. Their results revealed a superior impact of neighborhood size relative to orthotactic and phonotactic measures. Oganian and colleagues (2016) further stated that in order to investigate language membership decisions in bilinguals, variables that are differently distributed between their two languages are especially relevant. Conducting a corpus analysis based on the German and the English Subtlex databases (Brysbaert et al., 2011), they demonstrated that neighborhood size served as the best source of language membership information. More than 90% of the words of each language had more orthographic neighbors in their own language than in the other one, whereas the distributions of bigram frequencies showed a high overlap between both languages. For the purpose of discriminating between the orthographic structures of German and English, therefore, it seems advisable to select neighborhood size over orthographic frequency measures.

A promising approach to investigating the effect of sub-lexical information is to study the processing of nonwords. A classical finding within this area of research is that in a lexical decision task (LDT), nonwords are rejected the faster the less word-like they are (Coltheart, Davelaar, Jonasson, & Besner, 1977; Forster & Shen, 1996). This observation was first explained by the Multiple Read-Out Model (Grainger & Jacobs, 1996) – an IA-type model which postulates that the more word-like a word is, the more representations (e.g. orthographic neighbors) it will activate. The underlying theory suggests that if, at a certain point in time, the search for a matching word candidate in the lexicon has remained unsuccessful, the stimulus will be rejected as a word. This temporal deadline is set later the more word-like a stimulus is. Within the framework of leaky competing accumulator models (Usher & McClelland, 2001), Dufau, Grainger, and Ziegler (2012) revised this theory by proposing a dynamic deadline account. Accordingly, the rejection of a stimulus as a word is equal to a constant value that optimizes the speed and accuracy in an LDT minus the activation of the stimulus as a word. Nonword recognition, hence, is a function of the amount of lexical activity generated by a stimulus. Manipulating the word-likeness of nonwords according to German and English neighborhood sizes, Lemhöfer and Radach (2009) asked German–English bilingual adults to perform a seemingly monolingual German, a monolingual English, and a mixed LDT. They found that English-like nonwords were more difficult to reject in the English relative to the German task, and vice versa. The authors concluded that the bilingual word recognition system makes a distinction between languages before their actual recognition or rejection. In line with temporal deadline accounts, German-like nonwords were less word-like in the English task, which is why their temporal deadline for rejection was set earlier than for English-like stimuli. In other words, the more English-like a nonword was in the English task, the harder it was for the recognition system to reject it as a word, and vice versa. In the mixed task, responses were generally slower, but participants reacted faster and more accurately to German-like compared to English-like stimuli. The authors explained this finding by the fact that participants were unbalanced bilinguals with a

greater proficiency in German compared to English. They concluded that if stimuli resemble the weaker language, their temporal deadline is set later, which is why they take more time to be processed than stimuli resembling the stronger language. Taken together, Lemhöfer and Radach's findings indicate that rejection criteria for nonwords depend on the language context, which provides further evidence for the view that bilingual lexical access is language-selective if language-specific sub-lexical information is given.

So far, research on language detection in bilinguals has been exclusively conducted with adults. To the best of our knowledge, there is no study that has ever investigated early language detection in bilingual children. Exploring how the degree of cross-linguistic orthographic overlap influences bilingual word recognition at different stages of reading development, Duñabeitia, Ivaz, and Casaponsa (2016) recently demonstrated that the cognate effect as a marker for language co-activation declined as a function of increasing exposure to print. The authors interpreted these findings in terms of different language interference suppression skills of younger and older children. They hypothesized that in a still immature bilingual language control system, top-down regulatory activity from the language nodes is impoverished, leading to a lack of inhibitory regulation at the lexical level. Yet, especially with regard to the development of sub-lexical language nodes as postulated by the BIA+ extended model, knowledge on the sensitivity to orthographic information in children is scarce. The goal of the present study was to fill this gap by investigating the presence of an early language detection mechanism at the beginning of reading development. Linking to previous research on nonwords, we conducted two language-specific LDTs, one in German and one in English, and manipulated nonwords according to their word-likeness in both languages. To rule out proficiency effects, we recruited balanced bilingual children who had started reading acquisition in German and English at the same time. We predicted that in a seemingly monolingual context, a fast-operating sub-lexical route sensitive to orthographic information would perceive differences in word-likeness. In other words, if bilingual children were sensitive to language-specific sub-lexical information like adults, lexical access would be language-selective. That is, in the English LDT, German-like nonwords, which activate less word candidates in English than English-like nonwords, should be rejected faster and more accurately than English-like nonwords. In the German LDT, the reverse should be true. If, on the other hand, there was no performance difference between German-like and English-like nonwords, this would be evidence for language-nonselective access. In that case, language-specific sub-lexical information would not be used to speed up the recognition process. This, in turn, would indicate that language detection in bilingual children depends on lexical information, which would argue for the absence of sub-lexical nodes in the early stages of the bilingual lexicon.

Method

Participants

Participants were recruited from a bilingual school, in which the language of instruction was 50% German and 50% English. Forty-six third-graders (21 female, $M = 7.65$ years, $SD = 0.48$) participated in the study, which was conducted during regular school hours and comprised two sessions each lasting 45 minutes. As part of an admission requirement, all children proved to be fluent speakers of German and English upon entering school. At the time of testing, they had received two years of formal reading instruction in each language. All of them reported using both languages equally on a daily basis and to have normal or corrected-to-normal vision.

To rule out sampling effects, we assessed nonverbal intelligence by administering the CFT 20-R (Weiß, 1998). Participants did not differ from the norm for monolinguals of the same age group

(sample: $M = 5.04$, norm sample: $M = 5.4$, $t < 1$, $p = .31$). To ensure equal language proficiency, we measured vocabulary knowledge using the CFT 20-R Vocabulary Test (Weiß, 1998) for German and the British Picture Vocabulary Scale (Dunn & Dunn, 2009) for English. Both tests consisted of multiple-choice items that required participants to select the closest-matching equivalent for a given target word. The mean percentile was 30.0 ($SD = 21.6$) for German and 26.4 ($SD = 20.9$) for English. As often reported for bilingual children's vocabulary knowledge (Bialystok, Luk, Peets, & Yang, 2010), scores were lower than the monolingual norm. However, results were comparable in German and English ($t < 1$, $p = .42$), which indicated equal vocabulary knowledge in both languages. Additionally, we assessed children's word and nonword reading fluency in each language through computerized speed reading tests, which require participants to name single words and nonwords as fast as possible. In German, the Salzburger Leserechtschreibtest (Moll & Landerl, 2010) for German revealed a mean raw score of 56.6 ($SD = 24.3$) for words and 36.2 ($SD = 13.5$) for nonwords. In English, the Test of Word Reading Efficiency (Torgeson, Wagner, & Rashotte, 1999) yielded a mean raw score of 55.5 ($SD = 16.8$) for words and 36.1 ($SD = 12.6$) for nonwords. We interpreted these results as an indication for participants' equal reading fluency in both languages (all $ts < 1$).

Stimuli

Words for the LDTs were taken from the childLex corpus for German (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015) and from the TASA corpus for English (Zeno, Ivens, Millard, & Duvvuri, 1995), which are both solely based on children's literature. We selected 128 English and 128 German nouns that were matched on length and frequency. Nonwords were constructed from these words for each language separately using the multilingual pseudoword generator Wuggy (Keuleers & Brysbaert, 2010), which is based on an algorithm that replaces sub-syllabic elements (i.e. onset, nucleus, or coda) of words with equivalent elements from other words of the same language. To avoid language-unique graphemes, for each word 10 nonwords were generated, from which we hand-picked the most optimal one. All nonwords were pronounceable, ranged from three to eight letters ($M = 4.5$, $SD = 1.2$), and did not differ in length between English and German ($t < 1$).

Language-specificity was verified using two measures of orthographic neighborhood. Comparisons of orthographic neighborhood size (N) (Coltheart et al., 1977) and orthographic Levenshtein distance 20 ($OLD20$) (Yarkoni, Balota, & Yap, 2008) for nonwords between the languages showed that nonwords had an overall greater lexical similarity to the language they were supposed to resemble. English-like nonwords had more orthographic neighbors and a smaller Levenshtein distance in English than German-like nonwords, and vice versa, as verified by one-sided t -tests (all $ps < .03$). Language-specificity was additionally validated through a rating study performed by 12 adult native speakers of English and German respectively. Participants rated nonwords according to their English- or German-likeness on a five-point Likert scale in their native language. Results showed that English-like as well as German-like nonwords were rated higher in the language they were supposed to resemble ($ps < .01$). Characteristics for the final set of nonwords in both languages are provided in Table 1.

For each language, words and nonwords were randomly assigned to two lists each including 64 words and 64 nonwords. For nonwords, one of the two lists was then replaced with a list from the other language. That is, stimuli for the English LDT consisted of 128 English words, 64 English-like nonwords, and 64 German-like nonwords, while stimuli for the German LDT included 128 German words, 64 German-like nonwords, and 64 English-like nonwords. As in German nouns are always capitalized, English-like nonwords in the German LDT were capitalized, while German-like nonwords in the English LDT were uncapitalized. Due to a technical error, 10 words had to be

Table 1. Characteristics of English-like and German-like nonwords.

	English-like nonwords	German-like nonwords
Length	4.6 (1.2)	4.5 (1.1)
Mean <i>N</i> in English	7.6 (7.4)	5.6 (6.2)
Mean <i>N</i> in German	4.6 (6.0)	7.6 (8.5)
Mean <i>OLD20</i> in English	1.7 (0.5)	1.8 (0.5)
Mean <i>OLD20</i> in German	1.9 (0.6)	1.7 (0.5)
Mean rating for word-likeness in English	3.2 (0.8)	2.3 (1.1)
Mean rating for word-likeness in German	2.3 (0.9)	3.2 (0.7)

OLD20: orthographic Levenshtein distance 20; N: Coltheart's N.

excluded from all analyses because they were duplicates (3) or existing words (4) in each language. For the complete set of nonwords used in each LDT, see the Appendix.

Procedure

Participants performed two seemingly monolingual LDTs in a counterbalanced order. The experiment was conducted using IBM-compatible laptops, which recorded reaction times (RTs) automatically while participants used the keyboard to respond. Items were presented in Courier New font on a 15-inch TFT-screen in white 28-point letters on a black background. Children were instructed to decide whether or not a presented letter string formed a correct word in German (German LDT) or in English (English LDT), and asked to perform as quickly and accurately as possible. To further boost the level of activation of the target language, the language of instruction was English during the English LDT and German during the German LDT. For every language, participants completed a practice block with four trials. Words and nonwords were randomly assigned to three blocks of 46 items each. Each trial began with the presentation of a fixation cross for 500 ms, followed by another 500 ms until the item appeared, which remained on screen until a response was given.

Results

Accuracy scores and RTs were analyzed for nonwords only using mixed-effects models (Baayen, Davidson, & Bates, 2008) as implemented in the *lme4* package (version 1.0-4) (Bates, Maechler, Bolker, & Walker, 2015) in the statistical software *R*. RT data were log-transformed and analyzed using a linear model, while accuracy data were logit-transformed and analyzed using a generalized linear model with a binomial link function. Stimuli and participants served as random effects, whereas language (German nonwords vs. English nonwords) and task (German LDT vs. English LDT) were included as fixed effects. Contrasts for post-hoc comparisons were estimated using the general linear hypotheses test generated with the *multcomp* package (Hothorn, Bretz, & Westfall, 2008). To control for the impact of vocabulary knowledge and nonword reading fluency, we fitted two additional models by separately adding vocabulary and reading fluency as fixed effects. Factors were generated for each language by centering participants' raw scores of the tests on vocabulary knowledge and nonword reading fluency and included as main effects in the model.

Table 2 contains the mean results for English-like and German-like nonwords in both LDTs from the main model. There was a main effect for task in accuracy data, $\chi^2(1) = 24.67, p < .01$, indicating that accuracy scores were higher in the German LDT than in the English LDT, which is a finding usually observed for transparent orthographies. However, there was no main effect for

Table 2. Mean reaction times (in ms) and accuracy scores (in % correct) for nonwords in both lexical decision tasks (LDTs). (Standard errors are given in parentheses.).

	Reaction time		Accuracy	
	English LDT	German LDT	English LDT	German LDT
English-like nonwords	1742 (285)	1708 (279)	87.28 (4.27)	92.90 (2.67)
German-like nonwords	1692 (276)	1717 (280)	85.21 (4.88)	92.31 (2.85)

language, $\chi^2(1) = 0.96, p = .33$, and no interaction between language and task, $\chi^2(1) = 0.04, p = .85$. These findings persisted after controlling for vocabulary and reading fluency in each language. For the RT analysis, incorrect trials and trials that deviated more than 2.5 *SDs* from either the stimulus or participant mean were discarded, accounting in sum for 18% of the raw data. There was neither a main effect for task, $\chi^2(1) = 0.02, p = .88$, nor for language, $\chi^2(1) = 0.46, p = .49$, in RT data. Although there was a tendency in the English LDT for German-like nonwords to be processed faster than English-like nonwords, the interaction between language and task, $\chi^2(1) = 0.98, p = .32$, did not reach significance. Again, these results persisted after controlling for vocabulary and reading fluency in each language. Additional analyses, which also accounted for children's chronological age as an indicator for differences in their time of exposure to print as well as for differences in their nonword recognition skills, revealed the same pattern of results with regard to the absence of a significant interaction between language and task.

To test whether the non-significance of our results actually points to the lack of language-specific sub-lexical processing in children, or merely indicates data insensitivity, we calculated the Bayes factor for general linear models. According to Dienes (2014), the Bayes factor compares the null hypothesis to an alternative hypothesis by providing a factor *B* by which the obtained results are more likely under the alternative than under the null. Dienes states that:

Bayes factors allow three different types of conclusions: There is strong evidence for the alternative (*B* much greater than 1); there is strong evidence for the null (*B* close to 0); and the evidence is insensitive (*B* close to 1). (p. 4)

We calculated *B* for the RT model including the main effects of Task and Language and the language \times task interaction while accounting for participants and stimuli as random factors. Using the function for general linear mixed-effects models as implemented in the *BayesFactor* package (Morey & Rouder, 2015), B_{RT} was 0.02, indicating that our results provide support for the null hypothesis rather than for insensitive data.

Discussion

The goal of the present study was to investigate the presence of an early language detection mechanism in balanced bilingual children. German–English bilingual third-graders performed two seemingly monolingual lexical decision tasks, one in each language, which each included German-like and English-like nonwords. We hypothesized that if children were sensitive to word-likeness as a language-specific cue, they would use this information to speed up their recognition process. Accordingly, English-like nonwords would be more easily identified as non-German words in the German LDT, and thus be rejected faster and more accurately than German-like nonwords. Given that participants were equally proficient in both languages, in the English LDT the reverse should be true.

Overall, results suggest that bilingual children are not sensitive to language-specific sub-lexical information. In both tasks, performance did not differ as a function of word-likeness, which indicates that children do not benefit from language-specific information on a sub-lexical level. This finding differs from observations on nonword processing in bilingual adults, who were found to be able to use orthographic cues in order to speed up the recognition process (e.g. Lemhöfer & Dijkstra, 2004; Lemhöfer & Radach, 2009). All results persisted after controlling for vocabulary and reading fluency in both languages, which rules out poor linguistic skills as an explanation for our findings. Additional models accounting for differences in children's chronological age and their nonword reading skills revealed no effect of word-likeness either, which indicates that findings are stable with regard to inter-individual differences. Also, the pattern of results was the same for German and English, which is what we expected to be the case in balanced bilinguals. We thus interpret our data as evidence for the absence of sub-lexical nodes in the early stages of the bilingual lexicon. With regard to the BIA+ extended model, we propose that initially there are only lexical nodes, and that sub-lexical nodes first emerge in the course of reading development. Accordingly, the fine-grained sensitivity to language-specific orthographic structures found in bilingual adults is the result of their extensive exposure to print in both languages. Based on the theory of statistical learning, beginning readers, in contrast, seem to not yet have the expertise to make use of this kind of information. This finding is in line with observations made by Duñabeitia et al. (2016), who demonstrated that cross-language activation on the lexical level diminished in the course of reading development. Young bilingual readers showed a greater reliance on cross-linguistic similarity than their older peers, which the authors ascribed to their still immature language control system.

Our findings further suggest that the word recognition system in bilingual children solely relies on information at the lexical level. Whereas for bilingual adults it is assumed that the mechanisms of lexical access are shaped by sub-lexical stages of orthographic processing (Casaponsa & Duñabeitia, 2016), this view does not seem to hold true for children. In contrast to adults, who show language-selective lexical access if language-specific information is given, our data provide evidence that lexical access in children is language-nonspecific despite the presence of language-specific cues. This challenges the applicability of the BIA+ extended model for children, which predicts that orthographically salient information immediately activates language nodes, which are then read out by the decision system. Given the absence of sub-lexical language nodes in beginning readers, as we propose, children rely on lexical language nodes only. From this it follows that language detection in bilingual children depends on lexical information and thus can first occur at the (post-) lexical stage in the word recognition process.

In sum, based on the present findings, we argue that bilingual lexical access is initially language-nonspecific, and that sensitivity to language-specific orthographic structures first emerges over time. In contrast to bilingual adults, who demonstrate the ability to detect language membership at an early stage in the word recognition process, we found that language detection in bilingual children is exclusively based on lexical information. To conclude, the present study provides first data on the detection mechanism of language membership at the early stages of bilingual reading development. We are aware that our findings are based on a limited set of nonwords and that replications are urgently needed to confirm our results as well as to extend them with regard to different language combinations and age levels. However, we demonstrate an important difference in the architecture of the bilingual lexicon between children and adults, which poses limitations to the generalizability of the BIA+ extended model. Further research should therefore include direct comparisons between bilingual and monolingual children as well as adults. Especially given that in today's world more and more children are being raised bilingually, data such as that we have provided are important to better understand the development of bilingual reading.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.
- Bailey, T. M., & Hahn, U. (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language*, *44*, 568–591.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2009). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*, 525–531.
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, *58*, 412–424.
- Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: The impact of orthographic markedness. *Frontiers in Psychology*, *5*(136), 1–10.
- 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*, 589–604.
- Chauncey, K., Grainger, J., & Holcomb, P. (2008). Code-switching effects in bilingual word recognition: A masked priming study with event-related potentials. *Brain and Language*, *105*, 161–174.
- Coltheart, M., Davelaar, E., Jonasson, T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention & performance IV* (pp. 535–555). Hillsdale, NJ: Erlbaum.
- Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 1–17.
- Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*, 496–518.
- Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language*, *42*, 445–464.
- 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.
- Dijkstra, T., van Jaarsveld, H., & Ten Brinke, S. T. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51–66.
- Dufau, S., Grainger, J., & Ziegler, J. C. (2012). How to say “no” to a nonword: A leaky competing accumulator model of lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1117–1128.
- Duñabeitia, J. A., Ivaz, L., & Casaponsa, A. (2016). Developmental changes associated with cross-language similarity in bilingual children. *Journal of Cognitive Psychology*, *28*(1), 16–31.
- Dunn, L. M., & Dunn, D. M. (2009). *The British Picture Vocabulary Scale: Third Edition*. London: GL Assessment.
- Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: Evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1340–1359.
- Forster, K. I., & Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory neighborhood effects in lexical decision and semantic categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 696–713.

- Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, *103*, 518–565.
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*, 346–363.
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*, 627–633.
- Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, *32*, 533–550.
- Lemhöfer, K., & Radach, R. (2009). Task context effects in bilingual nonword processing. *Experimental Psychology*, *56*, 41–47.
- 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*, 375–407.
- Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009a). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics*, *22*, 281–300.
- Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009b). Masked repetition and translation priming in second language learners: A window on the time-course of form and meaning activation using ERPs. *Psychophysiology*, *46*, 551–565.
- Moll, K., & Landerl, K. (2010). *SLRT-II: Lese- und Rechtschreibtest: Weiterentwicklung der Salzburger Lese- und Rechtschreibtests (SLRT): Manual*. Bern, Switzerland: Verlag Hans Huber.
- Morey, R., & Rouder, J. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs. R package vers. 0.9.12-2*. Retrieved from <https://CRAN.R-project.org/package=BayesFactor>.
- Oganian, Y., Conrad, M., Aryani, A., Heekeren, H. R., & Spalek, K. (2016). Interplay of bigram frequency and orthographic neighborhood statistics in language membership decision. *Bilingualism: Language and Cognition*, *19*, 578–596.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods*, *47*, 1085–1094.
- Torgesen, J. K., Wagner, R., & Rashotte, C. (1999). *TOWRE: Test of Word Reading Efficiency*. Austin, USA: Pro-Ed.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.
- van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? Cognate effects in a sentence context. *Psychological Science*, *20*, 923–927.
- van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, *9*, 780–789.
- van Heuven, W. J. B., Dijkstra, T., & 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.
- von Studnitz, R. E., & Green, D. W. (2002). The cost of switching language in a semantic categorization task. *Bilingualism: Language and Cognition*, *5*, 241–251.
- Weiß, R. H. (1998). *Wortschatztest (WS) und Zahlenfolgentest (ZF): Ergänzungstests zum Grundintelligenztest CFT 20*. Germany: Hogrefe, Verlag für Psychologie.
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*, 971–979.
- Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. New York, NY: Touchstone Applied Science Associates. Inc. My Book.

Author biographies

Pauline Schröter received her PhD in Educational Science from the Freie Universität Berlin in 2016 and currently works at the Institute for Educational Quality Improvement (IQB) at the Humboldt-Universität zu

Berlin. From 2012 to 2016, she was a research fellow in the research group REaD (Reading Education and Development) at the Max Planck Institute for Human Development Berlin, where she investigated the development of visual word recognition in German bilinguals. Her research focuses on orthographic processing in beginning readers of L2 German as well as in German-English balanced bilingual children.

Sascha Schroeder received his PhD in Psychology from the University of Cologne in 2008 and habilitated in 2011 at the Freie Universität Berlin. After working at the Max Planck Institute for Human Development Berlin and the University of Kassel, as well as holding guest professorships at the University of Potsdam and the Freie Universität Berlin, he established the Max Planck Research Group REaD in 2012. He investigates the cognitive processes underlying written language acquisition and the cognitive determinants of reading development in childhood and adolescence. His research combines longitudinal and experimental approaches and uses linguistic as well as cognitive methods to analyze and simulate developmental data. He is a member of the International Max Planck Research School LIFE and supervises students at both the undergraduate and graduate levels.

Appendix

Nonword stimuli for LDTs in English and German.

English LDT

English-like nonwords

agleered, ath, bealing, bix, bize, burder, cally, cheems, cian, clissars, clobes, coneme, dag, debroo, doy, ducket, eak, evat, famiday, faquid, fengal, foom, foy, fuds, gath, geason, hant, homp, ith, lew, meaves, mized, moice, mook, mourt, municean, muth, nerm, nood, oze, palk, pean, phes, pight, prac, pud, pum, rawn, rean, rike, selfand, sloon, sloor, smic, snirge, snode, soa, soat, sosh, sweel, tady, trawn, urage, wuns

German-like nonwords

bage, (bans), bauns, bips, blossom, borz, bute, dauge, fub, gein, gerl, (hams), heet, hehne, helb, hok, imme, japf, kalmt, kawe, keffe, kihi, laum, leed, meife, mims, nekien, nilete, nis, nuge, ogel, ohl, pafe, pahme, pauner, pazo, pids, plad, (pluck), posel, rahl, rak, relm, rolpe, sittam, sokat, sond, sor, spreime, tuwe, ubu, vakke, wehl, weik, woch, wosen, wotz, wuklimus, wutimer, zach, zaffel, zebel, zise, zotter

German LDT

German-like nonwords

Adrille, Baft, (Bags), Biet, Breif, Dage, Dilastor, Dist, Dite, Dond, Firg, Foge, Folz, Gaflik, (Gan), Gane, Gause, Giel, Goks, Henk, (Herk), Hest, Hiser, Hon, Irf, Junter, Kall, Kast, Kims, Kland, Kontus, Krock, Lans, Lis, Lumt, Mand, Ming, Moge, Nakafe, Noge, Nolf, Pand, Plie, Plin, Pok, Pulser, (Rit), Sarz, Spirm, Spoch, Stapem, Stebs, Stort, Stralt, Susid, Taf, Tam, Tenribel, Tock, Ulsel, Urm, Wond, Wose, Zafel

English-like nonwords

Angic, Awd, Awn, Bacel, Bame, Baw, Benane, Bestus, Bicer, Blee, Boof, Bove, Broaf, Catter, Crind, Dimaster, Fadric, Fank, Fism, Fobe, (Gan), Gice, Gope, (Herk), Hoke, Hud, Hur, Jit, Kide, Lale, Lape, Mish, Mude, Nace, Nesh, Noke, Nole, Nuncer, Nust, Oppriss, Pault, Plun, Ransible, Rish, (Rit), Rosh, Rudic, Sarn, Sath, Shiple, Snosh, Stape, Stewn, Stoth, Stoze, Strile, Tay, Tob, Toke, Vape, Wagh, Wibs, Woil, Woze

Note: Nonwords in parentheses were excluded from the analyses.