Title

The Influence of Stimulus Type on Language Processing in Comprehension

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors

Liu, Hong
Chaouch-Orozco, Adel

Publication Date

2024

Peer reviewed

The Influence of Stimulus Type on Language Processing in Comprehension

Hong Liu (Hong.Liu@xjtlu.edu.cn)
Department of Applied Linguistics, Xi'an Jiaotong-Liverpool University
Suzhou, China

Adel Chaouch-Orozco (adel.chaouchorozco@polyu.edu.hk)
Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University
Hong Kong

Abstract

Numbers and pictures are the two most frequently used types of experimental stimuli in bilingual language control studies. However, the potential qualitative differences in the representation and processing of these stimuli could involve the recruitment of divergent cognitive mechanisms. This paper investigates the influence of stimulus type (numbers vs pictures) on language processing in bilingual comprehension, specifically examining whether semantic connections between numbers impact language switching. We tested Chinese-English-Spanish trilinguals in two cross-modal matching tasks (i.e., a picture-word matching task and a magnitude-number matching task) in the context of the n-2 language switching paradigm. Contrary to the n-2 repetition cost observed in previous studies employing the same paradigm, our findings reveal an n-2 repetition benefit. Crucially, the n-2 repetition effect was observed only with numbers. We discuss the findings in relation to the prevalent language control mechanisms and how lexical associations between numbers may give rise to the observed difference.

Keywords: bilingualism; language control; mental lexicon; language comprehension; number representation

Introduction

Empirical research suggests that bi-/multilinguals (hereafter bilinguals) activate all their languages in language processing, hence the bilingual non-selective processing hypothesis (Costa, 2005). This parallel activation results in the need to apply some sort of control mechanism to prevent non-target languages from interfering during language processing (Declerck & Koch, 2022; Declerck & Philipp, 2015). The most prevalent mechanism proposed in the literature for effectively keeping at bay non-target influence is inhibitory control (Dijkstra & Van Heuven, 1998; Green, 1998). Essentially, in order to process messages in one language, bilinguals must inhibit the unintended language. Alternatively, activation-based accounts have also been proposed to explain how bilinguals process their languages (Blanco-Elorrieta & Caramazza 2021; Philipp, Gade, & Koch, 2007). In these views, inhibition is not necessary; language selection occurs due to the stronger activation of the target language.

The language-switching paradigm is the most commonly used method for investigating bilingual language control. This paradigm includes two variants: n-1 and n-2 language switching. In n-1 language switching, subjects, upon seeing or hearing a cue, switch between two languages to perform a task (e.g., picture naming). Longer response latencies and/or higher error rates are usually observed in switch trials, where subjects respond in a different language from the one used in the previous (n-1) trial. This effect is referred to as switch costs. Importantly, an asymmetry is often reported by which these switch costs are larger in the more proficient language (e.g., Meuter & Allport, 1999; Slevc, Davey, & Linck, 2016), although conflicting results have also been reported (Liu et al., 2019; Mosca & Clahsen, 2016). Asymmetrical switch costs are believed to arise because the more proficient language is activated more strongly. This means that when responding in the less proficient language, there is a need for greater inhibition to suppress the dominant language. Subsequently, when switching back to the more proficient language, it requires more effort to overcome this strong inhibition than when switching from the less proficient to more proficient language.

The procedure in the n-2 language switching paradigm is slightly different, as subjects switch between three languages while performing a task. As such, the analysis entails comparing performance in sequences like AB*A* and CB*A*, where each letter refers to a language to be employed in each subsequent trial (e.g., Chinese-English-Chinese vs Spanish-English-Chinese sequences), being the third trial in the sequence (i.e., A here) the critical one. The key difference between these two types of sequence is whether A is repeated two trials back (i.e., the n-2 trial). Generally, better performance in A is found in CB*A* sequences compared to AB*A* ones. This is often referred to as the n-2 repetition cost. As it occurs with n-1 switch costs, n-2 repetition costs are assumed to be caused by persisting inhibition (Philipp, Gade, & Koch, 2007; Philipp & Koch, 2009).

While both the n-1 and n-2 switching paradigms are mostly used with production tasks, it should be noted that the empirical effects are often diminished or disappear altogether when they are used in comprehension (e.g., Declerck et al., 2019; Declerck & Philipp, 2018).

Influence of Stimulus Type

In both n-1 and n-2 language switching experiments, pictures and numbers are the two most frequently used stimuli. However, the potential *qualitative* differences in the representation and processing of these stimulus types are often overlooked. Consequently, findings derived from using both stimulus types are usually compared without considering these differences. This point is far from trivial. These intrinsic differences could involve the recruitment of

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

distinct cognitive mechanisms during language processing. This situation is concerning, especially when considering the implications of such comparisons for the development of cognitive theories of language control.

Numbers possess key features that set them apart from most picture sets. For instance, numbers constitute a specific semantic category, and their activation may lead to the co-activation of adjacent numbers on a mental number line, potentially influencing language control mechanisms. Moreover, numbers co-occur more frequently than other concepts, and they are usually learned in a sequential order. As such, they are likely to be connected by strong associative links at the lexical level. Consequently, activating a number lexical form may quickly co-activate other related numbers at the word level—even when semantic involvement is not observed (Herrera & Macizo, 2011, 2012; Liu & Chaouch-Orozco, 2023).

Crucially, as noted, these idiosyncratic divergencies between numbers and pictures may have introduced a confound in previous studies employing these stimuli, blurring our understanding of the cognitive mechanisms behind bilingual language control. However, to date, only two studies have investigated this question empirically. Declerck, Koch and Philipp (2012) compared the effect of stimulus type on switch costs in a series of n-1 language switching experiments in which German-English bilinguals named either Arabic numbers or pictures. Their results showed that numbers yielded smaller switch costs compared to pictures—a so-called "number effect."

Digging further into the reasons for this difference, the authors compared numbers with two control picture stimulus sets. Because German and English share many cognate numbers (e.g., *six* and "sechs", *six* in German), one such control set contained pictures labelled by cognates in the two languages. In addition, a set with pictures denoting semantically related concepts was used. In doing so, the authors aimed at balancing off the semantic relationships potentially present in numbers from 1 to 9.

Notably, the initial difference in switch costs between the two types of stimuli remained when numbers and the semantic control set were compared. However, the difference disappeared when numbers were contrasted with the cognate set. Thus, the authors concluded that the difference between numbers and pictures originated at the phonological level and was caused by the cognate status of the numbers in German and English. In this context, naming a cognate number would lead to the phonological activation of numbers in the other language. Therefore, when switching to that language, this phonological co-activation would facilitate switching, resulting in the reduction of the switch cost.

To confirm that the stimulus difference was not driven by semantics, the authors conducted further analyses to test whether switch costs varied for numerical concepts that are more closely related (i.e., a numerical distance effect; Brysbaert, 1995). The results revealed larger switch costs for numbers that are more closely related, which contrasted with smaller switch costs for numbers in comparisons with pictures. Based on these findings, Declerck, Koch and Philipp (2012) argued that the difference in switch costs for stimuli was not semantically driven. Yet, the presence of cognates was an inherent limitation of their study, leaving an open question: Would the same pattern of results be observed in languages not related phonologically?

This question was pursued in Liu and Chaouch-Orozco's (2023) follow-up study. The authors investigated Arabic number and picture naming in Chinese and English, two languages where cognates are absent from their numbers. Importantly, their results revealed a difference between numbers and pictures in switch costs in the opposite direction than in Declerck, Koch and Philipp (2012). That is, their data revealed that digits yield *larger* (and not smaller) switch costs.

Having ruled out a phonological explanation for this result given the controlled experimental setup, the authors further examined whether the larger switch costs with numbers were driven by semantics. If so, the effect should be even larger for numerical concepts that are more related (i.e., a numerical distance effect). Crucially, the authors did not observe such an effect, concluding that semantic co-activation was not present during Arabic number naming.

Instead, to explain their results, Liu and Chaouch-Orozco resorted to word-level associations. More specifically, in their view, strong associations are established between the numbers in the proficient language(s). When a given number is being processed, these strong links lead to the rapid co-activation of other numbers from the same language at the lexical level, making switching away from this language more effortful.

At this point, it should be noted that despite the valuable insights gained from production studies, a significant gap persists in our understanding of bilingual language processing, particularly regarding language comprehension (Declerck & Koch, 2022). While production tasks have offered substantial evidence on bilingual language control, they do not entirely capture the full complexity of bilingual language processing. Therefore, investigating the nature of the number effect in the context of language comprehension represents a unique opportunity to shed light on essential yet currently unexplored aspects of the interaction between different cognitive mechanisms in the bilingual mind.

The Present Study

We investigate the influence of stimulus type in bilingual comprehension, specifically examining whether semantic connections between numbers impact language switching. Importantly, instead of relying on the more traditional n-1 language switching paradigm, here we adopt the n-2 version, as it has been shown to more consistently yield n-2 repetition costs in multiple studies (e.g., Declerck et al., 2015; Declerck & Philipp, 2018; Philipp, Gade, & Koch, 2007). As such, recent discussions have started to regard the n-2 paradigm as a more reliable indicator of bilingual language control (Declerck & Koch, 2022; Declerck & Philipp, 2015).

We tested Chinese-English-Spanish trilinguals in two cross-modal matching tasks. This task has been previously used to investigate language control in bilingual comprehension (Jiao et al., 2021; Jiao et al., 2022). In it, subjects see visual stimuli (e.g., a picture) and listen to a word with the goal of determining if they match. To ensure semantic processing, the visually displayed numerical items were presented with magnitudes, represented by dots. Crucially, our materials did not include any cognates.

By ensuring that both number and picture processing involves semantics without including any phonological confounds, our design effectively tests whether the semantic relationship between numbers influences language control mechanisms. If semantics play a role, and given the inherent semantic relationships between numbers, we expect to observe differences in the n-2 repetition effect based on the type of stimulus, with no clear expectations as per the directionality of the effect. Conversely, similar n-2 repetition effects should be observed for both numbers and pictures if the semantic relationships between numbers do not lead to differences in processing.

Methods

Participants

The study was approved by the ethics committee at XX University (university name and approval number anonymised for review). All participants gave informed consent in accordance with the Declaration of Helsinki. Thirty-seven Chinese-English-Spanish trilinguals (*Mage* = 21.52, *SD* = 2.37) from several universities in mainland China participated. Chinese was their dominant language and the language of daily communication. They were undergraduate students in Spanish, and they had learned English as a compulsory subject at school before entering university. They were exposed to Spanish daily in the classroom. During the first two years of university, they had two hours of English class each week as part of the mandatory requirement. After that, English was optional. Before the experiment, we measured the participants' language proficiency in each language with the Multilingual Naming Test (MINT; Gollan et al., 2012). The MINT is a standardised naming test where participants name 68 pictures of varying word frequencies. It has been validated as a reliable measure for capturing variance in bilinguals' language proficiency and language dominance in English, Spanish, and Mandarin (Gollan et al., 2012; Ivanova, Salmon, & Gollan, 2013; Sheng, Lu, & Gollan, 2014). Participants completed the MINT first in Chinese, then in English, and lastly in Spanish.

The MINT mean scores in Chinese, English, and Spanish were 60.82 (*SD* = 2.43), 38.67 (*SD* = 2.47), 29.67 (*SD* = 10.17), and they significantly differed across the three languages (L1 vs L2, $t = 34.08$, $p < 0.001$; L1 vs L3, $t = 18.00$, $p < 0.001$; L2 vs L3, $t = 4.74$, $p < 0.001$), indicating that the participants' proficiency in the three languages correlated with their order of acquisition. The subjects completed the Language History Questionnaire 3 (LHQ3, Li et al., 2020) too. Information on participants' language use and proficiency is presented in Table 1.

Tasks and Materials

The participants completed two tasks. In task 1, they matched pictures and words, whereas in task 2, they performed magnitude-number matching. The order of the two tasks was counterbalanced. In both cases, the participants listened to a word or a digit number and decided whether a picture or magnitude expressed by dots on the screen matched the heard stimuli. For the magnitude-digit matching task, there were six numerical stimuli (2, 4, 5, 7, 8, 9), selected to avoid cognate words between Spanish and English (e.g., six and "seis," Spanish for six). For the picture-word matching task, six pictures depicting common entities were taken from MultiPic (Duñabeitia et al., 2018). The nouns labelling each of these entities in the three languages were non-cognates and were comparable in word frequency and the number of syllables. The frequencies for the Chinese, English and Spanish names for the pictures were based on SUBTLEX-CH (Cai & Brysbaert, 2010), SUBTLEX-UK (van Heuven et al., 2014), and SUTBLEX-ESP (Cuetos et al., 2011), respectively. The pronunciation of the stimuli words was created on the website https://soundoftext.com/, using Mandarin, British English, and Peninsular Spanish.

There were six blocks in each task and 36 trials in each block. Each number/picture word was repeated 12 times in each language as audio stimuli; each magnitude/picture was repeated 36 times as visual stimuli. The sequence of trials was pseudorandomized across language, stimulus type, language sequence (e.g., ABA vs CBA), and answer type (match vs non-match). Immediate repetition of a language and immediate repetition of stimuli (both visual and audio) either in the subsequent trial or the one following it were not allowed.

Procedure

The experiment was created and presented on Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). After providing their consent, the participants performed an audio check, in which they listened to a few words and adjusted the sound volume. Then, a brief task description in Chinese was presented, emphasising the need to respond as accurately and quickly as possible. The participants were then presented

Table 1. Participants' language use and proficiency information.

	L1 Chinese	L2 English	L3 Spanish
Years of learning/using	21.06 (2.54)	13.64 (3.14)	3.30 (2.04)
Self-reported proficiency (on a scale of 10)	9.55 (0.60)	6.77 (1.00)	6.35 (1.37)
MINT scores	60.82 (2.43)	38.67 (2.47)	29.67 (10.17)

with 12 practice trials and were allowed to repeat the practice as many times as needed. Right-handed subjects had to press "0" to indicate "match" and "1" for "non-match". The order was inverted for left-handed participants. Each trial started with a fixation point (600 ms), followed by the stimulus. The stimulus remained on the screen until a response (key press) was given. The participants were allowed to rest between blocks.

Data Analysis

The data and analysis code can be found in the OSF repository (https://osf.io/8kjgr/?view_only=dd0817d553794a34830cbb25843b7322). Importantly, different measures were taken to ensure the quality of the data. First, participants' responses were carefully examined to ensure they were not blatantly random. Second, we controlled for the total length of the experiment to avoid including data of participants who took relatively excessive time to complete the tasks—which could indicate they were distracted while completing the experiment. We calculated the average and the standard deviation time taken to finish (i) each task separately and (ii) the two tasks combined. Then, we excluded subjects who took more than the average plus two standard deviations to complete each and/or the two tasks. Following this strict exclusion criteria, one participant's data were removed. Thus, our final data contained responses from 36 subjects for each of the two tasks.

For each participant, response times 2.5 *SD* below and above their mean were removed. This resulted in the removal of 392 observations (out of 14154; 2.77% of the data). We further removed 2 observations below 200 ms (0.02 % of the data) that could hardly reflect a conscious response (e.g., Baayen & Milin, 2010).

In the RT analysis, incorrect responses were removed. Response times were inverse- and log-transformed. Q-Q plots and Shapiro-Wilk tests indicated that the inverse transformation provided a better skewness correction.

We analysed RTs and error rates with (generalized) linear mixed-effects models (Baayen et al., 2008) using R (version 3.6.1; R Core Team, 2021) with the lme4 package (Bates et al., 2015). For each analysis, a maximal model was fitted. In the case of non-convergence, the random factor that explained the least variance was removed until the model converged. We further checked models' assumption (which were met in all cases). The fixed effects in our models included main effects and interactions of interest (Brauer & Curtin, 2018). The grouping factors were trial (i.e., repetition v. non-repetition), stimulus type (i.e., number words vs picture words), and their interactions. We specified the full structure with random intercepts and random slopes for subjects and each grouping factor. All contrasts for the fixed factors with two levels assessed the difference between the two levels of each factor (coded as –0.5, 0.5). For the accuracy analysis, we dummy-coded the variable (1 for "correct" and 0 for "incorrect) and employed generalized linear mixed-effects models with a binomial family fit.

Results

The response times and error rates can be seen in Table 2. Overall, the main effects were significant for trial (β = .00, t = 6.05, p < .001) and stimulus type (β = .00, t = 8.78, p < .001). The response times in the repetition trials, contrary to previous reports, were faster compared to the non-repetition trials. In other words, we observed an n-2 repetition benefit effect. Additionally, response times were faster with pictures compared to numbers. The interaction between trial and stimulus type was also significant (β = -.00, t = -4.42, p < .001), illustrated in Figure 1. When looking at the post-hoc pairwise comparisons, we observed that the effect reached significance with the numerical stimuli only (β = -.00, t = -7.45, p < .001), but not with the pictures (β = -.00, t = -1.15, p = 1.00). In the error rates analysis, only the main effect of stimulus type approached significance (β = -0.20, z = -1.78, p = .07)**.**

Table 2. Mean response times (RTs, in milliseconds; standard errors), error rates (%), and n-2 repetition effects (in milliseconds)

	N-2 repetition		N-2 non-repetition		
Stimulus	RT	Error rate	RT	Error rate	N-2 repetition effect
Numbers	1050 (439)	2.21%	1109 (484)	1.93%	-59*
Pictures	1018 (400)	2.16%	1022 (396)	2.72%	-4

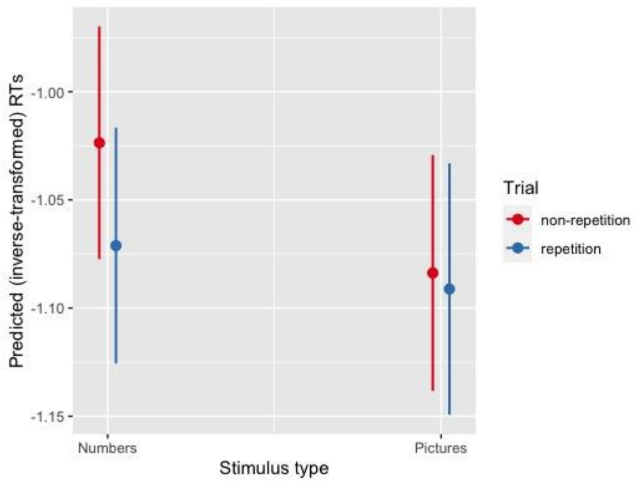

Figure 1: The interaction between trial and stimulus type.

Discussion

In the present study, we investigated the influence of stimulus type (numbers vs pictures) on language control in bilingual comprehension. To do so, we employed a cross-modal matching task with the n-2 language switching paradigm and tested Chinese-English-Spanish trilinguals. Contrary to the n-2 repetition costs observed in previous studies employing the same paradigm (Babcock & Ballesi, 2015; Declerck & Philipp, 2018; Philipp, Gade, & Koch, 2007; Timmer et al., 2018), our findings reveal an n-2 repetition benefit. Our participants responded significantly faster in the n-2 repetition trials as compared with the switch trials. Importantly, this effect was modulated by stimulus type, such that it was only observed with the numerical items.

The first thing to note about the present findings is that they do not align with the predictions posed by inhibition-based accounts. To accommodate our results, we resort to persisting activation accounts, as proposed by, for example, Koch et al. (2010) and Philipp, Gade and Koch (2007). In this light, the observed benefit in trial n for language A, would stem from activation in trial n-2. Specifically, responding in trial n-2 activates language A beyond its baseline level. This additional activation, not being suppressed upon switching to language B, takes time to dissipate. Consequently, when participants respond in trial n to language A again, this residual activation facilitates performance compared to the situation where such persisting activation is absent (i.e., in the CB*A* sequence).

With regard to our primary research question concerning the number effect, our results indicate that the n-2 repetition effect was observed with numbers but not with pictures. This discrepancy invites two interpretations. On the one hand, (i) the observed difference may be driven by semantic relationships between the numbers. If so, the digit effect should be larger when numbers are in close proximity. Alternatively, as recently suggested by Liu and Chaouch-Orozco (2023), (ii) numbers may present unique associative relationships due to their co-occurrence across many different contexts (Herrera & Macizo, 2011, 2012). Moreover, the associations may not be limited to numbers in close numerical proximity, as suggested by Herrera and Macizo (2012). Instead, they could extend to all numbers within the natural sequence from 1 to 10.

To examine the semantic hypothesis (i), we conducted an additional analysis with trial and numerical distance as fixed factors on response times. In line with Declerck, Koch and Philipp (2012), numerical distances of three or less were considered small, while those larger than three were deemed large. Our analysis found no significant interaction between these factors ($\beta = .00$, $t = 1.29$, $p = .20$), indicating no distance effect. In other words, the n-2 repetition effect did not differ as a matter of the semantic relationship among adjacent numbers.

Given the lack of a numerical distance effect, it seems more plausible to attribute our findings to associative links between numbers, particularly within the 1 to 10 range (see Liu & Chaouch-Orozco, 2023, for further discussion). In the current n-2 language switching task, naming a number in language A during trial n-2 activates related numerical concepts through associative links, particularly within that same language. This activation is not suppressed even during a switch to language B in the subsequent trial, and it persists until language A is required again in trial n. Consequently, this task demands less cognitive effort compared to processing the same sequence with unrelated pictures, hence, explaining the number effect we obtain.

In summary, our findings indicate that the cognitive mechanisms underlying the processing of picture and number stimuli are likely to differ qualitatively, aligning with prior research with different emphases (Herrara & Macizo, 2011, 2012; Liu & Chaouch-Orozco, 2023). In this context, it is crucial to recognise that the conclusions based on specific types of stimuli might not be widely applicable due to the unique nature of these stimuli. Thus, caution must be exercised when comparing results from different stimuli, and further research is imperative to explore how the cognitive processes related to language control, and language processing more broadly, vary with each stimulus type.

Acknowledgments

This paper was supported by the Research Development Fund (RDF) awarded by Xi’an Jiaotong-Liverpool University (grant number: RDF-22-01-026) to both authors.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52,* 388–407.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*, 12-28.

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.

Babcock, L., & Vallesi, A. (2015). Language control is not a one-size-fits-all languages process: Evidence from simultaneous interpretation students and the n-2 repetition cost. *Frontiers in Psychology, 6,* 1622.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1 - 48.

Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition, 213,* 104625.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23,* 389–411.

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General, 124,* 434–452.

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS One, 5,* e10729.

Costa, A. (2005). Lexical access in bilingual production. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic Approaches*. New York: Oxford University Press.

Cuetos, F., Glez-Nosti, M., Barbón, A., & Brysbaert, M. (2012). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica, 33,* 133-143.

Declerck, M., & Koch, I. (2022). The concept of inhibition in bilingual control. *Psychological Review, 130,* 953-976.

Declerck, M., Koch, I., Duñabeitia, J. A., Grainger, J., & Stephan, D. N. (2019). What absent switch costs and mixing costs during bilingual language comprehension can tell us about language control. *Journal of Experimental Psychology: Human Perception and Performance, 45,* 771-789.

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review, 22,* 1630-1645.

Declerck, M, Koch, I, & Philipp, A. M. (2012). Digits vs. pictures: The influence of stimulus type on language switching. *Bilingualism: Language and Cognition, 15,* 896–904.

Declerck, M., & Philipp, A. M. (2018). Is inhibition implemented during bilingual production and comprehension? N-2 language repetition costs unchained. *Language, Cognition and Neuroscience, 33,* 608-617.

Declerck, M., Thoma, A. M., Koch, I., & Philipp, A. M. (2015). Highly proficient bilinguals implement inhibition – Evidence from n-2 language repetition costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1911-1916.

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist Connectist Approaches to Human Cognition*. Mahwah, NJ: Lawrence Erlbaum.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology, 71,* 808–816.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition, 15,* 594–615.

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

Herrera, A., & Macizo, P. (2011). Naming digits in a semantic blocking paradigm. *Quarterly Journal of Experimental Psychology, 64,* 328–338.

Herrera, A., & Macizo, P. (2012). Semantic processing in the production of numerals across notations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 40–51.

Ivanova, I., Salmon, D. P., & Gollan, T. H. (2013). The multilingual naming test in Alzheimer's disease: clues to the origin of naming impairments. *Journal of the International Neuropsychological Society, 19,* 272-283.

Jiao, L., Duan, X., Liu, C., & Chen, B. (2022). Comprehension-based language switching between newly learned languages: The role of individual differences. *Journal of Neurolinguistics, 61,* 101036.

Jiao, L., Liu, C., de Bruin, A., & Chen, B. (2020). Effects of language context on executive control in unbalanced bilinguals: An ERPs study. *Psychophysiology, 57,* Article e13653.

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin & Review, 17,* 1–14.

Li, P., Zhang, F., Yu A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition, 23,* 938-944.

Liu, H., & Chaouch-Orozco, A. (2023). Is the digit effect a cognate effect? Digits (still) differ from pictures in non-phonologically mediated language switching. *Bilingualism: Language and Cognition, 26,* 469-475.

Liu, C., Jiao, L., Wang, Z., Wang, M., Wang, R., & Wu, Y. J. (2019). Symmetries of bilingual language switch costs in conflicting versus non-conflicting contexts. *Bilingualism: Language and Cognition, 22,* 624-636.

Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: evidence from switching language-defined response sets. *European Journal of Psychology of Education, 19,* 395–416.

Philipp, A. M., & Koch, I. (2009). Inhibition in language switching: What is inhibited when switching between languages in naming tasks? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1187–1195.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25–40.

Declerck, M., Koch, I., & Philipp, A. M. (2012). Digits vs. Pictures: The influence of stimulus type on language switching. *Bilingualism: Language and Cognition, 15,* 896–904.

R Core Team (2021) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/.

Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in Mandarin-English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition, 17,* 364–383.

Slevc, L. R., Davey, N. S., & Linck, J. A. (2016). A new look at "the hard problem" of bilingual lexical access: evidence for language-switch costs with univalent stimuli. *Journal of Cognitive Psychology, 28,* 385-395.

Timmer, K., Calabria, M., Branzi, F. M., Baus, C., & Costa, A. (2018). On the reliability of switching costs across time and domains. *Frontiers in Psychology, 9,* 1032.

van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67,* 1176-1190.