

Research Article 

**Cite this article:** Timmer, K., Wolna, A. and Wodniecka, Z. (2024). The impact of cues on language switching: do spoken questions reduce the need for bilingual language control? *Bilingualism: Language and Cognition*, 1–19  
<https://doi.org/10.1017/S1366728924000841>


Received: 04 March 2024  
 Revised: 16 September 2024  
 Accepted: 20 September 2024

**Keywords:**

bilingualism; conversational language switching; bilingual language control; switch costs; reversed language dominance



**Corresponding author:**

Kalinka Timmer;  
 Email: [k.timmer@uw.edu.pl](mailto:k.timmer@uw.edu.pl)

 This article has earned badges for transparent research practices: Open Data. For details, see the Data Availability Statement.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

# The impact of cues on language switching: do spoken questions reduce the need for bilingual language control?

Kalinka Timmer<sup>1,2</sup> , Agata Wolna<sup>2,3,4</sup>  and Zofia Wodniecka<sup>2</sup>

<sup>1</sup>Faculty of Psychology, University of Warsaw, Warsaw, Poland; <sup>2</sup>Psychology of Language and Bilingualism Lab, Jagiellonian University, Krakow, Poland; <sup>3</sup>Department of Brain and Cognitive Sciences, MIT, Cambridge, MA, USA and <sup>4</sup>McGovern Institute for Brain Research, MIT, Cambridge, MA, USA

**Abstract**

The classical language switching paradigm using arbitrary cues to indicate the language to speak in has revealed switching between languages comes at a cost (i.e., switch cost) and makes one slower in the first than in the second language (i.e., reversed language dominance). However, arbitrary cues can create artificial requirements not present during everyday language interactions. Therefore, we investigated whether simulating elements of real-life conversations with question cues ('Co?' versus 'What?') facilitates language switching in comparison to the classical paradigm (Experiment 1: red versus blue outline; Experiments 2 and 3: low versus high tone). We revealed a dissociation between the two indices of language control: (1) question cues, compared to arbitrary cues, reduced switch costs but (2) did not modulate (in Experiment 1) or increase the reversed language dominance (Experiments 2 and 3). We propose that this conversational switching paradigm could be used as a conceptually more 'true' measure of language control.

**Highlights**

- The classical language-switching paradigm shows language switching comes at a cost.
- Why do bilinguals switch languages in everyday conversation with seeming ease?
- Do spoken question cues reduce the need for bilingual language control in the lab?
- Auditory question cues facilitate language selection and reduce switch costs.
- Conversational switching paradigm might be a more 'true' measure of language control.

**1. Introduction**

Bilinguals have two words for each concept, one in each language. Therefore, when bilinguals want to produce even a simple utterance such as 'strawberry ice cream', they need a mechanism to help select the words in the correct language and avoid intrusions from the unintended language (e.g., Gollan et al., 2011). This mechanism has been referred to as Bilingual Language Control (BLC) (Christoffels et al., 2007; Green, 1998). Researchers often use the classical language switching paradigm in which pictures are named in the first (L1) and second language (L2) based on arbitrary cues (e.g., colored frames around the pictures), to measure BLC. During speech production in this paradigm, switching from one language to another incurs a cost over staying in the same language (Bobb & Wodniecka, 2013). Yet, bilinguals easily switch between their languages during everyday conversations, where speech production does not occur in isolation but as part of a conversation between people. One possible explanation for greater difficulties in switching between languages in the classical switching paradigm over seemingly effortless switches in everyday conversation is that the arbitrary cues in the classical language switching paradigm create additional artificial requirements for the cognitive system, which are absent in everyday language interaction.

Although the task-switching literature has investigated the impact of cues on switch costs (Arrington et al., 2007; Arrington & Logan, 2004; Grange & Houghton, 2010), to the best of our knowledge, only two studies compared the impact of cues on language switching (Blanco-Elorrieta & Pykkänen, 2017; Liu et al., 2019). Both studies used faces that were culturally congruent with the language but revealed inconsistent findings. Blanco-Elorrieta and Pykkänen (2017) showed that the switch cost disappeared when faces were presented as cues (compared to color cues), whereas Liu et al. (2019) revealed a reduction (but not absence) of the L1 switch cost and no impact on the L2 switch cost. However, faces might not be consistently linked to a language to be spoken. Therefore, we wanted to use a cue that has a very obvious preexisting association with both the specific languages and the goal of communicating through speech, while still being comparable to the controlled experimental manipulation with color cues. We chose auditory cues formulated as questions (e.g., 'What?'), which allows merging

comprehension and production in one task, which we will refer to as the conversational language switching paradigm. This introduces two essential elements of real-life conversations: (1) a question-and-answer interaction and (2) the language of an answer already present in the preceding question. At the same time, this modulation preserves the controlled lab settings necessary to investigate BLC.

We predicted the conversational paradigm would reduce the language switch cost compared to artificial cues typically used in the classical language switching paradigm. This is because the conversational paradigm presents some elements of the naturalistic setting in which speech is produced and switched between. We test the replicability of switch cost reduction due to auditory question cues in three experiments with large sample sizes (see Nosek et al., 2022 for the importance of replication for theory forming).

### 1.1. BLC mechanisms

Within the classical language switching task, typically employed in bilingualism research, two phenomena have played an essential role in informing about BLC mechanisms: the switch cost and the reversed language dominance effect (Bobb & Wodniecka, 2013; Declerck & Koch, 2022). The switch cost is reflected by slower naming on a language switch than repeat trial (Costa et al., 2006; Meuter & Allport, 1999; Philipp et al., 2007; Timmer et al., 2017). Reversed language dominance is reflected by slower naming in the dominant than in the non-dominant language (i.e., for our [and most] participants, respectively, L1 and L2) when the two languages are intermixed. Slower naming in the dominant than non-dominant language under the language mixing condition was present in 81.25% of cued picture naming studies with younger adults included in a recent meta-analysis (Goldrick & Gollan, 2023; calculated based on their open source data). This so-called reversed dominance contrasts consistently faster naming in the dominant than non-dominant language when only one language is at play (Christoffels et al., 2007; Costa & Santesteban, 2004; de Groot & Christoffels, 2006; Timmer et al., 2018; also called global slowing of L1). These two phenomena have been hypothesized to differ in the type of bilingual control involved: local and global control (Christoffels et al., 2007; Green, 1998; Timmer et al., 2018). Although the switch cost is thought to reflect local control during which language activation is reactively adjusted on a trial-by-trial basis, reversed language dominance is believed to reflect a more global and proactive adjustment of the relative language activations as a whole (Bobb & Wodniecka, 2013; Christoffels et al., 2007; Timmer, Christoffels et al., 2018).

### 1.2. The impact of cues on language switching

In the current paper, we investigate whether the observed local language control (i.e., switch costs) is partly a side effect of laboratory conditions that create artificial obstacles that increase switch costs compared to language switching in a natural context. Some researchers turned to study voluntary rather than cued switching to explore whether it reduces switch costs (e.g., de Bruin et al., 2018). However, most research has focused on the forced language-switching paradigm. From the latter category, almost all published studies (57/60) use visually presented cues to indicate the language to speak in (see [Appendices A1](#) and [A2](#)). Only seven out of the 60 cued language switching studies used word cues, of which only three auditory cues, like ‘What?’ (Tarlowski et al., 2012) or ‘say’ (Hernandez et al., 2000; Hernandez & Kohnert,

1999). Others implemented visual cues (i.e., flags, faces, or cultural objects) that were considered relatively ecological (Costa et al., 2008; Timmer, Wodniecka et al., 2021; Woumans et al., 2015). For example, suppose we know a particular person speaks Spanish. In that case, their face will activate Spanish over other languages the bilingual knows (Woumans et al., 2015). Naming culturally biased images in a culturally congruent language makes it more accessible than in an incongruent language (Jared et al., 2013). A Bengali background (iconic cultural images representing Bengali culture) also impedes naming pictures in English (Roychoudhuri et al., 2016). Face cues modulated local (switch cost) but not global (reversed language dominance) control. For the switch cost, Blanco-Elorrieta and Pylkkänen (2017) showed it disappeared when culturally congruent faces were presented compared to arbitrary color cues. However, others revealed that a Chinese face (familiar race) or a culturally biased picture over color cues (e.g., Chinese food) only helped Chinese – English bilinguals to switch back to their L1 but not L2 (Liu et al., 2019, 2021). Furthermore, when faces of virtual interlocutors (Peeters, 2020; Peeters & Dijkstra, 2018) were used, the switch cost size was similar in the two languages. One reason the results using faces as cues might be inconclusive is that faces are not always unambiguously related to a given language; hence, a more straightforward cue should be used to test whether the BLC engagement is reduced with the more context-congruent (as opposed to arbitrary) cues. One possibility is cues that mimic a conversation, i.e., auditory question cues, which in natural situations elicit speech production (response) in the interlocutors. Therefore, we used auditory question cues to initiate speech production and compared them to arbitrary cues used in a classical switching paradigm.

### 1.3. The interaction of bilingual language comprehension and production

The idea that comprehension and production processes are not entirely separable has already been formulated in monolingual theories of language processing. Pickering and Garrod (2013) proposed an integrated theory of language comprehension and production, suggesting they are interwoven, which enables people to predict their own and others’ speech and perform joint speech actions. The close interaction between language comprehension in production in conversation has also been confirmed by the neuroimaging experiments (Castellucci, Guenther et al., 2022; Castellucci, Kovach et al., 2022). To the best of our knowledge, only a handful of studies have tackled bilingual control mechanisms in a situation that more closely approximates bilingual dialogue (Gambi & Hartsuiker, 2016; Liu et al., 2018; Zhang et al., 2020). For example, the so-called joint switching paradigms studied the impact of a second person present during a task. In such a paradigm, two bilinguals name pictures while alternating between languages based on arbitrary cues. The results obtained in such paradigms demonstrate that when two different people name consecutive trials in different languages (between-speaker), a similar switch cost is observed when the same bilingual names two consecutive trials in different languages (within-speaker, as in the classical paradigm). This suggests that listeners plan their speech while listening to an interlocutor, which impacts bilinguals’ language selection during production. However, the results do not reveal how we switch languages in response to cues that are conversational in nature, as the switching occurred based on arbitrary cues (e.g., colors). Therefore, we make an explicit comparison between two types of cues and predict that the use of spoken questions reduces language switch

costs, revealing a more ‘true’ local language control mechanism. Furthermore, we will investigate why the reversed language dominance effect might be absent.

The theoretical implication of using an auditory question cue (compared to a classical visual cue) is that it triggers (1) the language the question is asked in and (2) a communication goal to name the picture you see. Both these factors make the studied phenomena more naturalistic (element of a conversation) and more transparent (i.e., language cue explicitly triggers a response in the same language) than non-language cues.

#### 1.4. Cue (to task) transparency

The greater the cue transparency to the task at hand, the more directly the cue stimulates the relevant working memory representation required to perform the task, which has been repeatedly shown in the non-linguistic task-switching literature (Arrington et al., 2007; Arrington & Logan, 2004; Grange & Houghton, 2010). Moreover, it has been suggested that preexisting associations between the cue and the task increase cue transparency (Logan & Schneider, 2006). For example, when words such as “low” and “high” are used as cues in a task requiring participants to decide if a given number is  $<5$  or  $>5$ , they are faster and more accurate than when random letters are used as cues. Similarly, we predict language questions implicitly activate the language to be spoken in the picture naming task.

Cue transparency depends not only on the properties of the cue but also on the task to be performed and is, therefore, not a fixed property. For example, although some faces (e.g., Joe Biden) can be a cue to switch to a particular language, faces of unknown or bilingual people (Spanish-Catalan population) have no direct connection to one language. Flags are arbitrary color combinations linked to specific countries but not necessarily to the act of speaking in that language. Thus, although the transparencies of faces as cues depend on the context, hearing a language in one’s environment is a very direct cue to name objects in that language (Grosjean, 2001; Timmer, Christoffels et al., 2018). The English written phrase ‘to speak’ is a transparent cue to activate this language for the consequent task to be performed (Hernandez et al., 2000; Tarłowski et al., 2012). Auditory question cues could be even more transparent than faces and flags, as they indicate the language to speak in and have a preexisting association with answers that commonly occur in daily conversations. Introducing an element of a conversation makes a paradigm more ecological, as such cues typically occur in daily life.

#### 1.5. The current conversational study

Our overarching aim was to investigate whether auditory language questions facilitate language selection and reduce the need for BLC compared to arbitrary color cues used in the seminal switching study by Meuter and Allport (1999), and most frequently in the literature (see Appendix A2). Therefore, in three experiments with large sample sizes, we compared participants’ performance in two language-switching paradigms: classical (with arbitrary cues) versus conversational (with question cues). We predicted a more consistent reduction of BLC than research that has been demonstrated with faces and cultural objects until now, because language is a more transparent cue.

Crucially, we predicted local but not global BLC mechanisms to be modulated by the type of cues presented during the switching paradigm. We predicted reduced local control (as indexed by a smaller switch cost) during the conversational paradigm compared

to the classical paradigm. Questions could more easily trigger the response to name a picture in the appropriate language, because the association between languages and between questions and answers is deeply entrenched by real-life language use. Therefore, simulating real-life language use could facilitate switching to the appropriate language. In contrast, based on previous findings (Liu et al., 2019, 2021), we did not predict modulations of global control (as indexed by reversed language dominance) with different cue types. The reversed language dominance is typically found when languages are intermixed and seem to be driven by a relative difference in the default activation level between the two languages (Costa et al., 2006; Costa & Santesteban, 2004; Goldrick & Gollan, 2023). In both switching paradigms, the two languages are intermixed. Therefore, the reversed language dominance should be of the same magnitude, regardless of the cueing paradigm.










All three experiments in the present study consisted of two cue paradigms: the classical language-switching paradigm and the modified conversational language-switching paradigm. The latter was the same in all three experiments, with the questions ‘Co?’ and ‘What?’ cueing Polish and English, respectively. In Experiment 1, we compared this to the classical switching paradigm, which most often uses colored outlines (e.g., blue and red) to indicate which language to name the picture in. Although we expected switch costs to be reduced for question cues, we did not have exact predictions on how question cues would impact the speed of response compared to color cues. Previous research showed that responses to cues in the auditory versus visual domain (i.e., tone versus asterisk) were slower in the auditory domain (Lukas et al., 2010). Therefore, auditory language cues could, on the one hand, hamper the speed of processing but, on the other hand, facilitate language selection. To dissociate the effects associated with the modality (auditory) from the cue transparency (questions), in Experiments 2 and 3, we compared the question cues to tones, as they are both in the auditory domain. Here we predicted questions would facilitate overall naming compared to tones and the reduced switch cost. Although Experiments 1 and 2 were run online, Experiment 3 replicated Experiment 2 in the laboratory (see Figures 1 and 2).

## 2. Methods

### 2.1. Participants

Polish native speakers currently residing in Poland, for whom English was the second language with good proficiency, were invited to take part in the study after a short pre-selection questionnaire about their language background and proficiency. They all had normal or corrected-to-normal vision and no history of neurological impairments or language disorders. Everyone was paid for their participation in the main study. The online participants were paid for their participation in the pre-selection regardless of whether they were invited to the main experiment.

A total of 260 participants were divided between three experiments as follows: (1) 80 took part in the online Color versus Question experiment, (2) 80 in the online Tone versus Question experiment and (3) 100 in the lab Tone versus Question experiment (see Table 1 and Figure 1). In Experiment 1, 5 of the 80 participants were excluded from the analysis: 3 due to high error rates ( $>25\%$  incorrect responses or hesitations (‘eesh’), 1 participant did not pass the attention checks and 1 did not finish the experiment. This left 75 participants (12 females) in the online Color versus Question experiment (average age: 24.43 years;  $SD = 3.22$ ). In Experiment 2, 5

Experiment 1	 online	 color cue task	 'What?' question cue task
Experiment 2	 online	 tone cue task	 'What?' question cue task
Experiment 3	 lab	 tone cue task	 'What?' question cue task

**Figure 1.** A schematic overview of the differences and similarities between the three experiments. The target pictures and languages were the same across all experiments.

**Table 1.** Mean answers (and standard deviations) on the self-rating language proficiency and social-economic background questionnaire

		Experiments		
		online C versus Q ( <i>N</i> = 76)	online T versus Q ( <i>N</i> = 75)	lab T versus Q ( <i>N</i> = 92)
Polish	Speaking <sup>a</sup>	8.65 (1.98)	8.51 (1.64)	9.72 (0.73)
	Listening <sup>a</sup>	8.96 (1.82)	8.97 (1.49)	9.83 (0.52)
	Reading <sup>a</sup>	9.33 (1.15)	9.25 (1.23)	9.84 (0.45)
	Writing <sup>a</sup>	8.59 (1.79)	8.44 (1.55)	9.61 (0.99)
	Passive use <sup>b</sup>	62.27 (20.02)	61.04 (21.72)	58.46 (18.10)
	Active use <sup>b</sup>	78.80 (17.55)	79.33 (21.62)	80.86 (14.52)
English	Age of Acquisition	6.72 (2.25)	6.65 (2.41)	6.10 (1.96)
	Speaking <sup>a</sup>	7.29 (1.52)	6.84 (1.71)	7.36 (1.39)
	Listening <sup>a</sup>	7.80 (1.37)	8.16 (1.34)	7.62 (1.39)
	Reading <sup>a</sup>	8.49 (1.11)	8.41 (1.20)	8.22 (1.29)
	Writing <sup>a</sup>	7.33 (1.41)	7.09 (1.54)	7.30 (1.44)
	Passive use <sup>b</sup>	36.31 (20.14)	37.23 (19.88)	37.57 (17.48)
	Active use <sup>b</sup>	20.16 (16.44)	19.20 (20.36)	16.68 (13.25)
	LexTALE	73.85 (10.60) <sup>c</sup>	72.52 (10.42)	72.06 (11.33) <sup>c</sup>

<sup>a</sup>A 10-point scale with 1 point being the lowest and 10 points being the highest self-rated proficiency.

<sup>b</sup>Percentage of time using a specific language in a normal week. The total does not always add up to 100% due to the potential usage of a third language.

<sup>c</sup>Due to technical problems, the LexTALE score of one participant is missing from the average for the lab tone versus question experiment and three from the online color versus question experiment.

of the 80 participants<sup>1</sup> were excluded: 3 due to technical problems and 2 due to high error rates (>25% incorrect responses). This left 75 participants (30 females) in the online Tone versus Question experiment (average age: 23.45 years; SD = 4.48). In Experiment 3, 8 of the 100 participants were excluded: 3 due to technical problems, 4 due to high error rates (>25% incorrect responses) and 1 participant resigned from the study. This left 92 participants (68 females) in the lab Tone versus Question experiment (average age: 22.66 years; SD = 4.16).

Subjective language experience was measured with a self-rating proficiency and social-economic background questionnaire, and objective proficiency in the second language was measured with LexTALE (see Table 1). This is a quick and valid tool to check

<sup>1</sup>Twenty-five participants were allowed to join a second time in the experiment. We did not include the data of their second participation.

second language proficiency for advanced English second language learners (Lemhöfer & Broersma, 2012). Participants saw 60 items within the task: 40 words and 20 non-words from the original version. Participants had a maximum of 3000 ms (instead of unlimited time in the original version) whether the item was a word or not by pressing the 'j' or 'f' key.

## 2.2. Procedure

Each participant signed a consent form with a description of the experiment and we informed the participants they were free to leave at any time without providing any explanation to the experimenter. All three experiments consisted of a pre-selection and the main language-switching study. The main study consisted of two tasks: the classical language-switching paradigm and the modified conversational language-switching paradigm. Both switching paradigms used



the same target pictures and languages in all experiments. The conversational paradigm also used the same question cues in all three experiments.

The two differences between the three experiments lay in the type of arbitrary cue used in the classical paradigm and where the study was conducted (online versus lab). The arbitrary Color cue was used in Experiment 1 and the arbitrary Tone cue in Experiments 2 and 3. Experiments 1 and 2 were conducted online and Experiment 3 was conducted in the lab (see Figure 1), to make sure the results obtained in the online study were replicable in the lab. Both parts of the online experiments were designed in Gorilla Experiment Builder (<https://gorilla.sc/>), a platform for creating and hosting online experiments. Participants were recruited via Prolific (<https://prolific.co>) (Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2019), an online participant recruitment platform. For the lab experiment, participants were selected through a short questionnaire in Office Forms and the main experiment took place in the lab.

### 2.2.1. Pre-selection

The pre-selection aimed at selecting the eligible participants to join the language-switching experiment. During pre-selection, the objective proficiency of English was assessed with the Lex-TALE (Lemhöfer & Broersma, 2012) for the online experiments (score > 50%) and with the General English Test (Cambridge assessment: <https://www.cambridgeenglish.org/test-your-english/general-english/>; minimum score of 18/25) for the lab experiment. For the pre-selection in the online studies, the quality of the audio recordings during a short picture naming task was also checked to make sure that the participants' setup was sufficient to run the main study. Participants who satisfied the requirements were invited to the main experiment within 24 hours after completing the pre-selection task.

### 2.2.2. Language switching experiment

The experimental task consisted of a classical and modified conversational language-switching paradigm (see the Design section). All pictures used in the switching paradigms and their Polish and English names were presented to the participants preceding the paradigms. Lastly, the participants completed a self-rating language proficiency and social-economic background questionnaire. In the online experiments, participants also checked their audio setup before the language switching paradigms and received eight attention checks (e.g., "Now write down this number: 9.") throughout the entire experiment to make sure they paid attention to the task.

### 2.3. Design of the language switching paradigms

For all three experiments, the main task consisted of two cued paradigms: the classical language-switching paradigm and the modified conversational language-switching paradigm. In a language-switching paradigm, participants named pictures based on cues, indicating the language in which to name the pictures. In switch trials, the current trial was to be named in a different language than in the previous trial, whereas in the repeat trials, two successive pictures were supposed to be named in the same language.

All participants performed the classical and conversational switching paradigms, each comprising 229 trials. Each cue paradigm started with a practice session consisting of 17 pictures (i.e., 1 filler and 16 experimental pictures) presented twice in random order and had to be named once in each language. After the practice phase, participants completed three experimental blocks of 65 trials.

After each block, participants could take a short break. Each block started with a filler trial, as it was impossible to determine whether they were repetition or switching trials. The order of the pictures was randomized within each block, while the order of presenting a Polish or English trial was kept constant, but not predictable, within each block to achieve an equal amount of repeat and switch trials. Each of the 16 pictures is repeated four times within a block, equally divided between switch and repeat trials, as well as between the languages (i.e., Polish and English) to name the pictures in. The order of the blocks was counterbalanced between participants. This design resulted in 48 trials per specific condition (e.g., Question cue – repeat trial in Polish) for each participant.

The trial sequence for all 458 trials each participant saw included (1) fixation cross (jittered at 300, 350, 400, or 450 ms; 375 ms during practice and first trials), (2) cue (317 ms) and pictures presented together (2300 ms) and (3) inter-trial interval fixation (300 ms). The sound was recorded from the onset of the cue and picture presentation until the end of the inter-trial interval. All visual stimuli were presented in the center of the screen and the auditory stimuli were pre-recorded and presented through speakers or headphones.

In all three experiments, the conversational switching paradigm used spoken question cues (i.e., 'Co?' versus 'What?'), which always indicated the language to name the picture in. The classical switching paradigm with arbitrary cues differed between experiments. In Experiment 1, blue and red frames presented around the target picture indicated whether to name the target picture in Polish or English. In Experiments 2 and 3, low and high pure tones were presented auditorily and indicated naming in Polish or English (see Figure 2). The arbitrary tone cues matched the domain (auditory) of the transparent cues (question cues) in the conversational switching paradigm. The assignment of arbitrary cue rules to language was counterbalanced across participants. The order of the classical and conversational paradigms was counterbalanced across participants. To conclude, the factors included in the design of each experiment consisted of Cue type (color/tone versus question), Language (Polish versus English) and Trial type (repeat versus switch). For the two sound experiments, the factor Experiment (online versus lab) is also included.

### 2.4. Materials

In all three experiments, the same 16 target pictures (colored line drawings) with non-cognate names between Polish and English were presented (see the Cross-Linguistic Lexical Task picture set [Haman et al., 2015, 2017]; see Appendix B). Mean word (Polish: 5.00 and English: 5.31 letters) and syllable length (Polish: 1.69 and English: 1.38 syllables) were similar in the two languages (respectively,  $p = 0.61$  and  $p = 0.12$ ). An additional picture was presented at the beginning of each block, as the first trial was not analyzed because it cannot be coded as a switch or repeat trial.

Cues indicated the language to name the target picture in. Cues were either arbitrary or transparent in terms of their relation to a specific language. Arbitrary cues were either in the visual domain (red versus blue outline that appeared around the target picture) or in the auditory domain (low [500 Hz] versus high [2000 Hz] tone). These color cues have been used in classical language-switching paradigms (Christoffels et al., 2007; Timmer, Christoffels et al., 2018) and these auditory tone stimuli have been used during task switching (Periáñez & Barceló, 2009). Transparent question cues "Co?" and "What?" were used in all three experiments and recorded by a native speaker of both Poland and English, and adjusted in Praat to have the same length (317 ms).

## 2.5. Data analysis

For all three experiments, response latencies were measured from the onset of the cue and target picture presentation with Praat (<https://www.fon.hum.uva.nl/praat/>). Speech onset was measured with an automatized script defining the loudness and duration of sound, and next checked and adjusted manually when necessary. First, trials with erroneous, incorrect language or no responses (color experiment: 3.64% of the data; sound experiments: 2.17% of the data) were removed from the analyses. Due to the high accuracy rate, we do not report the accuracy analyses here; however, the averages are in line with the naming latencies data. Next, hesitations (e.g., 'uuhhh') were also excluded from further analyses (color experiment: 1.10%; sound experiments: 0.55% of the data without errors). Note that trials after some error types (i.e., no responses and cross-linguistic errors) do not represent a proper repeat or switch condition anymore. Therefore, trials after these error types are not included in the analyses (color experiment: 2.98%; sound experiments: 1.52%). Subsequently, latencies shorter than 317 ms (cue duration) and longer than 2500 ms were discarded from the analyses (color experiment: 0.22%; sound experiments: 0.79% of the data without errors and hesitations). The total number of trials left in the analyses is 92.28% in the color experiment and 95.10% in the tone experiment.

Naming latencies were analyzed in R (version: 4.0.2; R Core Team, 2020) using linear mixed-effect models as implemented in the lme4 package (Bates et al., 2015). We fitted a separate model for Experiment 1 (comparing color versus question cues) and a joint model for Experiments 2 and 3 (comparing tone versus question cues). The predictors included in the models were: Cue (color/tone versus question), Language (L1/L2) and Trial type (repeat/switch). The joint model based on data from Experiments 2 and 3 also included the predictor Experiment (lab versus online). Additionally, both models included two-way interactions between Language and Trial type, Cue and Trial type, Cue and Language, and a three-way interaction between Cue, Language and Trial type. Lastly, for the model comparing tone versus question cues, interactions with the Experiment were also included.

Before the analysis, the dependent variable, *Naming latencies*, was log-transformed to correct for the model's assumption of the normal distribution of the residuals and all categorical predictors were deviation-coded (Cue: question = -0.5, color/tone = 0.5; Language: L1 = -0.5, L2 = 0.5; Trial type: repeat = -0.5, switch = 0.5; Experiment: lab = -0.5, online = 0.5). The final model for Experiment 1 (color versus question cue) included random intercepts by subjects and items, as well as random slopes for Cue, Language and Trial type by subjects, and Cue and Language by items (uncorrelated with the intercept). The final model for Experiments 2 and 3 (tone versus question cue) included random intercepts by subjects and items, and random slopes for Cue, Language and Trial type by subjects and Language by items (uncorrelated with the intercept). To further explore the results of interactions, we run a series of posthoc pairwise comparisons using the lsmeans() function from the lsmeans package (Lenth & Lenth, 2018).

To assess how L2 proficiency affects the reverse language dominance effect and the asymmetry of switch costs, we computed the indices of reversed language dominance and asymmetry of switch costs for each participant averaged over Cue type in the three experiments. The reversed language dominance index was defined as a difference between mean naming latencies in L1 and L2 (L1 - L2). Therefore, the larger the value of the index, the greater the relative difference between languages (L1 > L2). The index of asymmetry of switch costs was calculated as a difference between

mean switch costs (mean of RTs in switch - repeat trials) between L1 and L2 (switch costs in L1 - switch costs in L2). Therefore, the larger the value of the index, the bigger the L1 switch cost relative to the L2 switch cost. These two indices were subsequently used as dependent variables in two regression models, which assessed the relationship between proficiency and reverse dominance (Model 1) and asymmetry of switch costs (Model 2). We chose LexTALE, as it is an objective measure of proficiency for second language learners. In both cases, we fitted a linear regression using the following formula:

Reversed language dominance/Asymmetry of switch costs index ~ LexTALEscore. Before the analysis, the continuous predictor of LexTALE was demeaned.

## 3. Results

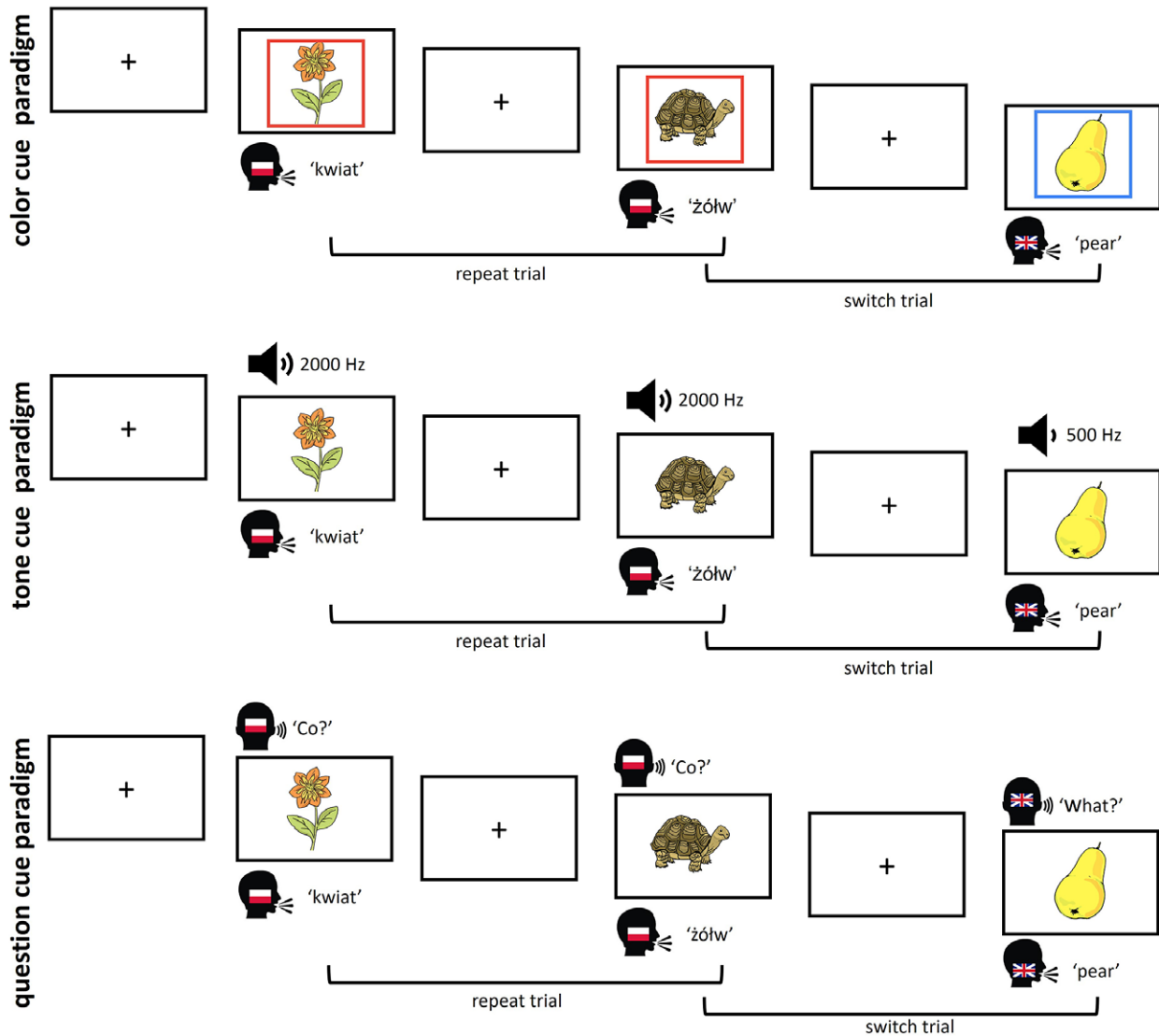
The mean naming latencies for all conditions by Cue, Language and Trial type are presented in Table 2 for each of the three experiments. The indices (A) switch costs (switch - repeat trials) and (B) reversed language dominance (L1 - L2 trials) for these experiments are presented in Figure 3.

### 3.1. Experiment 1: color versus question cue

The results of Experiment 1 revealed a main effect of Cue, showing that participants named pictures faster when presented with a color cue than with a question cue, a main effect of Language, showing that participants named pictures in L2 faster than in L1

**Table 2.** Mean naming latencies (ms) predicted by the models. Values in brackets correspond to 95% confidence intervals of the models

Experiment 1 (online): color versus question cue				
Trial type	color cue		question cue	
	L1	L2	L1	L2
repeat	969 [935, 1003]	920 [883, 958]	1082 [1041, 1125]	1022 [982, 1064]
switch	1044 [1007, 1082]	999 [958, 1041]	1145 [1100, 1191]	1090 [1047, 1136]
Experiment 2 (online): tone versus question cue				
Trial type	tone cue		question cue	
	L1	L2	L1	L2
repeat	1057 [1016, 1099]	1009 [968, 1052]	1046 [1007, 1086]	985 [946, 1026]
switch	1160 [1113, 1209]	1123 [1076, 1173]	1103 [1060, 1147]	1052 [1009, 1097]
Experiment 3 (lab): tone versus question cue				
Trial type	tone cue		question cue	
	L1	L2	L1	L2
repeat	935 [902, 970]	886 [853, 921]	922 [890, 954]	858 [827, 891]
switch	1023 [985, 1062]	984 [946, 1024]	972 [938, 1008]	917 [882, 953]



**Figure 2.** An example of a trial within each of the three Cue paradigms. In the color task, the classical language switching paradigm was used with blue and red colored frames around the target picture to indicate the language to name the picture in. In the tone task, high (2000 Hz) and low (500 Hz) tones indicated which language to use. In the question task, the language the question was asked in ('Co?' or 'What?') was the language to respond in. The target pictures are examples of the stimuli used in the experiment and come from the Cross-Linguistic Lexical Tasks (CLT) database (Haman et al., 2017).

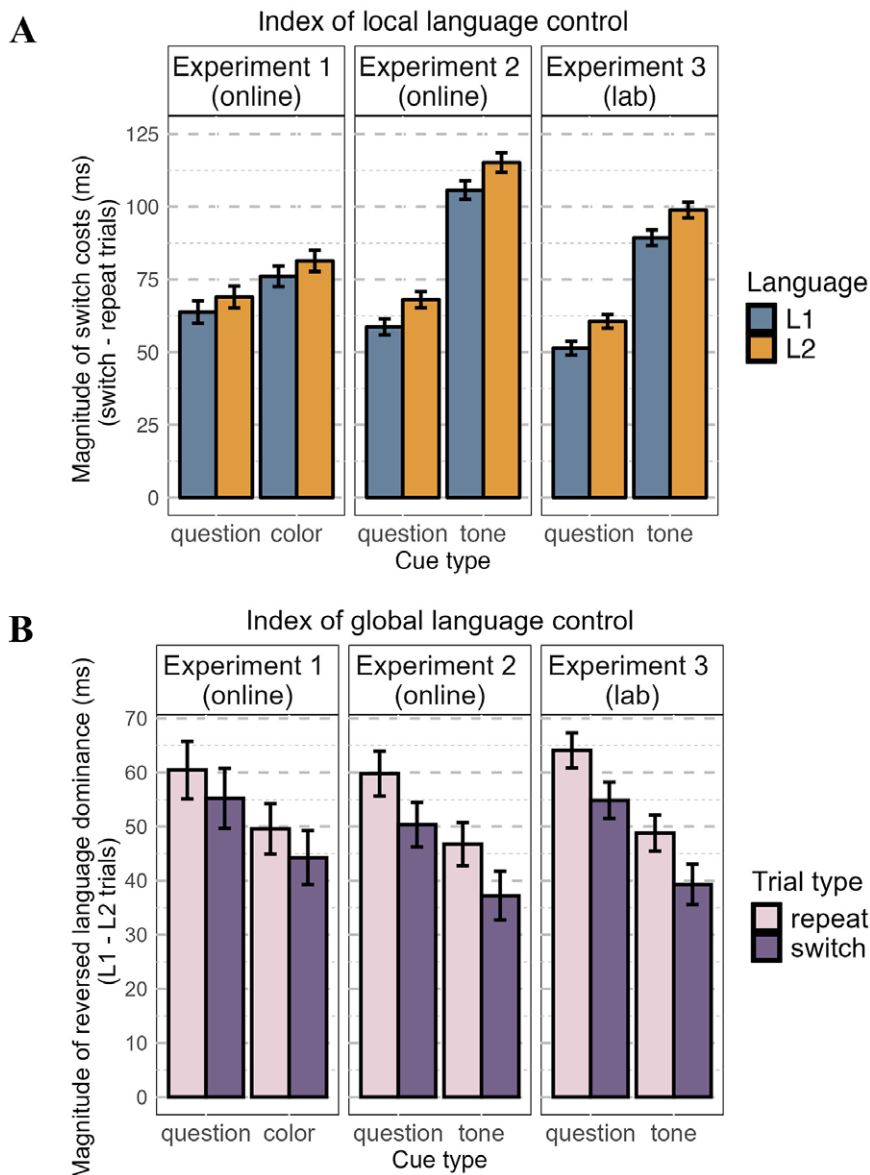
(i.e., reversed language dominance), and a main effect of Trial type, showing that participants were faster on the repeat than on switch trials (i.e., switch cost). We also found a significant interaction between the Cue and the Trial type. To further explore the interaction, we performed pairwise comparisons. Crucially, this revealed that the switch cost (naming latencies on switch – repeat trials) was smaller for the question cue (65 ms;  $z = -11.738$ ,  $p < 0.001$ ) than for the color cue (77 ms,  $z = -15.241$ ,  $p < 0.001$ ). The results of the main model are summarized in Table 3.

To conclude, Experiment 1 replicated the previous finding in the literature, showing a typical switch-cost effect and a reversed dominance effect. Importantly, the conversational paradigm with question cues revealed a smaller switch cost than the classical switching paradigm.

### 3.2. Experiments 2 and 3: tone versus question cue

The results of Experiments 2 and 3 revealed the following: (1) a main effect of Experiment, showing that participants were faster in

the lab study than in the online study; (2) a main effect of Cue, showing that participants named pictures faster when presented with a question cue than with a tone cue; (3) a main effect of Language, showing that participants named pictures in L2 faster than in L1; (4) and a main effect of Trial Type, showing that participants were faster on the repeat than on switch trials. We also found significant interactions between Cue and Trial type, between Cue and Language and between Language and Trial type. To further explore the interactions, we performed pairwise comparisons. Similar to experiment 1, the switch cost was smaller for the question cue (58 ms;  $z = -16.192$ ,  $p < 0.001$ ) than for the arbitrary (tone) cue paradigm (100 ms,  $z = -26.661$ ,  $p < 0.001$ ). The switch cost was larger for the L2 (84ms;  $z = -23.315$ ,  $p < 0.001$ ) than for the L1 (74 ms;  $z = -19.610$ ,  $p < 0.001$ ). In addition, surprisingly, we found that the effect of Language (i.e., reversed language dominance) was larger for the question cue (58 ms;  $z = 3.944$ ,  $p < 0.001$ ) than for the tone cue (43 ms;  $z = 2.846$ ,  $p < 0.001$ ). The results of the main model are summarized in Table 4.



**Figure 3.** Indices of (A) the local control (measured by the switch cost) and (B) the global control (measured by the reverse language dominance) in experiments 1, 2, and 3. Indices correspond to differences in mean naming latencies between (A) switch – repeat trials and (B) L1 – L2 trials. Error bars correspond to the standard error of the mean.

To conclude, the results of Experiments 2 and 3 replicate those of experiment 1 for local control: a typical switch cost (repeat < switch), which was smaller for question cues than arbitrary cues and also larger for L2 than for L1. The latter asymmetrical switch cost was driven by stronger facilitation for repeating L2 than L1 (55 ms,  $z = 3.848$ ,  $p < 0.001$ ) than switching to L2 than L1 (46 ms,  $z = 2.943$ ,  $p < 0.001$ ). In contrast to Experiment 1, the reversed dominance effect was greater for the question cues than the arbitrary tone cues in Experiments 2 and 3. Based on available evidence, this effect was unexpected, as we predicted a similar magnitude of reversed language dominance in the two cue paradigms. However, none of the previous studies compared question cues versus tone cues. We found that this effect is driven by the fact that question cues are overall faster than tone cues, with the facilitation of naming in L2 being greater (47 ms,  $z = -7.481$ ,  $p < 0.001$ ) than in L1 (32 ms,  $z = -5.154$ ,  $p < 0.001$ ). This translates to a greater difference

between L1 and L2 naming speed in response to the question cue than arbitrary tone cues.

### 3.3. The impact of L2 proficiency on reversed language dominance and asymmetric switch costs – a comparison across all experiments

The results revealed a significant effect of L2 proficiency on reversed language dominance ( $\beta = -3.278$ ,  $p = 0.001$ ), showing a smaller reversed language dominance effect in bilinguals with higher L2 proficiency (see Figure 5A). However, no significant effect of L2 proficiency was found on the asymmetry of switch costs ( $\beta = 1.398$ ,  $p = 0.164$ , see Figure 5B). It is noteworthy that the majority (84.0%) of participants show the reversed language dominance effect, while the direction of the asymmetric switch cost was more equally divided between participants (43.3% L1 > L2 and 56.7% L1 < L2).



**Table 3.** Results of the naming latencies analysis in experiment 1 (color versus question)

Predictors	Estimate	SE	95% CI	t-value	p-value
Intercept	0.03	0.02	0.00–0.06	1.84	0.066
Cue (question versus color)	0.10	0.01	0.08–0.12	8.45	<b>&lt;0.001</b>
Language (L1 versus L2)	–0.05	0.02	–0.08 – –0.02	–3.08	<b>0.002</b>
Trial type (repeat versus switch)	0.07	0.00	0.06–0.07	15.48	<b>&lt;0.001</b>
Cue × Trial type	–0.02	0.01	–0.03 – –0.01	–3.68	<b>&lt;0.001</b>
Cue × Language	–0.00	0.01	–0.01 – 0.01	–0.87	0.382
Language × Trial type	0.01	0.01	0.00–0.02	1.57	0.116
Cue × Language × Trial type	–0.00	0.01	–0.02 – 0.02	0.11	0.912

**Table 4.** Results of the naming latencies analysis in experiments 2 and 3 (tone versus question)

Predictors	Estimate	SE	CI	t-value	p-value
Intercept	0.00	0.01	–0.03 – 0.02	–0.12	0.907
Experiment (online versus lab)	0.13	0.02	0.08–0.18	5.45	<b>&lt;0.001</b>
Cue (question versus tone)	0.04	0.01	0.03–0.05	6.78	<b>&lt;0.001</b>
Language (L1 vs L2)	–0.05	0.01	–0.08 – –0.02	–3.42	<b>0.001</b>
Trial type (repeat versus switch)	0.08	0.00	0.07–0.09	24.39	<b>&lt;0.001</b>
Cue × Trial type	0.04	0.00	0.03–0.05	11.04	<b>&lt;0.001</b>
Cue × Language	0.02	0.00	0.01–0.02	4.67	<b>&lt;0.001</b>
Language × Trial type	0.02	0.00	0.01–0.02	3.86	<b>&lt;0.001</b>
Cue × Language × Trial type	0.00	0.01	–0.03 – 0.02	–0.38	0.705
Experiment × Cue	0.01	0.01	–0.01 – 0.03	1.06	0.290
Experiment × Language	0.00	0.01	–0.01 – 0.01	0.23	0.819
Experiment × Trial type	0.00	0.01	–0.01 – 0.02	0.19	0.851
Experiment × Cue × Trial type	0.00	0.01	–0.01 – 0.02	0.48	0.630
Experiment × Cue × Language	0.00	0.01	–0.02 – 0.01	–0.67	0.501

#### 4. Discussion

We explored whether spoken question cues reduce the presence of BLC by comparing the performance in two language-switching paradigms: classical (with arbitrary cues) versus conversational (with question cues). The classical language switching paradigm, introduced by Meuter and Allport (1999) and typically used in the literature (see Appendices A1 and A2 for a list of the cue and target types used in the literature), employs colored arbitrary cues to indicate the language to name the pictures in (e.g. Christoffels et al., 2007). This paradigm usually reveals that switching between languages comes at a cost (local control) and makes naming pictures slower in the L1 than in the L2 (global control) (Bobb & Wodniecka, 2013). In Experiment 1, we compared the findings from the classical paradigm with those obtained in a novel conversational language switching paradigm, which used auditorily

presented questions (i.e., ‘Co?’ versus ‘What?’) instead of the typical arbitrary cues (i.e., colors). As predicted, the conversational paradigm reduced the switch cost compared to the classical paradigm, suggesting a diminished demand for local control. However, we also found that pictures were overall named faster in the classical than conversational paradigm, most likely because the visual processing of colors is faster than the auditory processing of questions (Lukas et al., 2010). Therefore, to keep the modality of cues constant, in the following experiments (Experiments 2 and 3), we compared the auditorily presented question cues to arbitrary pure tone cues, often applied in the task-switching literature (Periáñez & Barceló, 2009). As expected, pictures were named slower when arbitrary tones rather than conversational questions were presented. Crucially, both experiments also revealed a reduced switch cost during the conversational paradigm. However, surprisingly, the magnitude of global control (i.e., L1 > L2) increased in

Experiments 2 and 3. As such, the conversational switching paradigm reveals reduced switch costs with conversational cues and a potentially intriguing dissociation between the two indices of language control.

We first discuss why auditory question cues facilitate language choice compared to auditory but arbitrary cues, next we discuss how questions modulate local and global BLC mechanisms compared to arbitrary cues and, lastly, we reflect on how the findings impact theoretical models and research in bilingualism.

#### 4.1. Language selection

In Experiment 1, pictures were named slower when auditory question cues (e.g., ‘Co?’ versus ‘What?’) compared to visual color cues (e.g., blue and red) indicated the language to name the target pictures in. This is in line with a previous study showing slower responses when the cue and target stimulus were presented in a different modality than the same one (Lukas et al., 2010). When controlling for (mis)matching cue-target modality, auditory cues were overall slower than visual cues. Therefore, we suggest that either cue modality (auditory cues) or the modality mismatch with the visual target stimulus could have slowed down the processing speed in the auditory paradigm. In Experiments 2 and 3, we compared two types of auditory cues (question cues and tone cues; see Figure 1) to dissociate the effects of the modality itself or cue-target modality mismatch. As predicted, when cue modality was kept the same, question cues facilitated language choice over pure tone cues (see Figure 4). A likely reason for this facilitation is that hearing a question in a specific language is a more transparent cue to name objects in that language than hearing a pure tone (Arrington et al., 2007; Arrington & Logan, 2004; Grosjean, 2001).

#### 4.2. Modulations of BLC mechanisms

For the local control index, namely the switch costs (Timmer, Calabria et al., 2018; Timmer, Christoffels et al., 2018), we observed that auditory question cues (in the conversational switching paradigm) reduced the switch cost in all experiments compared to the arbitrary cues used in the classical switching paradigms. This implies that the need for control is smaller when an auditory question cue triggers language selection compared to a visual cue. The consistent reduction is most likely related to the consistent reduction of local control in the conversational language switching paradigm across the three reported comes from the removal of artificially induced laboratory cues (e.g., colors). The questions in different languages directly boost the language because, in daily life, answering questions in the appropriate language is a normal part of conversations. Furthermore, the preexisting association between questions and answers also increased cue-target transparency (Logan & Schneider, 2006). Notably, the two previous studies that compared faces (Blanco-Elorrieta & Pykkänen, 2017; Liu et al., 2019) to arbitrary cues showed much variability in their findings. For example, for Chinese – English bilinguals, only Asian-looking faces reduced the switch cost in their native language, but Caucasian-looking faces did not impact switch costs (Liu et al., 2019), while faces did not reveal a switch cost at all for Arabic – English speaker (Blanco-Elorrieta & Pykkänen, 2017). With question cues, we revealed a reduction (but not full absence) of switch costs in three experiments with large sample sizes for question cues.

Although the conversational switching paradigm reduced the overall language switch cost, the direction of switch cost (symmetric [similar in L1 and L2] in Experiment 1 and asymmetric [larger for L2 than L1] in Experiments 2 and 3) remained the same regardless of the cue manipulation (i.e., type of switching paradigm) and experiment. A recent meta-analysis has suggested that the findings

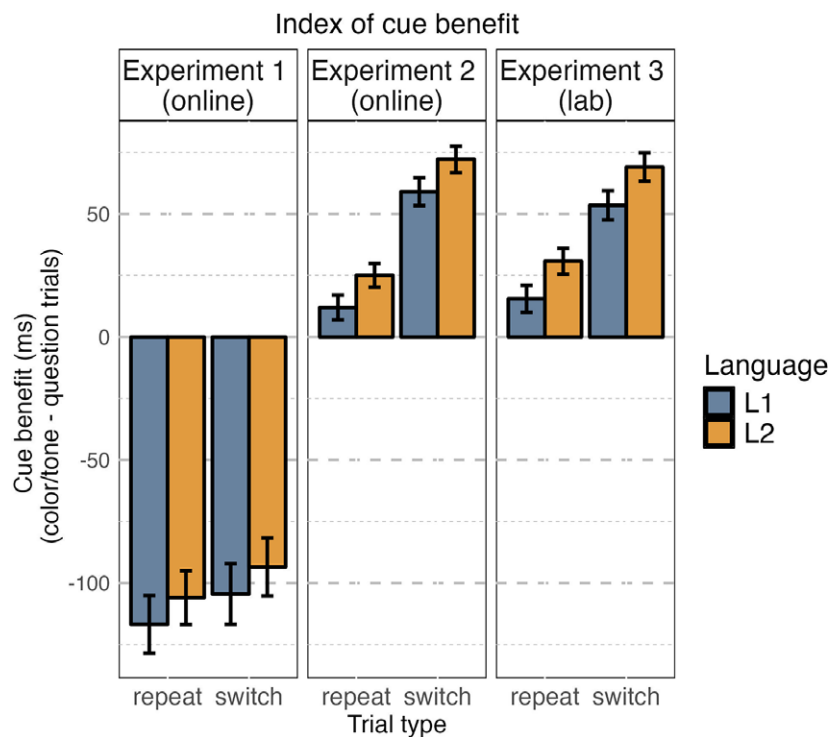
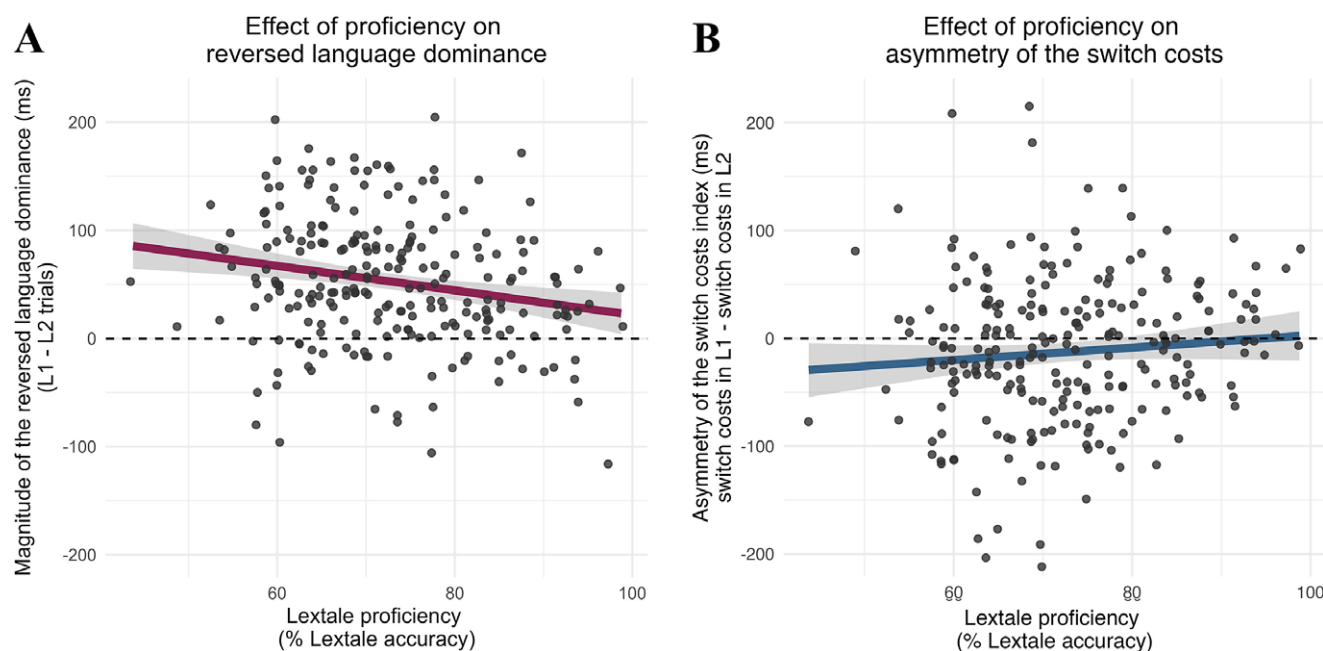


Figure 4. Index of cue benefit (arbitrary (color/tone) – question trials) in experiments 1, 2, and 3. Error bars correspond to the standard error of the mean.



**Figure 5.** The relationship between L2 proficiency (measured by LexTALE) and the reversed language dominance (A) and the asymmetry of switch costs (B). Participants above the dashed line showed a reversed language dominance (A) and asymmetric switch cost greater for L1 than L2.

regarding asymmetry related to bilinguals' language proficiency are unreliable (Gade et al., 2021a, 2021b). However, this meta-analysis also includes studies that predicted and found a reversed asymmetry (Liu et al., 2019; Timmer, Christoffels et al., 2018). For example, asymmetric switch costs when language switching occurred in a context mainly using the non-dominant second language (Timmer, Christoffels et al., 2018) or when cultural faces (i.e., Asian and Caucasian) indicated the language to speak in (Liu et al., 2019), but symmetric switch costs when color cues were used. Therefore, it is currently unclear how reliable the asymmetry of switch costs is. The theoretical interpretation of the asymmetric switch cost is also debated. The most common interpretation of the asymmetric switch cost is that the larger switch cost to L1 is due to overcoming greater inhibition of the stronger language when reactivating the L1. However, our follow-up analyses for the asymmetric switch cost in the tone versus question cue experiments showed that the larger switch cost to L2 originated from stronger L2 facilitation for repeating the L2 than switching to the L2 (as compared to repeating and switching to L1). Thus, we suggest that the language switch cost asymmetry ( $L2 > L1$ ) is due to increased facilitation of L2 repetition.

For the global control index (reversed language dominance), in Experiment 1, we observed no modulation of the reversed language dominance (i.e., L1 naming is slower than L2) due to the paradigm type. This was in line with earlier findings in the literature: (1) the robustness of the effect and (2) the absence of a modulation due to cue type. First, a recent meta-analysis by Goldrick and Gollan (2023) showed that 81.25% of observations revealed a reversed language dominance ( $L1 > L2$ ) in mixed language contexts. It is sometimes suggested that the reversed language dominance should be compared to single language blocks, which is not what we did. However, the meta-analysis showed that when the reversed language dominance in mixed blocks (81.25%) was compared to language dominance in single blocks (75% showed a dominant language advantage [ $L1 < L2$ ]), the reduction of this dominance in the mixed language context was present in 88% of observations

(Goldrick & Gollan, 2023). Thus, both the reversed language dominance and the reduction concerning single language blocks are often replicated, and so Goldrick and Gollan consider both effects as reflecting the same underlying cognitive mechanism. Second, cultural objects and culture-specific faces did not modulate global control compared to arbitrary color cues (Liu et al., 2019, 2021), because the effect is driven by the relative activation of two languages (Costa et al., 2006; Costa & Santesteban, 2004; Goldrick & Gollan, 2023), which are equally present in both cue paradigms.

However, contrary to our prediction and finding in Experiment 1 (i.e., global control was not modulated by cues), global control was modulated between the question cues and tone cues (Experiments 2 and 3). Specifically, the magnitude of the reversed language dominance was greater in response to the conversational question cue than the tone cue. Interestingly, our follow-up analyses revealed that hearing the question 'What?' facilitated naming in L2 more than the question 'Co?' facilitated naming in L1 compared to tone cues. For the Polish-English bilinguals tested, the L2 is weaker than the L1 and less present in their daily environment. Therefore, it makes sense that hearing an English (L2) question is more beneficial than hearing a Polish (L1) question, as the latter is their standard. This could mean the greater magnitude of the reversed language dominance for questions than tone cues is driven by greater facilitation for L2 than L1. This modulation was not observed in Experiment 1, possibly because naming in the conversational paradigm is overall slower than in the classical color cue paradigm. Thus, future research must establish the robustness of modulations of the reversed language dominance effect due to the transparency of cues, as it was not fully consistent across experiments.

In an additional analysis of all experiments combined, we checked whether L2 proficiency modulates the asymmetry of the switch cost and the magnitude of the reversed language dominance effect. We found that L2 proficiency did not modulate the asymmetry of switch cost (see Figure 5B). The absence of asymmetry modulations due to L2 proficiency replicates that of a meta-analysis

(Gade et al., 2021b). However, L2 proficiency did modulate the magnitude of the reversed language dominance (see Figure 5A). When participants have a higher objective L2 proficiency, as measured by the LexTALE score (Lemhöfer & Broersma, 2012), the magnitude of the reversed language dominance is reduced (i.e., a smaller difference between L1 – L2). This replicates the finding of Goldrick and Gollan (2023) – who found a smaller reversed language dominance when bilinguals were more balanced – with a smaller sample, different measure and a different task. Goldrick and Gollan (2023) showed bilinguals made more intrusion errors in their L1 than L2 during mixed language paragraph reading. The difference in intrusion errors between languages was larger for more unbalanced bilinguals. We replicate a similar pattern during speech production (i.e., picture naming) using the more sensitive measure of the naming latencies and a larger sample size. At the same time, our pattern contradicts the one found in the study by Declerck et al. (2020), who found that more balanced bilinguals reversed language dominance more than less balanced bilinguals. The different participant populations used in the studies may be a possible reason for this discrepancy. We tested a group of unbalanced, L1-dominant bilinguals whose L1 is also their first acquired native language. On the other hand, Declerck et al. (2020) tested a group of highly proficient Spanish-English bilinguals living in San Diego. For most (91.6%) participants, L2 (English) became their dominant language and they lived immersed in the L2 environment. The fact that bilinguals were more balanced in their study is confirmed by an absence of reversed language dominance (i.e., no reversed language dominance), while we showed that 84% of our participants showed a reversed language dominance pattern. Thus, we conclude that less balanced bilinguals set the relative activation of their L1 and L2 differently than more balanced bilinguals.

#### 4.3. Theoretical and methodological implications

Our findings suggest that existing theoretical models of BLC (Blanco-Elorrieta & Caramazza, 2021; Green, 1998; Green & Abutalebi, 2013) should integrate language comprehension and production to explain how bilinguals perform a speech act in a conversational context. Notably, such models are already proposed for monolinguals (Levinson & Torreira, 2015; Pickering & Garrod, 2013), but must be extended to bilinguals. For example, the bilingual IC model states that language switching during production is initiated by a goal (i.e., to speak in a specific language) and that the Supervisory Attentional System (SAS) monitors the language goal and transmits it to the language schemes (e.g., Polish and English), which in turn activate the correct lemma from the bilingual lexico-semantic system. We propose that a revised model should clarify the stage at which speech is initiated in a conversational setting. Our data suggest that the goal is triggered more efficiently when the cue is directly linked to the goal, i.e., the cue is formulated as a question and the goal is to answer the question by naming a picture in the correct language. This cue – goal relationship is specified in the turn-taking model of Levinson and Torreira (2015). According to their model, the task is initiated endogenously (i.e., top-down communicative goal) and quickly when the input – goal relationship occurs as in real-life conversations.

Furthermore, the auditory question cue is inherently in a specific language. Therefore, we propose auditory questions not only activate the goal of speaking (endogenous) but (also) activate words in the corresponding language exogenously (i.e., bottom-up language activation), thus likely boosting the language as a whole, not only a single word in the mental dictionary of the Inhibitory

Control (IC) model (Green, 1998). Interestingly, we found that language selection was boosted more by L2 than L1 in the conversational compared to the classical paradigm. This suggests the nature of reversed language dominance might not necessarily reflect inhibition applied to L1, but instead, it may reveal increased L2 activation, which fluctuates depending on specific task demands and opportunities. To conclude, conversational cues may boost language production through both exogenous (bottom-up language activation) and endogenous (top-down communicative goal to produce a word in a specific language) input. This bears similarity with the Bilingual Interactive Activation model (Grainger et al., 2010).

Another model of bilingual speech production, the Adaptive Control Hypothesis (ACH), focuses on the changes in control demand depending on the interactional context a bilingual speaker is in. In situations with both languages present but spoken by different interlocutors (i.e., a dual-language context), the detection of salient cues (e.g., the arrival of a new interlocutor speaking a different language) is crucial. Detecting a cue indicating a language switch may trigger other control processes, such as selective response inhibition and task (dis)engagement, to start speaking another language. These later control mechanisms suppress the activation of the non-target language and sustain attention on the target language. According to ACH, this creates a control dilemma as the suppressive state may also reduce sensitivity to relevant external cues (i.e., salient cue detection).

The pattern of results in the present study shows that the salient cue detection process (i.e., defined in ACH) is easier when cues are more transparent rather than arbitrary. If cues are transparent to the goal (i.e., hearing a language and answering in that language), the control processes following cue detection that are necessary to switch languages (e.g., selective response inhibition) seem to be less demanding as well. This is in line with the non-linguistic task-switching literature, which showed that with greater cue transparency, less working memory is required to make a switch between tasks (Arrington et al., 2007; Arrington & Logan, 2004; Grange & Houghton, 2010). Furthermore, we propose that the control dilemma mentioned in ACH (i.e., suppressing the non-target language also suppresses the ability to detect cues) may not arise for transparent cues, such as language cues, because bilinguals monitor their language environment for potentially relevant cues that indicate naturally in which language to respond (Timmer, Costa et al., 2021; Timmer, Wodniecka et al., 2021). Therefore, we propose the degree of transparency is an important feature that impacts the degree of engagement in control processes. This distinction in cue detection is important to include in ACH, as it could potentially impact the proposed cascading of control mechanisms in different ways.

Lastly, when speakers receive questions and have to answer, the turn-taking model suggests the speakers are dual-tasking between listening and speaking (Levinson & Torreira, 2015). However, the IC model (Green, 1998) and ACH (Green & Abutalebi, 2013) do not propose any mechanism to manage this situation. We propose that a plausible candidate for this role is the SAS, which monitors performance during (non)linguistic switching tasks in the IC model.

#### 4.4. Future directions

The current results provide an important first step showing that using existing mappings (i.e., auditory questions) to trigger actions (i.e., speaking) in experimental paradigms impacts the indices



(i.e., switch cost and reversed language dominance) we use to draw theoretical conclusions. Our data suggest that when arbitrary cues are used, these indices say little about the difficulty to switch between languages and more about the difficulty to switch between cues. This can impact conclusions drawn in the existing literature. For example, the extent to which the overlap found between language and task switching (Timmer et al., 2017) could be due to arbitrary cue processing instead of the actual language-switching mechanisms between the two domains. We suggest that the conversational paradigm as a comparison with the task-switching paradigm could reveal the actual shared control mechanisms between language and task switching.

At the same time, we acknowledge that the paradigm we use involves only one element of conversational switching. In daily life, bilinguals switch due to many reasons (i.e., not only in response to questions) and there is much more variability in both questions (or other input) and answers. Therefore, using more than one cue for each language will allow to separate the cost of switching between languages and assigning meaning to the cue (i.e., cue switch cost) in future studies. To the best of our knowledge, only two studies in the language-switching literature used two cues (i.e., faces) for each language. They showed that switching face cues while staying in the same language incurred a cost (i.e., cue-switch cost), whereas switching between languages and faces also incurred a cost (Heikoop et al., 2016; Peeters, 2020).

Furthermore, we hope this work provides a new avenue for future research. Although most studies used arbitrary cues, some studies already used more ecological cues (Appendix A). We predict that auditory cues like 'say' used by Hernandez et al. (1999, 2000, 2001) would similarly activate the speech goal as question cues used in the present study and by Tarlowski et al. (2012). To understand whether speech production in response to auditory cues is crucial to capture BLC, the auditory questions could be compared to visual questions as used by Zhu and Soman (2020). Furthermore, to understand whether a communicative cue, like a question or the word 'say', has an additional benefit over non-communicative words could be tested by comparing auditory question cues to non-question words in the respective languages (i.e., the latter used by Peeters et al., 2014). We hypothesize that when only a few words are used, both (non-)communicative words are quickly linked to the speech goal. However, when many different words are presented this might slow down overall processing speed. It could also be interesting to what extent faces (e.g., Peeters, 2020; Peeters & Dijkstra, 2018) and flags (e.g., Prior & Gollan, 2013; Timmer et al., 2019; Weissberger et al., 2015), which also have a preexisting relation to the cued language, would produce similar results. We hope our work in which we chose for auditory question cues in different languages, as they have the clearest link to the task at hand, provides a base for more research into the relation between the goal and action during bilingual switching.

## 5. Conclusions

In three experiments, we revealed a stable reduction of switch costs for questions over arbitrary cues over multiple experiments in contrast to the diverse results from the transparent face cues. Therefore, the present results systematically remove the artificially induced cost due to the opaque cue to language mapping for arbitrary cues. This is crucial for theoretical models as communication is the core goal of language. Specifying how the input node from the IC model activates language and task in theoretical models

helps to understand that a conversational cue: an auditory question, activates the goal of answering the question, the word that constitutes the response, and the language, congruent to that of the question. On the contrary, arbitrary cues only activate the goal through the first process. The need for local language control can be diminished when the input cue activates both the language and the goal. Counterintuitively, global control is increased with these conversational cues. In our results, in comparison to hearing arbitrary cues, hearing a question in L2 helps to speak in a foreign language more than hearing a question in L1. Therefore, the reversed language dominance may actually reflect L2 facilitation instead of L1 inhibition. Thus, the modified conversational switching paradigm reveals an intriguing dissociation between the two indices of language control.

The methodological conclusion of our study is that researchers trying to understand speech production mechanisms should turn to tasks that better reflect a speaker's communicative goal instead of employing tasks where the speech act is initiated by arbitrary cues, which artificially induce costs associated with cue processing. Considering why speech is initiated – or, more broadly, what triggers people to take specific actions in their daily lives – should help better reflect the cognitive mechanisms used to maneuver through our daily activities and introduce relatively more ecological laboratory paradigms. We believe that our new conversational switching paradigm may provide a more accurate measure of local language control. However, construct validity is yet to be measured in future studies by correlating it with another measure of how easy it is to switch between languages. However, up to date, there is no such well-established measure to compare it to, and to the best of our knowledge, no previous studies on language switching have provided such validity tests. To conclude, we propose that future research should consider using a cue that has the most apparent association with the goal of the task the participants will perform.

**Acknowledgments.** Kalinka Timmer was supported from Polish National Agency For Academic Exchange Narodowa Agencja Wymiany Akademickiej) in Poland with an Ulam grant (PPN/U LM/2019/1/00215) and by the National Science Centre (Narodowe Centrum Nauki) with an OPUS grant (2022/45/B/HS6/01931). The research in the lab experiment was funded by the Priority Research Area Heritage under the program Excellence Initiative – Research University at the Jagiellonian University in Krakow awarded to Kalinka Timmer. The research in the online experiments was funded by the National Science Centre with the SONATA BIS grant awarded to Zofia Wodniecka (2015/18/E/HS6/00428) and the Faculty of Psychology at the University of Warsaw. We thank Rita Akl, Zofia Kania, Joanna Kowacz, and Antonina Witkowska for their help with data collection and Rita Akl, Zofia Kania, and Antonina Witkowska, and Patrick Ye for their help with data coding. We would also like to thank Alba Casado and Piotr Rutkowski for their help in creating the auditory stimuli. Lastly, we thank Anne Bolders for the helpful discussion of design differences between the visual and auditory domains.

**Data availability statement.** All data and scripts necessary for reproducing the analyses and plots in this paper are available at <https://osf.io/qswfk>.

**Competing interest.** The author(s) declare none.

## References

- Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388–407. <https://doi.org/10.3758/s13428-019-01237-x>
- Arrington, C. M., & Logan, G. D. (2004). Episodic and semantic components of the compound-stimulus strategy in the explicit task-cuing procedure.

- Memory and Cognition*, 32(6), 965–978. <https://doi.org/10.3758/BF03196874/METRICS>
- Arrington, C. M., Logan, G. D., & Schneider, D. W. (2007). Separating cue encoding from target processing in the explicit task-cuing procedure: are there “true” task switch effects? *Journal of Experimental Psychology: Learning Memory and Cognition*, 33(3), 484–502. <https://doi.org/10.1037/0278-7393.33.3.484>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/JSS.V067.I01>
- Blanco-Elorrieta, E., & Caramazza, A. (2021). A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition*, 213, 104625. <https://doi.org/10.1016/j.COGNITION.2021.104625>
- Blanco-Elorrieta, E., & Pyllkkänen, L. (2017). Bilingual language switching in the laboratory versus in the wild: the spatiotemporal dynamics of adaptive language control. *The Journal of Neuroscience*, 37(37), 9022–9036. <https://doi.org/10.1523/JNEUROSCI.0553-17.2017>
- Blanco-Elorrieta, E., Emmorey, K., & Pyllkkänen, L. (2018). Language switching decomposed through MEG and evidence from bimodal bilinguals. *Proceedings of the National Academy of Sciences*, 115(39), 9708–9713. <https://doi.org/10.1073/pnas.1809779115>
- Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning, 25(5), 568–585. <https://doi.org/10.1080/20445911.2013.792822>
- Campbell, J. I. D. (2005). Asymmetrical language switching costs in Chinese–English bilinguals’ number naming and simple arithmetic. *Bilingualism: Language and Cognition*, 8(1), 85–91. <https://doi.org/10.1017/s136672890400207x>
- Castellucci, G. A., Guenther, F. H., & Long, M. A. (2022). A theoretical framework for human and nonhuman vocal interaction. *Annual Review of Neuroscience*, 45, 295–316. <https://doi.org/10.1146/ANNUREV-NEURO-111020-094807>
- Castellucci, G. A., Kovach, C. K., Howard, M. A., Greenlee, J. D. W., & Long, M. A. (2022). A speech planning network for interactive language use. *Nature*, 602(7895), 117–122. <https://doi.org/10.1038/s41586-021-04270-z>
- Chang, S., Xie, J., Li, L., Wang, R., & Liu, M. (2016). Switch Costs Occur at Lemma Stage When Bilinguals Name Digits: Evidence from Language-Switching and Event-Related Potentials. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01249>
- Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, 1147(1), 192–208. <https://doi.org/10.1016/j.BRAINRES.2007.01.137>
- Contreras-Saavedra, C. E., Koch, I., Schuch, S., & Philipp, A. M. (2020). The reliability of language-switch costs in bilingual one- and two-digit number naming. *International Journal of Bilingualism*, 25(1), 272–285. <https://doi.org/10.1177/1367006920951873>
- Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition*, 106(1), 59–86. <https://doi.org/10.1016/j.COGNITION.2006.12.013>
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50(4), 491–511. <https://doi.org/10.1016/j.JML.2004.02.002>
- Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning Memory and Cognition*, 32(5), 1057–1074. <https://doi.org/10.1037/0278-7393.32.5.1057>
- de Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language*, 103, 28–43. <https://doi.org/10.1016/j.jml.2018.07.005>
- de Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2020). Examining bilingual language switching across the lifespan in cued and voluntary switching contexts. *Journal of Experimental Psychology: Human Perception and Performance*, 46(8), 759–788. <https://doi.org/10.1037/xhp0000746>
- de Groot, A. M. B., & Christoffels, I. K. (2006). Language control in bilinguals: Monolingual tasks and simultaneous interpreting. *Bilingualism: Language and Cognition*, 9(2), 189–201. <https://doi.org/10.1017/S1366728906002537>
- Declerck, M., & Koch, I. (2022). The concept of inhibition in bilingual control. *Psychological Review*. <https://doi.org/10.1037/REV0000367>
- Declerck, M., Kleinman, D., & Gollan, T. H. (2020). Which bilinguals reverse language dominance and why? *Cognition*, 204, 104384. <https://doi.org/10.1016/j.COGNITION.2020.104384>
- Declerck, M., Koch, I., & Philipp, A. M. (2012). Digits vs. pictures: The influence of stimulus type on language switching. *Bilingualism: Language and Cognition*, 15(4), 896–904. <https://doi.org/10.1017/s1366728912000193>
- Declerck, M., Grainger, J., Koch, I., & Philipp, A. M. (2017). Is language control just a form of executive control? Evidence for overlapping processes in language switching and task switching. *Journal of Memory and Language*, 95, 138–145. <https://doi.org/10.1016/j.jml.2017.03.005>
- Fink, A., & Goldrick, M. (2014). Pervasive benefits of preparation in language switching. *Psychonomic Bulletin & Review*, 22(3), 808–814. <https://doi.org/10.3758/s13423-014-0739-6>
- Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(5), 1075–1089. <https://doi.org/10.1037/0278-7393.32.5.1075>
- Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021a). Assessing the evidence for asymmetrical switch costs and reversed language dominance effects – A meta-analysis. *Journal of Cognition*, 4(1), 1–32. <https://doi.org/10.5334/JOC.186>
- Gade, M., Declerck, M., Philipp, A. M., Rey-Mermet, A., & Koch, I. (2021b). Erratum: Assessing the evidence for asymmetrical switch costs and reversed language dominance effects – A meta-analysis (Journal of Cognition (2021) 4:1 (55) DOI:10.5334/joc.186). *Journal of Cognition*, 4(1). <https://doi.org/10.5334/JOC.195>
- Gambi, C., & Hartsuiker, R. J. (2016). If you stay, it might be easier: Switch costs from comprehension production in a joint switching task. *Journal of Experimental Psychology: Learning Memory and Cognition*, 42(4), 608–626. <https://doi.org/10.1037/XLM0000190>
- Goldrick, M., & Gollan, T. H. (2023). Inhibitory control of the dominant language: Reversed language dominance is the tip of the iceberg. *Journal of Memory and Language*, 130, 104410. <https://doi.org/10.1016/j.JML.2023.104410>
- Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science*, 22(9), 1155–1164. <https://doi.org/10.1177/0956797611417002>
- Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). What’s easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. *Journal of Experimental Psychology: General*, 143(6), 2167–2195. <https://doi.org/10.1037/a0038006>
- Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In *Language acquisition across linguistic and cognitive systems* (Vol. 52, pp. 267–283). John Benjamins Publishing Company.
- Grange, J. A., & Houghton, G. (2010). Cue-switch costs in task-switching: Cue priming or control processes? *Psychological Research*, 74(5), 481–490. <https://doi.org/10.1007/S00426-009-0270-Y/TABLES/2>
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67–81. <https://doi.org/10.1017/S1366728998000133>
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530. <https://doi.org/10.1080/20445911.2013.796377>
- Grosjean, F. (2001). One mind, two languages: bilingual language processing. In J. Nicol (Ed.), *The bilingual’s language modes* (pp. 1–22). Blackwell.
- Haman, E., Luniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., Blažienė, A., Chyl, K., Dabašinskienė, I., Engel de Abreu, P., Gagarina, N., Gavarró, A., Håkansson, G., Harel, E., Holm, E., Kapalková, S., Kunnari, S., Levorato, C., Lindgren, J., Mieszowska, K., Montes Salarich, L., Potgieter, A., Ribu, I., Ringblom, N., Rinker, T., Roch, M., Slančová, S., Southwood, F., Tedeschi, R., Tuncer, A. M., Ůnal-Logacev, Ů., Vuksanović, J., & Armon-Lotem, S. (2017). Noun and verb knowledge in monolingual

- preschool children across 17 languages: Data from Cross-linguistic Lexical Tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics*, **31**(11–12), 818–843. <https://doi.org/10.1080/02699206.2017.1308553>
- Haman, E., Łuniewska, M., & Pomiechowska, B. (2015). Designing cross-linguistic lexical tasks (clts) for bilingual preschool children. In *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment* (pp. 196–240). Multilingual Matters. <https://doi.org/10.21832/9781783093137-010/MACHINEREADABLECITATION/RIS>
- Heikoop, K. W., Decerck, M., Los, S. A., & Koch, I. (2016). Dissociating language-switch costs from cue-switch costs in bilingual language switching. *Bilingualism: Language and Cognition*, **19**(5), 921–927. <https://doi.org/10.1017/s1366728916000456>
- Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish–English bilinguals: An fMRI study. *NeuroImage*, **14**(2), 510–520. <https://doi.org/10.1006/NIMG.2001.0810>
- Hernandez, A. E., & Kohnert, K. J. (1999). Aging and language switching in bilinguals. *Aging, Neuropsychology, and Cognition*, **6**(2), 69–83. <https://doi.org/10.1076/ANEC.6.2.69.783>
- Hernandez, A. E., Martinez, A., & Kohnert, K. (2000). In search of the language switch: An fMRI study of picture naming in Spanish–English bilinguals. *Brain and Language*, **73**(3), 421–431. <https://doi.org/10.1006/BRLN.1999.2278>
- Jared, D., Pei Yun Poh, R. & Paivio, A. (2013). L1 and L2 picture naming in Mandarin–English bilinguals: A test of bilingual dual coding theory\*. *Bilingualism: Language and Cognition*, **16**(2), 383–396. <https://doi.org/10.1017/S1366728912000685>
- Jevtović, M., Duñabeitia, J. A., & de Bruin, A. (2019). How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Bilingualism: Language and Cognition*, **23**(2), 401–413. <https://doi.org/10.1017/s1366728919000191>
- Jylkkä, J., Lehtonen, M., Lindholm, F., Kuusakoski, A., & Laine, M. (2017). The relationship between general executive functions and bilingual switching and monitoring in language production. *Bilingualism: Language and Cognition*, **21**(3), 505–522. <https://doi.org/10.1017/s1366728917000104>
- Kang, C., Ma, F., Li, S., Kröll, J. F., & Guo, T. (2020). Domain-general inhibition ability predicts the intensity of inhibition on non-target language in bilingual word production: An ERP study. *Bilingualism: Language and Cognition*, **23**(5), 1056–1069. <https://doi.org/10.1017/s1366728920000085>
- Kang, C., Ma, F., & Guo, T. (2017). The plasticity of lexical selection mechanism in word production: ERP evidence from short-term language switching training in unbalanced Chinese–English bilinguals. *Bilingualism: Language and Cognition*, **21**(2), 296–313. <https://doi.org/10.1017/s1366728917000037>
- Kang, C., Fu, Y., Wu, J., Ma, F., Lu, C., & Guo, T. (2017). Short-term language switching training tunes the neural correlates of cognitive control in bilingual language production. *Human Brain Mapping*, **38**(12), 5859–5870. <https://doi.org/10.1002/hbm.23765>
- Kaufmann, E., Mittelberg, I., Koch, I., & Philipp, A. M. (2017). Modality effects in language switching: Evidence for a bimodal advantage. *Bilingualism: Language and Cognition*, **21**(2), 243–250. <https://doi.org/10.1017/s136672891600122x>
- Khateb, A., Shamshoum, R., & Prior, A. (2017). Modulation of language switching by cue timing: Implications for models of bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **43**(8), 1239–1253. <https://doi.org/10.1037/xlm0000382>
- Klecha, A. (2013). Language and Task Switching in Polish-English Bilinguals. *Psychology of Language and Communication*, **17**(1), 17–36. <https://doi.org/10.2478/plc-2013-0002>
- Kleinman, D., & Gollan, T. H. (2016). Speaking Two Languages for the Price of One: Bypassing Language Control Mechanisms via Accessibility-Driven Switches. *Psychological Science*, **27**(5), 700–714. <https://doi.org/10.1177/0956797616634633>
- Lemhöfer, K. & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, **44**(2), 325–343. <https://doi.org/10.3758/S13428-011-0146-0/TABLES/9>
- Lenth, R., & Lenth, M. R. (2018). Package ‘lsmeans’. *The American Statistician*, **34**(4), 216–221.
- Levinson, S.C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, **6**. <https://doi.org/10.3389/fpsyg.2015.00731>
- Linck, J. A., Schwieter, J. W., & Sunderman, G. (2011). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition*, **15**(3), 651–662. <https://doi.org/10.1017/s136672891100054x>
- Liu, H., & Chaouch-Orozco, A. (2022). Is the digit effect a cognate effect? Digits (still) differ from pictures in non-phonologically mediated language switching. *Bilingualism: Language and Cognition*, **26**(3), 469–475. <https://doi.org/10.1017/s1366728922000608>
- Liu, C., Li, L., Jiao, L., & Wang, R. (2021). Bilingual language control flexibly adapts to cultural context. *Frontiers in Psychology*, **12**, 4954. <https://doi.org/10.3389/FPSYG.2021.744289/BIBTEX>
- Liu, C., Timmer, K., Jiao, L., Yuan, Y., & Wang, R. (2019). The influence of contextual faces on bilingual language control. *Quarterly Journal of Experimental Psychology*, **72**(9), 2313–2327.
- Liu, H., Xie, N., Zhang, M., Gao, X., Dunlap, S., & Chen, B. (2018). The electrophysiological mechanism of joint language switching: Evidence from simultaneous production and comprehension. *Journal of Neurolinguistics*, **45**, 45–59. <https://doi.org/10.1016/j.jneuroling.2017.09.002>
- Logan, G. D., & Schneider, D. W. (2006). Interpreting instructional cues in task switching procedures: The role of mediator retrieval. *Journal of Experimental Psychology: Learning Memory and Cognition*, **32**(2), 347–363. <https://doi.org/10.1037/0278-7393.32.3.347>
- Lukas, S., Philipp, A. M., & Koch, I. (2010). Switching attention between modalities: Further evidence for visual dominance. *Psychological Research*, **74**(3), 255–267. <https://doi.org/10.1007/S00426-009-0246-Y/FIGURES/5>
- Ma, F., Li, S., & Guo, T. (2016). Reactive and proactive control in bilingual word production: An investigation of influential factors. *Journal of Memory and Language*, **86**, 35–59. <https://doi.org/10.1016/j.jml.2015.08.004>
- Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, **40**(1), 25–40. <https://doi.org/10.1006/JMLA.1998.2602>
- Mosca, M., & Clahsen, H. (2015). Examining language switching in bilinguals: The role of preparation time. *Bilingualism: Language and Cognition*, **19**(2), 415–424. <https://doi.org/10.1017/s1366728915000693>
- Mosca, M., Manawamma, C., & de Bot, K. (2022). When Language Switching is Cost-Free: The Effect of Preparation Time. *Cognitive Science*, **46**(2). <https://doi.org/10.1111/cogs.13105>
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M., grave le, B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D. & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, **73**, 719–748. <https://doi.org/10.1146/ANNUREV-PSYCH-020821-114157>
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, **41**(6), 407–420. <https://doi.org/10.1016/j.wocn.2013.07.005>
- Peeters, D. (2020). Bilingual switching between languages and listeners: Insights from immersive virtual reality. *Cognition*, **195**, 104107. <https://doi.org/10.1016/J.COGNITION.2019.104107>
- Peeters, D., & Dijkstra, T. (2018). Sustained inhibition of the native language in bilingual language production: A virtual reality approach. *Bilingualism: Language and Cognition*, **21**(5), 1035–1061. <https://doi.org/10.1017/S1366728917000396>
- Peeters, D., Runqvist, E., Bertrand, D., & Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of Experimental Psychology: Learning Memory and Cognition*, **40**(1), 284–292. <https://doi.org/10.1037/a0034060>
- Periáñez, J. A., & Barceló, F. (2009). Updating sensory versus task representations during task-switching: Insights from cognitive brain potentials in humans. *Neuropsychologia*, **47**(4), 1160–1172. <https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2009.01.014>
- 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**(5), 1187–1195. <https://doi.org/10.1037/a0016376>



- Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, *19*(3), 395–416. <https://doi.org/10.1080/09541440600758812>
- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329–347. <https://doi.org/10.1017/S0140525X12001495>
- Prior, A., & Gollan, T. H. (2011). Good Language-Switchers are Good Task-Switchers: Evidence from Spanish–English and Mandarin–English Bilinguals. *Journal of the International Neuropsychological Society*, *17*(4), 682–691. <https://doi.org/10.1017/s1355617711000580>
- Prior, A., & Gollan, T. H. (2013). The elusive link between language control and executive control: A case of limited transfer. *Journal of Cognitive Psychology*, *25*(5), 622–645. <https://doi.org/10.1080/20445911.2013.821993>
- R Core Team. (2020). *R: A language and environment for statistical computing* (Version 4.0.2) [Software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Roychoudhuri, K. S., Prasad, S. G., & Mishra, R. K. (2016). Iconic native culture cues inhibit second language production in a non-immigrant population: Evidence from Bengali–English bilinguals. *Frontiers in Psychology*, *7*, 1516. <https://doi.org/10.3389/fpsyg.2016.01516/BIBTEX>
- Segal, D., Stasenko, A., & Gollan, T. H. (2019). More evidence that a switch is not (always) a switch: Binning bilinguals reveals dissociations between task and language switching. *Journal of Experimental Psychology: General*, *148*(3), 501–519. <https://doi.org/10.1037/xge0000515>
- Shen, Q. (2023). Investigating the modulation of active preparation and passive dissipation on inhibitory control processes in the language switching paradigm. *Frontiers in Psychology*, *14*. <https://doi.org/10.3389/fpsyg.2023.1065268>
- Shen, Q., & Chen, Y. (2023). Investigating the modulation of stimulus types on language switching costs: Do semantic and repetition priming effect matter? *Frontiers in Psychology*, *14*. <https://doi.org/10.3389/fpsyg.2023.1090744>
- Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, *110*(1), 84–99. <https://doi.org/10.1016/j.cognition.2008.10.013>
- Tarlowski, A., Wodniecka, Z., & Marzecová, A. (2012). Language switching in the production of phrases. *Journal of Psycholinguistic Research*, *42*(2), 103–118. <https://doi.org/10.1007/S10936-012-9203-9/METRICS>
- 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*. <https://doi.org/10.3389/fpsyg.2018.01032>
- Timmer, K., Calabria, M., & Costa, A. (2019). Non-linguistic effects of language switching training. *Cognition*, *182*. <https://doi.org/10.1016/j.cognition.2018.09.001>
- Timmer, K., Christoffels, I. K., & Costa, A. (2018). On the flexibility of bilingual language control: The effect of language context. *Bilingualism*. <https://doi.org/10.1017/S1366728918000329>
- Timmer, K., Costa, A., & Wodniecka, Z. (2021). The source of attention modulations in bilingual language contexts. *Brain and Language*, *223*, 105040. <https://doi.org/10.1016/J.BANDL.2021.105040>
- Timmer, K., Grundy, J. G., & Bialystok, E. (2017). Earlier and more distributed neural networks for bilinguals than monolinguals during switching. *Neuropsychologia*, *106*. <https://doi.org/10.1016/j.neuropsychologia.2017.09.017>
- Timmer, K., Wodniecka, Z., & Costa, A. (2021). Rapid attentional adaptations due to language (monolingual vs bilingual) context. *Neuropsychologia*, *159*, 107946. <https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2021.107946>
- Timmermeister, M., Leseman, P., Wijnen, F., & Blom, E. (2020). No Bilingual Benefits Despite Relations Between Language Switching and Task Switching. *Frontiers in Psychology*, *11*. <https://doi.org/10.3389/fpsyg.2020.01832>
- Weissberger, G. H., Wierenga, C. E., Bondi, M. W., & Gollan, T. H. (2012). Partially overlapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging*, *27*(4), 959–974. <https://doi.org/10.1037/a0028281>
- Weissberger, G. H., Gollan, T. H., Bondi, M. W., Clark, L. R., & Wierenga, C. E. (2015). Language and task switching in the bilingual brain: Bilinguals are staying, not switching, experts. *Neuropsychologia*, *66*, 193–203. <https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2014.10.037>
- Woumans, E., Martin, C. D., Bulcke, C. v., van Assche, E., Costa, A., Hartsuiker, R. J., & Duyck, W. (2015). Can faces prime a language? *Psychological Science*, *26*(9), 1343–1352. <https://doi.org/10.1177/0956797615589330>
- Wu, J., Kang, C., Ma, F., Gao, X., & Guo, T. (2018). The influence of short-term language-switching training on the plasticity of the cognitive control mechanism in bilingual word production. *Quarterly Journal of Experimental Psychology*, *71*(10), 2115–2128. <https://doi.org/10.1177/1747021817737520>
- Zhang, M., Wang, X., Wang, F., & Liu, H. (2020). Effect of cognitive style on language control during joint language switching: An ERP study. *Journal of Psycholinguistic Research*, *49*(3), 383–400. <https://doi.org/10.1007/S10936-019-09682-7/TABLES/2>
- Zheng, X., Roelofs, A., & Lemhöfer, K. (2017). Language selection errors in switching: language priming or cognitive control? *Language, Cognition and Neuroscience*, *33*(2), 139–147. <https://doi.org/10.1080/23273798.2017.1363401>
- Zhang, H., Kang, C., Wu, Y., Ma, F., & Guo, T. (2015). Improving proactive control with training on language switching in bilinguals. *NeuroReport*, *26*(6), 354–359. <https://doi.org/10.1097/wnr.0000000000000353>
- Zhu, J. D., & Sowman, P. F. (2020). Whole-language and item-specific inhibition in bilingual language switching: The role of domain-general inhibitory control. *Brain Sciences*, *10*(8), 517. <https://doi.org/10.3390/BRAINS10080517>
- Zhu, J. D., Seymour, R. A., Szakay, A., & Sowman, P. F. (2020). Neurodynamics of executive control in bilingual language switching: An MEG study. *Cognition*, *199*, 104247. <https://doi.org/10.1016/j.cognition.2020.104247>
- Zhu, J. D., Blanco-Elorrieta, E., Sun, Y., Szakay, A., & Sowman, P. F. (2022). Natural vs forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage*, *247*, 118797. <https://doi.org/10.1016/j.neuroimage.2021.118797>



## Appendix A1. Articles using the cued language switching paradigm with cue and stimulus type indicated

Authors (year)	DOI	Cue type	Cue details	Stimulus type
Hernandez and Kohnert (1999)	<a href="https://doi.org/10.1076/anec.6.2.69.783">https://doi.org/10.1076/anec.6.2.69.783</a>	auditory words	(say vs diga)	pictures
Hernandez, Martinez et al. (2000)	<a href="https://doi.org/10.1006/brln.1999.2278">https://doi.org/10.1006/brln.1999.2278</a>	auditory words	(say vs diga)	pictures
Tarlowski, Wodniecka et al. (2012)	<a href="https://doi.org/10.1007/s10936-012-9203-9">https://doi.org/10.1007/s10936-012-9203-9</a>	auditory words	(what? vs co?)	pictures
Chang, Xie, Li, Wang et al. (2016)	<a href="https://doi.org/10.3389/fpsyg.2016.01249">https://doi.org/10.3389/fpsyg.2016.01249</a>	colors		digits
Christoffels, Frirk, Schiller (2007)	<a href="https://doi.org/10.1016/j.brainres.2007.01.137">https://doi.org/10.1016/j.brainres.2007.01.137</a>	colors		picture
Costa and Santesteban (2004)	<a href="https://doi.org/10.1016/j.jml.2004.02.002">https://doi.org/10.1016/j.jml.2004.02.002</a>	colors		picture
de Bruin, Samuel et al. (2018)	<a href="https://doi.org/10.1016/j.jml.2018.07.005">https://doi.org/10.1016/j.jml.2018.07.005</a>	colors		pictures
Declerck, Grainger et al. (2017)	<a href="https://doi.org/10.1016/j.jml.2017.03.005">https://doi.org/10.1016/j.jml.2017.03.005</a>	colors		digits
Declerck, Koch, and Philipp (2012)	<a href="https://doi.org/10.1017/S1366728912000193">https://doi.org/10.1017/S1366728912000193</a>	colors		pictures/digits
Fink & Goldrick (2014)	<a href="https://doi.org/10.3758/s13423-014-0739-6">https://doi.org/10.3758/s13423-014-0739-6</a>	colors		digits
Finkbeiner, Almeida et al. (2006)	<a href="https://doi.org/10.1037/0278-7393.32.5.1075">https://doi.org/10.1037/0278-7393.32.5.1075</a>	colors		digits
Jylkkä, Lehtonen et al. (2017)	<a href="https://doi.org/10.1017/S1366728917000104">https://doi.org/10.1017/S1366728917000104</a>	colors		picture
Kang, Fu, Wu, Ma, Lu et al. (2017)	<a href="https://doi.org/10.1002/hbm.23765">https://doi.org/10.1002/hbm.23765</a>	colors		pictures
Kang, Ma, and Guo (2017)	<a href="https://doi.org/10.1017/S1366728917000037">https://doi.org/10.1017/S1366728917000037</a>	colors		pictures
Kang, Ma, Li, Kroll et al. (2020)	<a href="https://doi.org/10.1017/S1366728920000085">https://doi.org/10.1017/S1366728920000085</a>	colors		pictures
Kaufman, Mittelberg et al. (2017)	<a href="https://doi.org/10.1017/S136672891600122X">https://doi.org/10.1017/S136672891600122X</a>	colors		pictures
Klecha (2013)	<a href="https://doi.org/10.2478/plc-2013-0002">https://doi.org/10.2478/plc-2013-0002</a>	colors		picture
Linck, Schwieter, Sunderman et al. (2011)	<a href="https://doi.org/10.1017/S136672891100054X">https://doi.org/10.1017/S136672891100054X</a>	colors		pictures
Liu and Chaouch-Orozco (2022)	<a href="https://doi.org/10.1017/S1366728922000608">https://doi.org/10.1017/S1366728922000608</a>	colors		pictures/digits
Ma, Li, and Guo (2016)	<a href="https://doi.org/10.1016/j.jml.2015.08.004">https://doi.org/10.1016/j.jml.2015.08.004</a>	colors		digits
<b>Meuter and Allport (1999)</b>	<a href="https://doi.org/10.1006/jmla.1998.2602">https://doi.org/10.1006/jmla.1998.2602</a>	<b>colors</b>		<b>digits</b>
Mosca and Clahsen (2015)	<a href="https://doi.org/10.1017/S1366728915000693">https://doi.org/10.1017/S1366728915000693</a>	colors		pictures
Mosca, Manawamma et al. (2022)	<a href="https://doi.org/10.1111/cogs.13105">https://doi.org/10.1111/cogs.13105</a>	colors		pictures
Olson (2013)	<a href="https://doi.org/10.1017/S0142716415000223">https://doi.org/10.1017/S0142716415000223</a>	colors		pictures
Shen (2023)	<a href="https://doi.org/10.3389/fpsyg.2023.1065268">https://doi.org/10.3389/fpsyg.2023.1065268</a>	colors		pictures
Shen and Chen (2023)	<a href="https://doi.org/10.3389/fpsyg.2023.1090744">https://doi.org/10.3389/fpsyg.2023.1090744</a>	colors		pictures
Timmer, Calabria, and Costa (2019)	<a href="https://doi.org/10.1016/j.cognition.2018.09.001">https://doi.org/10.1016/j.cognition.2018.09.001</a>	colors		picture
Timmer, Christoffels et al. (2018)	<a href="https://doi.org/10.1017/S1366728918000329">https://doi.org/10.1017/S1366728918000329</a>	colors		picture
Timmer, Grundy et al. (2017)	<a href="https://doi.org/10.1016/j.neuropsychologia.2017.09.017">https://doi.org/10.1016/j.neuropsychologia.2017.09.017</a>	colors		picture
Wu, Kang, Ma, Gao et al. (2018)	<a href="https://doi.org/10.1177/1747021817737520">https://doi.org/10.1177/1747021817737520</a>	colors		pictures
Zhang, Cao, Yue, Dai et al. (2020)	<a href="https://doi.org/10.3389/fpsyg.2020.00791">https://doi.org/10.3389/fpsyg.2020.00791</a>	colors		pictures
Zhang, Kang, Wu, Ma et al. (2015)	<a href="https://doi.org/10.1097/WNR.0000000000000353">https://doi.org/10.1097/WNR.0000000000000353</a>	colors		pictures
Zheng, Roelofs et al. (2017)	<a href="https://doi.org/10.1080/23273798.2017.1363401">https://doi.org/10.1080/23273798.2017.1363401</a>	colors		pictures
Blanco-Elorrieta et al. (2018)	<a href="https://doi.org/10.1523/jneurosci.0553-17.2017">https://doi.org/10.1523/jneurosci.0553-17.2017</a>	faces	(race)	picture
Heikoop, Declerck et al. (2016)	<a href="https://doi.org/10.1017/S1366728916000456">https://doi.org/10.1017/S1366728916000456</a>	faces	(female vs male)	pictures
Liu, Timmer, Jiao et al. (2019)	<a href="https://doi.org/10.1177/1747021819836713">https://doi.org/10.1177/1747021819836713</a>	faces	(race)	pictures
Peeters (2020)	<a href="https://doi.org/10.1016/j.cognition.2019.104107">https://doi.org/10.1016/j.cognition.2019.104107</a>	faces	(virtual agents)	pictures
Peeters and Dijkstra (2018)	<a href="https://doi.org/10.1017/S1366728917000396">https://doi.org/10.1017/S1366728917000396</a>	faces	(virtual agents)	pictures
Timmermeister et al. (2020)	<a href="https://doi.org/10.3389/fpsyg.2020.01832">https://doi.org/10.3389/fpsyg.2020.01832</a>	faces	(girl vs boy)	pictures
Zhu, Blanco-Elorrieta et al. (2022)	<a href="https://doi.org/10.1016/j.neuroimage.2021.118797">https://doi.org/10.1016/j.neuroimage.2021.118797</a>	faces	(race)	pictures

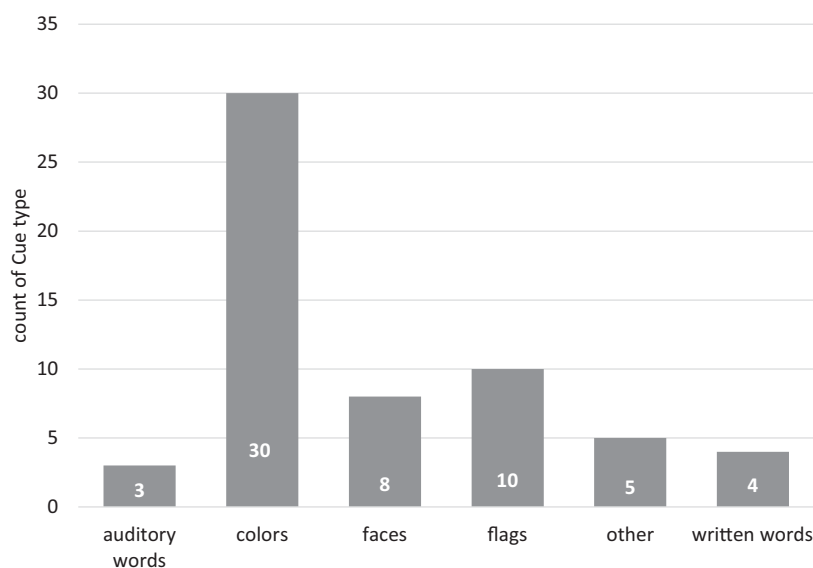
(Continued)

(Continued)

Authors (year)	DOI	Cue type	Cue details	Stimulus type
Zhu, Seymour et al. (2020)	<a href="https://doi.org/10.1016/j.cognition.2020.104247">https://doi.org/10.1016/j.cognition.2020.104247</a>	faces	(race)	digits
Contreras-Saavedra et al. (2020)	<a href="https://doi.org/10.1177/1367006920951873">https://doi.org/10.1177/1367006920951873</a>	flags		digits
de Bruin, Samuel et al. (2020)	<a href="https://doi.org/10.1037/xhp0000746">https://doi.org/10.1037/xhp0000746</a>	flags		pictures
Gollan, Kleinmann et al. (2014)	<a href="https://doi.org/10.1037/a0038006">https://doi.org/10.1037/a0038006</a>	flags		pictures
Jevtic, Duñabeitia et al. (2019)	<a href="https://doi.org/10.1017/S1366728919000191">https://doi.org/10.1017/S1366728919000191</a>	flags		pictures
Kleinman and Gollan (2016)	<a href="https://doi.org/10.1177/0956797616634633">https://doi.org/10.1177/0956797616634633</a>	flags		pictures
Prior and Gollan (2011)	<a href="https://doi.org/10.1017/S1355617711000580">https://doi.org/10.1017/S1355617711000580</a>	flags		digits
Segal, Stasenko, and Gollan (2019)	<a href="https://doi.org/10.1037/xge0000515">https://doi.org/10.1037/xge0000515</a>	flags		digits
Timmer, Calabria et al. (2018)	<a href="https://doi.org/10.3389/fpsyg.2018.01032">https://doi.org/10.3389/fpsyg.2018.01032</a>	flags		pictures
Verhoef, Roelofs et al. (2009)	<a href="https://doi.org/10.1016/j.cognition.2008.10.013">https://doi.org/10.1016/j.cognition.2008.10.013</a>	flags		pictures
Weissberger et al. (2012)	<a href="https://doi.org/10.1037/a0028281">https://doi.org/10.1037/a0028281</a>	flags		digits
Blanco-Elorrieta et al. (2018)	<a href="https://doi.org/10.1073/pnas.1809779115">https://doi.org/10.1073/pnas.1809779115</a>	other	(mouth vs hands)	pictures
Campbell (2005)	<a href="https://doi.org/10.1017/S136672890400207X">https://doi.org/10.1017/S136672890400207X</a>	other	colors and written words together	digits
Olson (2013)	<a href="https://doi.org/10.1016/j.wocn.2013.07.005">https://doi.org/10.1016/j.wocn.2013.07.005</a>	other	colors and written words together	pictures
Philipp and Koch (2009)	<a href="https://doi.org/10.1037/a0016376">https://doi.org/10.1037/a0016376</a>	other	shapes (cross vs diamond)	digits
Philipp, Gade, and Koch (2007)	<a href="https://doi.org/10.1080/09541440600758812">https://doi.org/10.1080/09541440600758812</a>	other	shapes (square vs triangle)	digits
Hernandez, Dapretto et al. (2001)	<a href="https://doi.org/10.1006/nimg.2001.0810">https://doi.org/10.1006/nimg.2001.0810</a>	written words	(say vs diga)	pictures
Khateb, Shamshoum et al. (2017)	<a href="https://doi.org/10.1037/xlm0000382">https://doi.org/10.1037/xlm0000382</a>	written words	(Arabic vs Hebrew)	pictures
Peeters, Runnqvist et al. (2014)	<a href="https://doi.org/10.1037/a0034060">https://doi.org/10.1037/a0034060</a>	written words	(different words)	pictures
Zhu and Snowman (2020)	<a href="https://doi.org/10.3390/brainsci10080517">https://doi.org/10.3390/brainsci10080517</a>	written words	(whats this?)	pictures

Note: Only articles using speech production with pictures and digits are included in the list. The seminal cued language switching article is marked in **bold**. The articles comparing the mentioned cue type against color cues are marked *ursive*.

## Appendix A2. A histogram of the cued language switching articles separated by cue type from Appendix A1



**Appendix B. Experimental stimuli language switching paradigm**

Polish	English
ogon	tail
klatka	cage
gruszka	pear
żółw	turtle
wiewiórka	squirrel
żołnierz	soldier
szklanka	glass
krzesło	chair
dywan	carpet
smok	dragon
oko	eye
kwiat	flower
duch	ghost
ucho	ear
pies	dog
wąż	snake