Journal of Experimental Psychology: Learning, Memory, and Cognition

Putting Language Switching in Context: Effects of Sentence Context and Interlocutors on Bilingual Switching

Angela de Bruin and Veniamin Shiron

Online First Publication, December 14, 2023. https://dx.doi.org/10.1037/xlm0001309

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

https://doi.org/10.1037/xlm0001309

Putting Language Switching in Context: Effects of Sentence Context and Interlocutors on Bilingual Switching

Angela de Bruin and Veniamin Shiron
Department of Psychology, University of York

Many bilinguals switch languages in daily-life conversations. Although this usually happens within sentence context and with another speaker, most research on the cognitive mechanisms underlying the production of language switches has studied individual words. Here, we examined how context influences both switching frequency and the temporal cost associated with it. Sixty Bulgarian-English bilinguals named pictures in their language of choice without any context, in a sentence context, and in interaction with another (recorded) bilingual. Switching frequency was lower, and costs higher, when bilinguals switched languages with context than without context. This suggests switching costs were not an artifact of tasks without context. Furthermore, both switching frequency and costs correlated across the tasks. In addition, we examined the potential influence of sentence context and the conversation partner. Predictability in sentence context has previously been argued to reduce language competition, which in turn could influence switching. We therefore compared sentences with a predictable or unpredictable target word. As hypothesized, bilinguals were less likely to switch languages when a word was predictable in its sentence context, potentially because words in the other language were less active. The conversation partner's overall switching behaviour had little impact on a bilingual's general switching rate, showing relatively low global alignment. However, local alignment was observed as switching was influenced by the partner's switching in the immediately preceding utterance. Together, these findings show that while production tasks without context can reliably measure switching costs, studying effects of context is necessary to better capture a bilingual's language-switching behavior.

Keywords: voluntary language switching, bilingualism, sentence context, predictability, bilingual interlocutors

Supplemental materials: https://doi.org/10.1037/xlm0001309.supp

One remarkable ability closely associated with bilingualism is the ability to switch between multiple languages. However, research studying the cognitive mechanisms underlying language switching can be low in ecological validity (e.g., Blanco-Elorrieta & Pylkkänen, 2018). Much research has focused on cued single-word production tasks, in which participants name individual pictures in response to a cue (e.g., a Spanish flag instructing them to name the picture in Spanish). While these tasks reflect one type of language switching, bilinguals also often switch freely in their conversations with other bilinguals who speak the same languages. Recent research has therefore started to compare these cued naming paradigms to production tasks where bilinguals are free to use the languages and switch as and when they want. In these voluntary naming tasks, participants still name individual pictures but can now use their own language of choice rather than having to respond to a cue. The advantage of picture-naming tasks is that they provide measures such as naming speed (naming onset relative to picture presentation), which can be relatively easily compared across tasks traditionally used in the literature. However, naming individual words without any context does not capture the many contextual factors

Angela de Bruin https://orcid.org/0000-0001-8326-0294

This research was funded by an Experimental Psychology Society (EPS) Small Grant awarded to Angela de Bruin. The research was also presented at the EPS meeting in Stirling, United Kingdom, July 2022 and at European Society for Cognitive Psychology in Porto, Portugal, September 2023.

The study's preregistration, stimuli (picture names and sentences), data, and analysis scripts can be found here: https://osf.io/nuqga/.

The data are available at https://osf.io/nuqga/.

The preregistered design is available at https://osf.io/nuqga/.

Open Access funding provided by University of York:

Correspondence concerning this article should be addressed to Angela de Bruin, Department of Psychology, University of York, York YO10 5DD, United Kingdom. Email: angela.debruin@york.ac.uk

that can influence language production in daily life. The current study therefore examined new paradigms to study the effects of context on bilingual switching in a more naturalistic setting that still provides the well-controlled and frequently used measure of naming times. In addition to comparing how bilinguals switch languages across different types of production with and without context, we also specifically studied the influence of sentence predictability and the bilingual interlocutor (conversation partner) on both switching behavior and costs.

Voluntary Language Switching Without Context

The way bilinguals use their languages and switch is likely related to various factors. It can be driven by bottom-up processes such as speed of lexical access (how fast a word can be produced in each language, e.g., de Bruin et al., 2018) and self-priming and priming by others (e.g., Fricke & Kootstra, 2016). It can also be governed by more top-down processes such as a speaker's communicative intentions (e.g., D. W. Green, 2018), which can include using language switches for more rhetorical functions such as emphasizing information (e.g., Poplack, 1988).

In order to study the cognitive mechanisms involved in language switching, research examining voluntary language switching during production typically presents participants with pictures they can name in their language of choice. In these free dual-language contexts, bilinguals usually generate switch trials (different language to name the picture on trial n than on $n-1$) and nonswitch trials (same language used on trials n and $n-1$). Comparing naming times between switch and nonswitch trials often shows a switching cost, with slower naming when switching languages than on nonswitch trials. The size of this cost, however, varies across studies. Some studies find a similar cost for cued and voluntary switching (e.g., de Bruin & McGarrigle, in press; de Bruin et al., 2018; Gollan et al., 2014) while others find larger cued than voluntary costs (e.g., de Bruin & Xu, 2023; Gollan et al., 2014; Jevtović et al., 2020). Some other studies have not observed voluntary switching costs at all (e.g., Blanco-Elorrieta & Pylkkänen, 2017), although this tends to be the case when bilinguals are encouraged, either through instructions or by choosing language-specific items, to always use the same language for each item (e.g., Kleinman & Gollan, 2016; Zhu et al., 2022).

The presence of voluntary switching costs suggests that bilinguals still need to coordinate the two languages so that they can be used interchangeably in a dual-language environment (D. W. Green & Wei, 2014). Typically, and similar to cued switching, switching costs are interpreted in light of bilinguals experiencing competition between languages when switching. To manage this competition, they might apply language control, for instance in the form of inhibition over the language not currently used (D. W. Green, 1998). However, other interpretations have been proposed too. Language switching might be used as a "tool" to use when speech planning is more difficult or demanding (cf. Beatty-Martínez et al., 2020; Johns & Steuck, 2021). For instance, a language switch could help to manage or recover from particularly high lexical competition or language interference. Following this interpretation, switching costs might reflect a switch being used in a more demanding utterance.

Furthermore, voluntary language production has also been shown to generally recruit less language control than other types of language production. For instance, naming in voluntary dual-language contexts is usually faster than in cued dual-language contexts, but sometimes also than single-language naming (e.g., de Bruin et al., 2018, cf. also Johns & Steuck, 2021). Freely using two languages might require less proactive (sustained) language control than having to use one language only. In the latter case, bilinguals proactively need to avoid interference from the nontarget language. This proactive control might be less necessary when two languages can be used freely (e.g., de Bruin et al., 2018). Thus, in some aspects, using two languages freely appears to place relatively low demands on language control, relative to other language contexts (cf. Adaptive Control Hypothesis, D. W. Green & Abutalebi, 2013).

The presence of switching costs in many picture-naming studies is the main indicator suggesting bilinguals continue to experience competition between languages. The current study therefore focuses on this competition (and the reactive control potentially used to manage it) in relation to the moment of language switching. However, it can also be questioned whether switching costs might be an artifact of the way these costs are measured, namely by having to produce individual words that are not related to each other. If this is the case, switching costs should be smaller or absent when words are produced in sentence context. Furthermore, in that scenario, switching costs and frequency (i.e., how often a bilingual switches) would be expected not to correlate (or to only show a low correlation) when measured in context versus no-context. The current study therefore examined switching frequency and costs in traditionally used picture-naming tasks without context, when naming pictures preceded by a sentence context, and when naming pictures preceded by a sentence in interaction with another bilingual.

Language Switching During Production in Sentence Context

Previous studies examining switching costs during production in sentence contexts have mostly focused on language switching in response to cues, rather than voluntary switching. Furthermore, results are very mixed. Gullifer et al. (2013) asked Spanish–English bilinguals to read sentences in Spanish and/or English. Participants were instructed to read the marked word in the sentence out loud. Naming times did not differ between sentences following a switch versus those preceded by the same language (no switching cost). Naming times were measured for words in the middle of the sentence that were not switches themselves, which could explain the absence of a switching cost. However, Zhang et al. (2014) assessed switching costs closer to the actual language switch and found no behavioral cost either (although neural differences were observed).

Other studies have found switching costs for word production in some types of sentence contexts but not in others. Declerck and Philipp (2015) compared scrambled sentences (no sentence context) to language-unspecific sentences (syntactic structure correct in both languages) and language-specific sentences (syntactic structure correct in one language but not the other). Participants memorized the languages and concepts they had to use to complete each sentence. Switching costs were found in the scrambled and language-specific sentences but not in the language-unspecific sentences. This study suggests that switching costs can be found in some sentence contexts, but perhaps not in the sentence contexts that more typically elicit switches in natural speech (language-unspecific sentences). There are multiple reasons why switching costs might be smaller

in a sentence context. One explanation for this specific study could be that language-specific sentences contained a syntactic structure that differed between languages, which might have increased interference. More generally, a sentence context might also allow more time for planning and response preparation than single-word production. Switching costs in single-word production studies have been found to be smaller—although not absent—when more preparation time is given (e.g., Costa & Santesteban, 2004). Sentence contexts that allow for less preparation, such as scrambled sentences, might therefore show the largest switching costs.

However, using various measures, other studies have shown switching costs across a range of sentence contexts (e.g., through language intrusions, Declerck et al., 2017; naming times, Tarłowski et al., 2013; cf. Declerck et al., 2021 for filled pauses as a measure of costs associated with using two languages). The vast majority of research instructed bilinguals which language to use (or assessed comprehension rather than production, which is not reviewed here). Three studies that assessed voluntary switching in sentence contexts or conversations all observed switching costs too. Faroqi-Shah and Wereley (2022) assessed intersyllabic durations and found these to be longer when participants switched languages in a conversation. Sánchez et al. (2022) assessed filled pauses (e.g., "um") and observed voluntary second language (L2) but not first language (L1) switching costs. Finally, Fricke et al. (2016) used corpus data and found that bilinguals' code switches were preceded by a slower speech rate and a larger cross-language influence on consonant voice onset times than sentences without switches.

Research looking at switching in sentences during language production thus remains limited, especially when considering voluntary switches (as opposed to switching in response to a cue or because a word is presented in a specific language). Most studies find switching costs, but not all. However, most studies do not compare sentence contexts to no-context and it thus remains unclear how exactly sentence contexts influence switching, both in terms of voluntary switching frequency and the costs often associated with it. A direct comparison across tasks is furthermore made more difficult because of the different measures and tasks used across studies in the current literature (e.g., filled pauses, speech rate, word-reading times), which also differ from the measures typically used in single-word studies (naming times in picture-naming paradigms).

On the one hand, sentence contexts might provide a more naturalistic reflection of how bilinguals typically use words and might therefore encourage switching. The coactivation of two grammars could further enhance activation of both languages (cf. MacSwan, 2000), which might increase switching frequency. With respect to switching costs, these might be smaller in sentence contexts than when individual words are not related to each other (Declerck & Philipp, 2015), potentially because sentence contexts allow for more preparation time. However, such effects of potential preparation time have been found to only influence cued but not voluntary language switching (de Bruin & Xu, 2023).

On the other hand, switching frequency might be lower and costs higher in a sentence context. Some types of sentence contexts have been found to decrease coactivation of the other language (e.g., Elston-Güttler, 2000; Van Hell & de Groot, 2008). The linguistic information presented in a sentence context might prime words in the corresponding target language and might partly (although not entirely, cf. Lauro & Schwartz, 2017) reduce the activation of the other language. For voluntary switching, this might reduce the frequency of switching and increase costs when a switch is made.

Thus, switching behavior/frequency and costs might differ between naming in isolation versus context, but it is currently unclear (a) whether switching costs during voluntary switching can be found reliably across different tasks with and without context, (b) whether an individual's switching frequency and costs relate to each other across (no) contexts, and (c) if differences between contexts are observed, how sentence context influences both switching frequency and costs.

Predictable Versus Unpredictable Sentence Contexts

In addition to studying how switching costs and frequency relate across contexts, we also examined two specific features of context that have been found to impact language production more generally. One important feature of sentence contexts that can modulate the activation of the two languages is predictability. Lexical activation is closely associated to language choice and voluntary switching. For instance, bilingual language choice depends on how quickly a word can be accessed in each language (de Bruin et al., 2018). We therefore studied the influence of predictable versus unpredictable sentence contexts on switching frequency and costs.

The effects of sentence predictability are often studied by assessing the cognate effect as a measure of language coactivation. For instance, Schwartz and Kroll (2006) presented participants with cognates, words that are similar in form and meaning across two languages (such as piano in English and Spanish), and control words. In some sentences (e.g., "When we entered the dining hall we saw the piano in the corner of the room"), the sentence context did not predict the occurrence of "piano" specifically. However, the target "piano" was predicted to occur in other sentences (e.g., "Before playing, the composer first wiped the keys of the piano at the beginning of the concert"). In line with previous studies, cognates were processed faster than control words in unpredictable sentences. However, this cognate facilitation effect was absent in predictable sentence contexts. This suggests that when the sentence is predicting a specific word, coactivation of the other language might be reduced (see also e.g., Elston-Güttler, 2000; Van Hell & de Groot, 2008, for similar findings). Although not all studies have found differences between predictable and unpredictable sentences (e.g., Van Assche et al., 2011), a recent meta-analysis (Lauro & Schwartz, 2017) showed that, across studies, cognate facilitation is present in both types of sentence contexts but smaller in predictable sentences.

These findings firstly suggest that even in predictable sentence contexts, the other language remains active and competes for selection. Thus, in line with nonselective models of bilingual processing and production (e.g., Dijkstra & Van Heuven, 2002; D. W. Green, 1998), neither type of sentence context fully restricts activation to the target language of the sentence. However, coactivation of the other language might be smaller when a word is predictable. In those contexts, a combination of semantic, syntactic, and language features and expectations can restrict coactivation of related word representations in the other language (cf. Dijkstra et al., 2015). However, these differences between sentence types might mostly affect later processing stages. Libben and Titone's (2009) eye-tracking study showed cross-linguistic effects in predictable sentences in early fixation measures but not for total reading times. This suggests coactivation between languages could influence the

early stages of lexical selection in both predictable and unpredictable sentences, with sentence effects more likely to emerge for later processing stages.

The literature regarding sentence predictability as reviewed above focuses on the listener or reader comprehending sentences. An open question therefore remains whether potential differences in language/lexical coactivation between predictable and unpredictable sentence contexts can influence language production and in particular also language switching (including earlier stages of planning and processing, i.e., language/lexical choice related to language switches). The findings from the comprehension literature regarding predictability modulating language coactivation would be expected to also apply to language production and switching. Focusing on switching frequency first, language switching might be more likely when the "other language" (i.e., the one not currently in use) is more active and competes more strongly for selection. If the other language is less active in predictable sentences, participants might be less likely to switch languages than in unpredictable sentences.

This account focuses on speaker-based factors, in the form of speaker-internal lexical activation influencing the speaker's choice. However, this prediction regarding switching frequency is also in line with listener-based factors focusing more on communicative strategies to help the listener. Using corpus data, Myslín and Levy (2015) examined the switching patterns of Czech-English bilinguals and found that word predictability was closely associated with switching. Bilinguals switched more often when a word's meaning was less predictable in that context or conversation. The explanation offered argues that speakers potentially switch languages to cue the listener that less predictable information is coming up. In this sense, language choice and switching can be used as a marker of informational content, with bilinguals potentially helping the listener by switching when words or meanings are less predictable or expected. Indeed, listeners can benefit from this. Tomić and Valdés Kroff (2022) used eye-tracking to examine how Spanish-English bilinguals process high- and low-frequency words in sentence contexts with and without language switches. Participants were more likely to look at low-frequency items in sentences containing switches, suggesting they were using these switches as a cue anticipating the presentation of less predictable words.

Thus, when words are less predictable, bilinguals might experience more coactivation from both languages and they might adjust their language-switching behavior to help their audience. Both speaker-based and listener-based accounts predict that bilinguals are more likely to switch when a target word is not strongly predicted by the preceding context. Furthermore, switching costs might also differ between predictable and unpredictable sentence contexts. If predictable sentence contexts reduce language coactivation, switching costs should be larger in predictable contexts. The activation of the other language (i.e., the language the bilingual is switching to) is argued to be lower in a predictable context while activation of the currently used language is higher and interfering more. As a consequence, switching to the other language might take more time in predictable contexts than in unpredictable contexts where coactivation of the other language is higher. However, a purely listener-based account could predict the opposite for naming times. If speakers switch in unpredictable contexts to help the listener in a manner that does not benefit the speaker, costs might be larger in unpredictable contexts instead, as a consequence of having to adjust your behavior in favor of the listener.

Bilingual Interaction With an Interlocutor

So far, we have focused on sentence context but language users also often communicate in interaction with a conversation partner (interlocutor), who can influence their lexical choice (sofa or couch) and syntactic alignment (e.g., Garrod & Pickering, 2004), as well as language choice (Kootstra et al., 2020) and switching (e.g., Kootstra et al., 2020). For instance, Kootstra et al. (2020) found that bilinguals were more likely to switch languages themselves immediately after a switch produced by the interlocutor. This lab-based evidence aligns with corpus data (Fricke & Kootstra, 2016). This interactive alignment can occur as the consequence of more automatic mechanisms closely related to priming. For instance, the interlocutor using Language A can increase the overall baseline activation of that language, thus facilitating a bilingual's retrieval of words in that language. However, alignment can also occur for more social or pragmatic reasons, including wanting to facilitate the listener's processing by using their preferred words or language (e.g., Kapiley & Mishra, 2019) and through aiming to create a shared identity or be viewed more positively (e.g., Chartrand & Bargh, 1999). Kootstra et al. (2020) suggested that the interactive alignment of switching might be more closely associated with priming and activation than with alignment for more social or pragmatic reasons. Their Experiment 2 showed that bilinguals were also more likely to switch to the nondefault language after hearing single-language sentences (without any switches) in that language. This suggests that the increased activation of the nondefault language increased the likelihood of using that language, and thus of switching, in their study.

However, their work focused on the immediate (local) effects of an interlocutor's use on a bilingual's utterance. Pragmatic or social alignment might occur at a more global level across a conversation, by "mimicking" the overall language behavior of the interlocutor. Furthermore, the pure presence of a bilingual interlocutor can influence the way bilinguals process language switching (Tomić & Kaan, 2022), although it is unclear how this relates to the production of switches. The current study therefore examined the influence of an interlocutor's global switching frequency on a bilingual's own global switching frequency. By keeping the actual language choice of the interlocutor consistent (half Language A, half Language B), we were able to study global alignment in switching frequency, irrespective of language activation. Furthermore, unlike previous studies, we were not only able to examine effects of the interlocutor's switching behavior on the participant's switching frequency, but also on switching costs. If bilinguals align their switching frequency to the interlocutor for pragmatic or social reason, this is likely to be a less automatic process and might therefore increase overall naming times and/or switching costs. However, if bilinguals align their switching frequency through more unmediated processes like priming, switching costs should not be influenced or even smaller in the high-switching interlocutor condition.

In addition, as more exploratory analyses, we also examined the immediate, local influence of the interlocutor's behavior on switching and language choice. With respect to local effects of language switching, similar to Kootstra et al. (2020), we examined whether bilinguals switched more often immediately after the interlocutor switched in their utterance. Importantly, any switching alignment effects in our study could not be because of "lexical boost" effects. Previous research has suggested alignment effects might be (partly)

driven by a lexical boost as the consequence of lexical items being repeated across the interlocutor's and participant's utterances (e.g., verb repetition, Pickering & Branigan, 1998). In the current study, the interlocutor's preceding utterance differed from the lexical items to be used in the bilingual's utterance, allowing us to examine switching alignment in the absence of such lexical boost. Furthermore, previous studies have shown that bilinguals can also align in terms of their language choice (e.g., Fricke & Kootstra, 2016). Given that our study was set up so that the interlocutor always used each language half of the time, we did not examine global language-choice alignment. However, we did examine whether bilinguals aligned their language choice on specific lexical items by studying if bilinguals were more likely to use the language the interlocutor previously used for that specific picture. Together, these planned and more exploratory analyses allowed us to examine alignment at a more global and more local level, in terms of both switching and language choice.

Current Study

In summary, the current study aimed to examine voluntary language switching across paradigms that differ in the presence and type of context. At a more practical level, we wanted to examine new tasks that allow for comparisons between switching across different contexts, and that allow for a comparison with previous work focusing on word production without context. A few different paradigms have been developed to study (cued and voluntary) language production in a sentence context (e.g., Sánchez et al., 2022), but these paradigms vary in the type of measures used, both from each other and from the no-context naming tasks typically used in the literature. Given that the vast majority of psycholinguistic research on language control during production examines naming onset times, here we asked participants to name pictures without context (no-context), after reading the start of a sentence out loud (context-sentence), and in interaction with another bilingual (context-interlocutor). Presenting bilinguals with a sentence, rather than allowing them to freely generate sentences, allowed us to control the content of those sentences and manipulate the role of sentence predictability.

As a first question (Q1), we examined how bilinguals switch languages (in terms of frequency and costs) in these different environments. We compared the influence of sentence context (naming a picture completing a sentence vs. no-context) and the influence of an interlocutor (naming a picture in a sentence context with or without a recording of another bilingual interlocutor). We examined whether these two types of context modulated the presence, size, and reliability of switching costs as well as the relationships across tasks.

On the one hand, context might provide a more naturalistic environment and additional time for response planning, which could result in lower switching costs (e.g., Declerck & Philipp, 2015) and potentially more frequent switching. Furthermore, if switching costs in voluntary picture naming tasks are artifacts of un-naturalistic tasks requiring naming without relationships between words or context, we would expect switching costs to only arise in the typically used no-context task. On the other hand, context might reduce language coactivation (e.g., Van Hell & de Groot, 2008), which could result in less switching and higher costs. Our main comparison of interest was therefore between the no-context and context-sentence (no interlocutor) tasks. The context-interlocutor task was mostly included to examine the effects of the interlocutor's switching behavior. However, we also compared the context-sentence and context-interlocutor tasks to examine if frequency and costs are influenced by the presence of an interlocutor more generally (cf. Tomić & Kaan, 2022). Due to the many differences between the no-context and context-interlocutor tasks, and as preregistered, we did not compare those two conditions directly.

Although we made the three tasks as similar as possible by working with the same dependent variables (DVs), stimuli, and participants, some differences between tasks remained that could influence the absolute size of switching frequency and costs. We therefore also examined the correlations and reliability of switching frequency and costs across tasks. If switching costs are observed in all tasks but are reflecting different underlying mechanisms, we would expect frequency and cost correlations across tasks to be low or absent. However, if switching frequency and costs correlate across tasks and are comparable in terms of internal reliability, this would provide support for more systematically arising switching behavior across context and no-context tasks.

Furthermore, we examined the influence of sentence context and interlocutors on language switching to assess the role of specific contextual factors within the same task. Within the context-sentence task (Q2), we compared predictable sentence contexts that predicted a specific target word to more neutral, unpredictable contexts that did not predict one specific final target word. If bilinguals experience less parallel language activation in predictable contexts (e.g., Van Hell & de Groot, 2008), we expected fewer switches and larger switching costs in predictable than unpredictable contexts. Within the context-interlocutor task (Q3), we compared interacting with a high- or low-switching interlocutor. If the overall switching pattern of an interlocutor influences a bilingual's global switching pattern, we expected lower switching frequency (and possibly higher switching costs) when interacting with a bilingual who switches rarely. Our main focus was on the overall behavior (switching frequency) of the interlocutor. However, in line with previous studies (e.g., Kootstra et al., 2020), we also conducted additional exploratory analyses examining the local influence of the interlocutor's switching in the preceding sentence and the influence of the interlocutor's language choice for a given item.

Method

Transparency and Openness

The study plan and analysis were preregistered. The data and analysis scripts, together with the preregistration, can be found here: https://osf.io/nuqga/de Bruin & Shiron (2023). A list of all stimuli used (picture names and sentences) can also be found on the Open Science Framework (OSF) page. The Methods section below reports all measures and manipulations used in the study explains how we determined sample size, and reasons for participant exclusion from data analysis. The data analysis and Results sections provide details about the software used to analyze the data. Exploratory analyses are presented in a separate section of the Results, labeled as being exploratory.

Participants

The study was completed by 60 Bulgarian-English bilinguals, who were recruited through Prolific.co. An additional 22 participants were tested but either did not complete all naming tasks ($N = 11$),

had problems with their recordings being inaudible, empty, or not aligning with the instructions regarding sentence reading ($N = 10$), or did not meet our eligibility requirements regarding not having any known language difficulties ($N = 1$). We initially intended to recruit 78 participants but included stopping at 60 participants as a possibility in our preregistration if recruitment proved difficult (which was the case, partly because of the large number of participants that had to be excluded, and partly because no more participants volunteered through Prolific.co). Power analyses based on language-switching in isolation (using simr, P. Green & MacLeod, 2016, with power estimates based on de Bruin et al., 2018) showed that ~30 participants gave over 80% power to detect switching costs of the size previously observed in voluntary switching tasks without context. Further power analyses were difficult as there are no previous studies providing clear effect sizes for effects of sentence predictability and global switching frequency of the interlocutor on switching cost and frequency. All participants ($M_{age} = 30.2$ years, $SD = 8.5$; 10 Male) had normal or corrected-to-normal vision, were not color blind, and had no known neurological, language, or reading impairment. All apart from six were right handed. The study was approved by the Ethics Committee in the Department of Psychology at the University of York. Participants provided informed consent online at the start of the study.

All participants were native speakers of Bulgarian, with a mean age of acquisition in English of 10.8 years ($SD = 7.4$). The majority of participants (53/60) were living in an English-dominant country (e.g., United Kingdom) at the moment of testing and participants, only including those living in an English-dominant country, on average had spent 10 years in that country (range 2–22 years). Proficiency was assessed in both languages through self-ratings and two objective measures. One assessed vocabulary through a written picture naming task based on de Bruin et al. (2017), in which they typed the English and Bulgarian names for 65 pictures. Correct responses with minor typos were accepted. Proficiency was also assessed through a cloze test. In this story completion task, participants read a story about Little Red Riding Hood with 25 blanks (based on Ivanov, 2012). They were asked to complete the blanks with a fitting word. Answers were accepted if they were grammatically correct and fitting the story. Participants' proficiency was high in both languages in all assessments (see Table 1).

Participants were also asked to indicate their language use on a scale of 1 (*all Bulgarian*) to 5 (*all English*) for different time frames (childhood or now), activities (e.g., reading, watching TV), and with different people (e.g., family and friends; a shorter version based on Anderson et al., 2018). For questions asking about language use during childhood/teenage years (four questions), participants reported more frequent use of Bulgarian ($M = 2.0$, $SD = 1.0$). For questions (16 questions) asking about current language use, participants reported balanced or more frequent English use ($M = 3.8$, $SD = 0.6$). Participants also reported (on a scale from 1—*never* to 7—*very frequently*) moderate to frequent language switching on a daily basis ($M = 5.3$, $SD = 1.6$), within conversations ($M = 5.0$, $SD = 1.9$), and within sentences ($M = 4.5$, $SD = 2.0$).

Table 1
Summary of the Participants' Language Proficiency in Bulgarian and English

Language proficiency measure	Bulgarian	English
Self-rated proficiency (1–10)[a]		
Speaking	9.7 (0.8)	8.6 (1.5)
Understanding	9.9 (0.5)	9.1 (1.3)
Writing	9.1 (1.9)	8.7 (1.6)
Reading	9.5 (1.3)	9.2 (1.0)
Written picture naming task (0%–100%)[b]	94.8% (2.8)	96.5% (3.4)
Cloze test (0%–100%)[c]	93.7% (5.4)	89.3% (13.4)

Note. Means (and *SD*s) are provided for self-rated proficiency, the written picture naming task, and the cloze test.
[a] Questionnaire data were missing for one participant. [b] Data are missing from five participants because they either did not provide any answers, their Bulgarian answers could not be scored because they did not answer in Cyrillic, or because they completed the task in the wrong language. [c] Data are missing from six participants because they did not complete the task in both languages (often because of keyboard-related issues when typing Bulgarian responses).

Design

Participants completed three different switching tasks. All analyses included two DVs: switching frequency (i.e., how often participants switched languages) and reaction times (RTs, i.e., naming onset time relative to the onset of picture presentation). We compared these DVs across the three versions of the switching task: naming pictures without any context ("no-context"); naming pictures after reading a sentence ("context-sentence"); and naming pictures after reading a sentence in interaction with another bilingual ("context-interlocutor"). Trial type (switch or nonswitch) was determined relative to the language used for the previous picture in the no-context task, in line with previous studies. However, for the context conditions in which the word was presented as part of a sentence, we determined trial type relative to the language used in the first part of the sentence. Here, a trial type was a switch if the language used to name the target picture differed from the language of the sentence (regardless of the language the participant used to name the previous picture).

Within the context-sentence task, we manipulated predictability as an additional independent variable within participants, with the sentence context predicting the upcoming target word or being neutral (not predicting a specific picture/word). Within the context-interlocutor task, we manipulated the switching frequency of the interlocutor, with half of the participants interacting with a frequent-switching interlocutor and the other half interacting with a low-switching interlocutor.

Materials

We selected a set of 160 target pictures. Each picture was presented in each task but not repeated within the task. Most pictures were taken from the MultiPic database (Duñabeitia et al., 2018), with a few additional pictures added through Google Images using a similar style. The full list of target words and sentences can be found on the OSF page, including an overview of the stimulus characteristics. Sentences in the predictable and unpredictable conditions were matched per language in terms of number of words, target frequency, and number of phonemes and syllables of the target (English only, Bulgarian words could not be matched on word length). Sentence length was also matched between the languages, but because of the nature of the two languages, target words could not be matched between English and Bulgarian in terms of frequency and length. However, this matching was not strictly necessary given that we

were not interested in the effects of or differences between languages in this study (see Data Analysis section).

In the no-context task, each picture was presented once without any context. In the context-sentence task, half of the pictures were presented in a predictable context while the other half was presented in an unpredictable context. Each sentence was presented once in this task. Finally, in the context-interlocutor task, each picture was presented within an unpredictable sentence context. To do this, 80 neutral sentences were presented twice (with two different target words). We opted for unpredictable sentence contexts in this task to simplify the design and because we expected this context to elicit more switches.

There were several reasons why we worked with Bulgarian and English. We aimed to use a language combination that allowed for the use of the same word order in the sentences while also giving us sufficient word options for our noncognate stimuli. For instance, working with more typically studied language pairs like French-English or Spanish-English would meet the word order requirements but, given the relatively high number of cognates, would have made word and sentence creation very difficult. Bulgarian-English allowed us to select noncognates only while still working with a subject–verb–object (SVO) sentence order in both languages. While multiple word orders are possible in Bulgarian, SVO is a common and neutral word order. The sentences we created therefore always started with a neutral subject (e.g., "she" or "the man"), followed by a verb. In most cases, the sentence context only included the subject and verb, with participants asked to generate the object in response to the pictures. In a few cases, sentences also included prepositions, which occur just before the noun phrase in both languages. One notable difference between English and Bulgarian concerns the position of the article, which is placed as a suffix at the end of the noun in Bulgarian. Definite articles in Bulgarian furthermore differ for masculine and feminine nouns and depend on the end of the noun being a consonant or vowel.

Sentence Predictability

Within the context-sentence task, half of the sentences predicted a specific target word (e.g., "She reads…the book," "He milks…the cow," or "She braids…the hair") while the other half did not predict a specific word (e.g., "She holds…the fork," "She sells…the mirror," or "He hides…the pear").

We ran a pilot study with six people to evaluate the predictability of the target word in the sentence, in addition to a separate pilot with nine participants conducted to ensure participants would switch languages in these sentence contexts. In the predictability checks, participants were presented with the beginning of the sentence and were asked to write down two possible target words. All participants saw all sentence-target pairs. Half of the participants saw one-half of the sentences in English and the other half in Bulgarian, with the language of the sentence presentation reversed for the other half of the participants.

For each participant, if one of their two responses contained the target word (including plural responses like "forks" instead of "fork"), they received a score of 1 for that item. If neither of their responses for that sentence contained our target word, their score was 0. If participants provided a similar word that was not identical to the target (e.g., "criminal" for "burglar" or "luggage" for "suitcase"), we did not count these as a target response. Several predictable sentences were therefore still given a score of 0, despite participants giving very similar responses to the target. Across languages, participants more often predicted our target words in predictable contexts (by-item $M = 26\%$, $SD = 27$) than in unpredictable contexts (by-item $M = 1.9\%$, $SD = 8.4$; a significant difference: Mann–Whitney $U = 5{,}111$, $p < .001$).

Comparisons for each language separately also showed similar differences between the predictable and unpredictable contexts. Two items were initially in the unpredictable condition but were moved to the predictable context when preparing the final set, as they were named as potential targets by multiple pilot participants. Almost all other unpredictable items (with the exception of three targets) were not named at all by the participants. Based on the pilot data, a few additional changes were made to the stimuli to increase the difference between predictable and unpredictable contexts. For the final sentence set, using latent semantic analysis (Landauer et al., 1998), we computed the relationship between the sentence's verb and the target noun. This relationship was higher for predictable sentences ($M = 0.30$, $SD = 0.17$) than for unpredictable sentences, $M = 0.11$, $SD = 0.08$; $t(158) = 8.690$, $p < .001$.

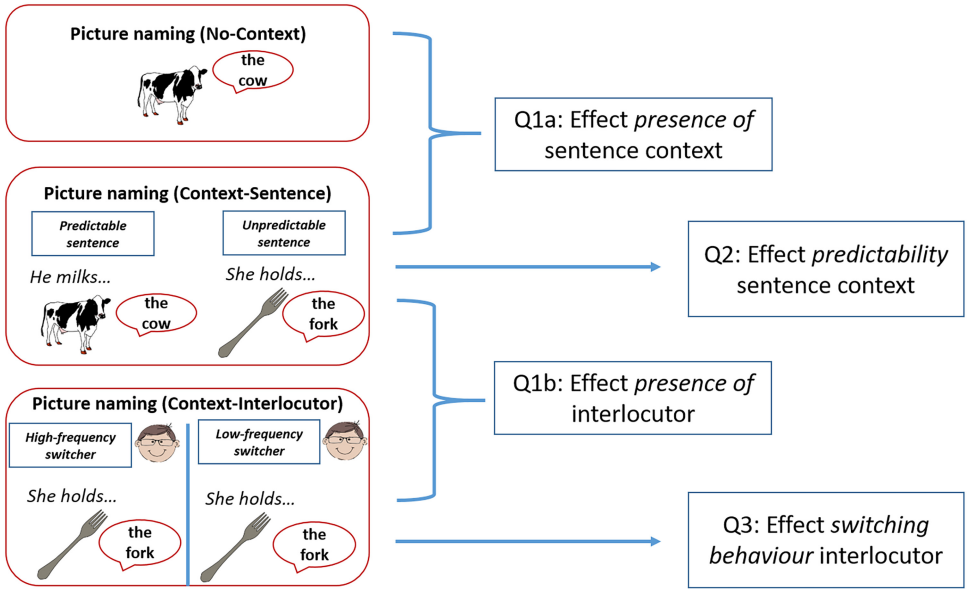
Due to the constraints in terms of items not always fitting a predictable sentence context, constraints around sentence complexity, avoidance of cognate words, and use of easy-to-name pictures only, we were not able to counterbalance pictures across the predictable and unpredictable sentence contexts. In the context-sentence task, each picture therefore occurred either in a predictable or in an unpredictable sentence context. To assess whether there were any baseline differences between these pictures in terms of the measures of interest, we compared switching costs and switching frequency between these sets of items during the baseline picture naming task without context. Neither switching costs, M predictable set = 120 ms, $SD = 135$; M unpredictable set = 133 ms, $SD = 126$, $t(58) = -0.728$, $p = .470$, nor switching frequency, M predictable set = 33.9%, $SD = 12.0$; M unpredictable set = 34.3%, $SD = 11.2$, $t(59) = -0.369$, $p = .714$, differed significantly between the two sets. Naming language, M predictable set = 55.1% English, $SD = 17.2\%$; M unpredictable set = 55.0% English, $SD = 18.6\%$, $t(59) = 0.159$, $p = .875$, did not differ either.

Procedure

The study was run online through Gorilla.sc (Anwyl-Irvine et al., 2020) and consisted of three sessions, which were counterbalanced in order across participants. Each session included one of the three naming tasks (see Figure 1; no-context, context-sentence, context-interlocutor) and lasted approximately 30–45 min. The first and second sessions were, on average, separated by 11 days ($SD = 9$) and the second and third sessions by 13 days ($SD = 9$). Each session started with a microphone check, asking participants to record themselves and to make sure they could hear their own recording before continuing. All tasks used the same set of 160 pictures (see "Materials") that participants were free to name in Bulgarian or English. Each picture was only presented once in each task and there was no familiarization phase. Participants were always instructed that they could use both languages interchangeably and that they could switch whenever they wanted. They were asked to just use the word that came to mind fastest regardless of the language. They were also asked to name the picture with the article

Figure 1

Overview of the Three Switching Tasks (Left) and the Questions Related to the Comparisons of Conditions Between and Within Tasks



Note. See the online article for the color version of this figure.

(e.g., "the dog"), to ensure comparability across the different contexts (i.e., article naming was required in the sentence contexts to ensure grammaticality). Each task was preceded by two practice trials using different items than the experiment and participants could take a break in the middle of each task.

Picture Naming Without Context

In this session, participants saw one picture at a time and named the picture without any context. Each trial started with a fixation cross for 500 ms, followed by the picture staying on the screen for 3 s. This task was followed by the proficiency measurements and questionnaire (see "Participants").

Picture Naming in Sentence Context (Without an Interlocutor)

In this task, participants were still instructed to name the pictures but now saw the beginning of a short sentence that they were asked to read out loud prior to each picture. Each trial started with a fixation cross for 500 ms, followed by the sentence presented on the screen. This sentence included the part prior to the noun phrase (article and noun) they had to use to name the picture (e.g., "They saw…"). Participants read the sentence out loud and pressed space when they finished reading. Sentence reading was recorded during the practice phase and was checked to make sure participants completed the task correctly, which was the case for all included participants. After they finished reading the sentence, another fixation cross was presented for 500 ms, followed by the picture staying on the screen for 3 s. The fixation cross was presented for 500 ms prior to picture presentation to avoid interference if participants already pressed space to see the picture while they were still finishing saying the final word of the sentence.

Half of the trials used predictable sentences; the other half used unpredictable sentences. Within each condition, half of the sentences were presented in Bulgarian and the other half in English, with the language of the sentences counterbalanced across participants. The order of trials was pseudorandomized so that no more than four trials of the same condition or language appeared in a row. Each picture was presented with a Bulgarian sentence to half of the participants and with an English sentence to the other half. We also ensured that language switches in terms of the start of the presented sentence (e.g., Bulgarian sentence followed by an English sentence) were distributed equally across predictable and unpredictable conditions, languages, and pictures.

Picture Naming in the Sentence Context With an Interlocutor

In the task with the interlocutor, participants and the interlocutor read sentences and named pictures interchangeably. Each trial started with an audio recording of the interlocutor reading a sentence including the target word. The participant was then presented with four pictures on the screen and asked to click on the one matching the target word. Next, the participant read the sentence and named their picture (similar to the sentence context without an interlocutor). After they finished naming their picture, they saw a screen saying "[name of interlocutor's] go!" for one second, in line with the instructions indicating that the interlocutor would also see four pictures and select one matching the participant's response. The participant and interlocutor used the same sentences and target words.

All participants interacted with the same interlocutor, but for 31 participants the interlocutor switched frequently throughout the

task and in the introduction video while he switched only rarely with the other 29 participants. Due to the COVID-19 pandemic, the entire study was conducted online. The interlocutor therefore introduced himself through a recorded video at the start of the study. In both conditions, the introduction video contained approximately 200 words, of which half were in Bulgarian and half in English. Regarding the introduction video, in the high-switching condition, approximately 10% of the words in the introduction text were a language switch. In the low-switching condition, there was only one language switch in the middle of the introduction text to make sure both languages were used. The introduction video contained some information about the interlocutor and explained the task the participant was going to do together with the interlocutor.

In the high-switch condition, the interlocutor switched on half of the sentences he produced in the task (40 switches to Bulgarian and 40 to English). Similar to the participant, the switch was always on the target word. Four different list versions were used. In the low-switch condition, the interlocutor switched on 10% of the trials (eight switches to Bulgarian and eight switches to English). Twenty different lists were used. All lists were set up so that, across the lists, the 160 interlocutor's target productions and the 160 participant's target word were distributed equally across the preceding interlocutor's English, Bulgarian, English-Bulgarian, and Bulgarian-English sentences. However, because of the participant number being lower than planned, there is some variability across target items with respect to how often they were preceded by each type of interlocutor utterance.

Data Analysis

Responses were scored and/or checked for accuracy by a Bulgarian-English bilingual. Given that there was no familiarization phase, we scored responses as correct if they were the target word or a similar alternative (e.g., "luggage" for "suitcase"). Responses were scored as incorrect if no response was given at all, if a completely different word was used, or if the response combined English and Bulgarian. RTs were scored in CheckVocal (Protopapas, 2007, using CheckFile).

Our analyses focused on switching frequency and RTs as the DVs. Switching frequency was based on trial type (0 = *nonswitch*, 1 = *switch*). RT analyses used log-transformed RTs as the raw data were not normally distributed. All means reported in the article (including figures and tables) are based on untransformed RTs. All analyses only included correct responses. For the no-context task, we also excluded trials preceded by a break or by a no response (or a response combining two languages) as trial type could not be determined in those cases. In the context tasks, trial type was determined relative to the language of the sentence and we therefore did not exclude trials after the break or after a no response. Prior to the RT analyses, we removed RT outliers that fell 2.5 *SD* above or below the mean by participant and condition (1.2% of correct trials, Grange, 2015).

Switching frequency data were analyzed using generalized linear mixed-effects models and RTs through linear mixed-effects models, using package lme4 (Version 1.1.21) and lmerTEST (Version 3.1.3) in R (Version 3.6.1.). We started with maximal models including by-participant and by-item intercepts and all within-participant/-item slopes. When models did not converge, we first removed correlations between slopes and intercepts, followed by the removal of item slopes that explained the lowest amount of variance. Details about the random-effects structure of the converging models are provided in the Results section per analysis.

First, we examined the effect of context on switching frequency and costs. As preregistered, we conducted two separate analyses. One compared the effect of context-sentence (no interlocutor) relative to no-context. The second compared the effect of having an interlocutor versus not having an interlocutor, comparing context-sentence to context-interlocutor. We opted for two separate analyses, rather than combining all three tasks in one analysis, because of the many differences between the context-interlocutor and the no-context tasks. These two separate analyses allowed us to focus specifically on the two core variables of interest: sentence context and interlocutor. The switching frequency analysis included Task (no-context = −0.5 vs. context-sentence = 0.5; context-interlocutor = −0.5 vs. context-sentence = 0.5) and language (Bulgarian = −0.5; English = 0.5) as main effects. As preregistered, we only included language as a main effect to take into consideration language-related differences when naming pictures. We did not allow language to interact with other variables in the model. This decision was made for a few reasons. First, as expected, the bilinguals we recruited ranged in their proficiency in and use of Bulgarian. While they all acquired Bulgarian from birth, many participants were using English more in their daily lives and had high proficiency in both languages, making it difficult to establish a clear first and second languages. Second, our stimuli could not be matched perfectly between the two languages and we therefore did not aim to examine specific interactions between language and, for example, predictability effects. Finally, because of the nature of the tasks, we could not work with more than 160 trials per task and allowing language to interact with other variables would likely lead to overfitting the data without collecting more trials.

The RT analysis was set up the same way, but now also included Trial type (nonswitch = −0.5; switch = 0.5) as a main effect and in interaction with Task. We also examined the correlation across the tasks by computing the mean switching frequency (% switch trials relative to number of correct trials) and switching cost (mean switch–nonswitch RT) per participant and task.

We furthermore examined the roles of sentence predictability (context-sentence task only; predictable = −0.5; unpredictable = 0.5) and interlocutor (context-interlocutor task only; low-switching interlocutor = −0.5; high-switching interlocutor = 0.5). Again, we examined the potential influence on switching frequency and RTs as the DVs.

Results

Accuracy was high in all three tasks (no-context: $M = 93.5\%$, $SD = 4.1$; context-sentence: $M = 92.8\%$, $SD = 5.1$; context-interlocutor: $M = 93.9\%$, $SD = 3.0$). Although accuracy was slightly below 95% (which we preregistered as the accuracy cutoff to analyze accuracy), we did not analyze the accuracy data further as most errors were because of participants using the wrong name for a picture because they did not recognize it. In all tasks, English was used somewhat more often than Bulgarian (no-context: $M = 55.1\%$, $SD = 17.5$; context-sentence: $M = 56.0\%$, $SD = 16.0$; context-interlocutor: $M = 61.2\%$, $SD = 16.4$; based on correct responses included in the frequency analysis).

Across tasks, participants ranged in their switching frequency (range 6%–49% of trials being a switch, although most participants on average switched between 20% and 40% of trials). Similarly,

items ranged in their switching frequency (19%–55% switches in the no-context task) and language choice (24%–79% named in English across participants). Across participants, all items were used to switch languages and all items were sometimes named in English and sometimes in Bulgarian.

Q1. Switching Frequency: Context-Sentence Versus No-Context and Context-Sentence Versus Context-Interlocutor

We first examined whether switching frequency differed when naming pictures in the no-context versus in the context-sentence task. This model converged with all intercepts and slopes apart from the by-item slope for language. There was a significant effect of task on switching frequency ($\beta = -.265$, $SE = 0.101$, $z = -2.629$, $p = .009$), with bilinguals switching less often in the context-sentence task ($M = 29.9\%$, $SD = 13.6$) than in the no-context task ($M = 34.1\%$, $SD = 11.0$, see Figure 2). Language did not relate to trial type ($\beta = -.097$, $SE = 0.061$, $z = -1.574$, $p = .115$), reflecting that switching frequency was comparable for Bulgarian and English.

Next, we examined the influence of interacting with a conversation partner versus not having an interlocutor on switching frequency. This model too converged after removing the by-item slope for language. The switching frequency did not differ significantly between the context-interlocutor task ($M = 28.0\%$, $SD = 12.7$) and the context-sentence task ($M = 29.9\%$, $SD = 13.6$; $\beta = .102$, $SE = 0.085$, $z = 1.193$, $p = .233$, see Figure 2). Switching frequency was higher in English ($M = 29.9\%$, $SD = 12.5$) than in Bulgarian ($M = 23.6\%$, $SD = 12.2$) in these tasks ($\beta = .353$, $SE = 0.083$, $z = 4.231$, $p < .001$). The context-sentence task included both predictable and unpredictable sentence contexts while the context-interlocutor task only included unpredictable sentence contexts. Given that switching frequency differed between predictable and unpredictable contexts (see next sets of analyses), as preregistered, we also ran the comparison between context-sentence and context-interlocutor including just the unpredictable sentences. In this analysis, there was a significant difference in switching frequency between tasks ($\beta = .219$, $SE = 0.084$, $z = 2.619$, $p = .009$), with a higher switching frequency in the context-sentence task than in the context-interlocutor task.

In terms of correlations across tasks, switching frequency was significantly related between the no-context and context-sentence tasks, $r(58) = 0.412$, $p < .001$, and between the context-sentence and context-interlocutor tasks, $r(58) = 0.556$, $p < .001$, see Figure 3.

Q1. Switching Costs: Context-Sentence Versus No-Context and Context-Sentence Versus Context-Interlocutor

We first compared RTs in the no-context and context-sentence tasks, where the maximal model converged after removal of correlations. There was no main effect of task ($\beta = -.023$, $SE = 0.015$, $t = -1.506$, $p = .137$). Overall RTs were comparable for the no-context ($M = 1{,}200$, $SD = 173$) and context-sentence tasks ($M = 1{,}162$, $SD = 197$). The main effect of language ($\beta = -.028$, $SE = 0.011$, $t = -2.613$, $p = .011$) reflected faster naming in English (across no-context and sentence-context tasks: $M = 1{,}166$, $SD = 162$) than Bulgarian ($M = 1{,}211$, $SD = 180$). Of main interest, there was a main effect of trial type ($\beta = .107$, $SE = 0.009$, $t = 12.376$, $p < .001$), reflecting a significant switching cost (see Table 2 and Figure 4). This cost interacted with Task ($\beta = .031$, $SE = 0.014$, $t = 2.242$, $p = .029$). The switching cost was larger in the context-sentence task (trial type computed relative to reading the sentence) than when naming without context (trial type computed relative to naming on the previous trial).

The comparison between context-sentence and context-interlocutor showed somewhat faster naming in the context-interlocutor task ($\beta = .035$, $SE = 0.017$, $t = 2.050$, $p = .044$; without interlocutor: $M = 1{,}162$, $SD = 197$; with interlocutor $M = 1{,}121$, $SD = 189$). The main effect of language ($\beta = -.053$, $SE = 0.013$, $t = -3.944$, $p < .001$) again reflected faster naming in English than Bulgarian. There was also a main effect of trial type ($\beta = .125$, $SE = 0.010$, $t = 12.889$, $p < .001$), reflecting that there was a significant switching cost (see Table 2 and Figure 4). This cost did not interact with Task ($\beta = .002$, $SE = 0.012$, $t = 0.212$, $p = .833$). The switching cost was comparable in the context-sentence and context-interlocutor tasks. This remained the case when only unpredictable sentences were included.

In terms of correlations across tasks, switching costs showed a significant correlation between the no-context and context-sentence tasks, $r(57) = 0.401$, $p = .002$, and between the context-sentence and context-interlocutor tasks, $r(57) = 0.317$, $p = .015$, see Figure 5. This analysis excluded the participants without switch trials in the no-context task.[1]

Finally, we assessed the internal consistency of RT switching costs using a permutation-based split-half approach (Parsons, 2021) with 5,000 random splits. The (Spearman–Brown corrected) internal consistency of the switching cost was 0.68, 95% confidence interval, CI [0.53–0.79] in the no-context task, 0.72, 95% CI [0.52–0.85] in the context-sentence task, and 0.65, 95% CI [0.47–0.79] in the context-interlocutor task.

Q2. Sentence Predictability (Context-Sentence Task)

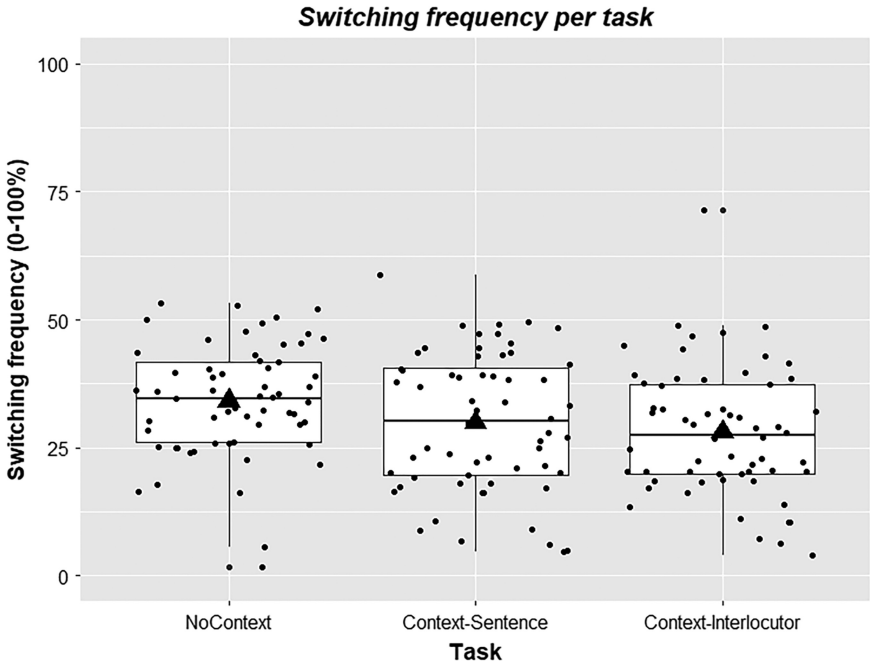
Within the context-sentence task, we first assessed the influence of sentence predictability on switching frequency. This model converged with participant and item intercepts and the participant slope for language. Switching frequency was higher in the unpredictable condition ($M = 32.0\%$, $SD = 13.1$) than in the predictable condition ($M = 27.8\%$, $SD = 14.8$; $\beta = .215$, $SE = 0.055$, $z = 3.928$, $p < .001$, see Figure 6). Similar to the previous analysis, there were more English than Bulgarian switches ($\beta = .184$, $SE = 0.086$, $z = 2.129$, $p = .033$).

The RT analysis converged with the intercepts and participants' slopes for Predictability, Language, and Trial type. RTs showed a significant main effect on predictability ($\beta = .069$, $SE = 0.018$, $t = 3.918$, $p < .001$). As expected, participants responded faster in the predictable ($M = 1{,}122$, $SD = 201$) than in the unpredictable condition ($M = 1{,}203$, $SD = 198$). Responses were faster in English than in Bulgarian ($\beta = -.034$, $SE = 0.014$, $t = -2.467$,

[1] After removal of incorrect responses and RT outliers, six participants produced <10 switch trials in at least one of the tasks. Removing these participants from the analysis still showed a larger switching cost in the sentence than no-context task and similar costs for the sentence tasks with and without an interlocutor. Furthermore, the correlations between switching costs remained significant, suggesting these were not caused by these low-switching participants.

Figure 2

Boxplots Showing the Switching Frequency (Percentage of Correct Trials Being a Language Switch) for the No-Context, Context-Sentence, and Context-Interlocutor Tasks



Note. The boxplot shows the interquartile range with the black dots representing individual participant scores. The median is indicated by the horizontal black line and the centers of the black triangles show the means.

$p = .017$). There was a significant switching cost ($\beta = .123$, $SE = 0.012$, $t = 10.564$, $p < .001$), which interacted with predictability ($\beta = -.024$, $SE = 0.011$, $t = -2.130$, $p = .033$). Switching costs were somewhat smaller in the unpredictable ($M = 159$, $SD = 163$) than in the predictable condition ($M = 166$, $SD = 159$; means excluding one participant with no switch trials left in the predictable condition, see Figure 6).

Q3. Switching and Naming Behaviour Interlocutor (Context-Interlocutor Task)

Finally, in the context-interlocutor task, we assessed the influence of the interlocutor on switching frequency and costs. Starting with switching frequency, the model converged after removing the by-item slope for language. Switching frequency was somewhat higher in the condition where the interlocutor switched frequently ($M = 30.5\%$, $SD = 12.7$) than in the low-switching condition ($M = 25.3\%$, $SD = 12.3$, see Figure 7) but this did not reach significance ($\beta = .334$, $SE = 0.177$, $z = 1.883$, $p = .060$). Similar to the previous analysis, there were more English than Bulgarian switches ($\beta = .530$, $SE = 0.106$, $z = 5.003$, $p < .001$).

In terms of RTs (full model converged), there was no significant difference between the high- ($M = 1{,}113$, $SD = 204$) and low-switch conditions ($M = 1{,}128$, $SD = 174$; $\beta = -.014$, $SE = 0.043$, $t = -0.337$, $p = .737$). Again, there were significant effects of trial type ($\beta = .128$, $SE = 0.011$, $t = 11.726$, $p < .001$) and language ($\beta = -.071$, $SE = 0.015$, $t = -4.885$, $p < .001$). There was no interaction between trial type and switching condition ($\beta = -.014$, $SE = 0.022$, $t = -0.638$, $p = .526$), reflecting similar switching costs in the high- ($M = 142$, $SD = 110$) and low-switching ($M = 189$, $SD = 139$) conditions (see Figure 7).

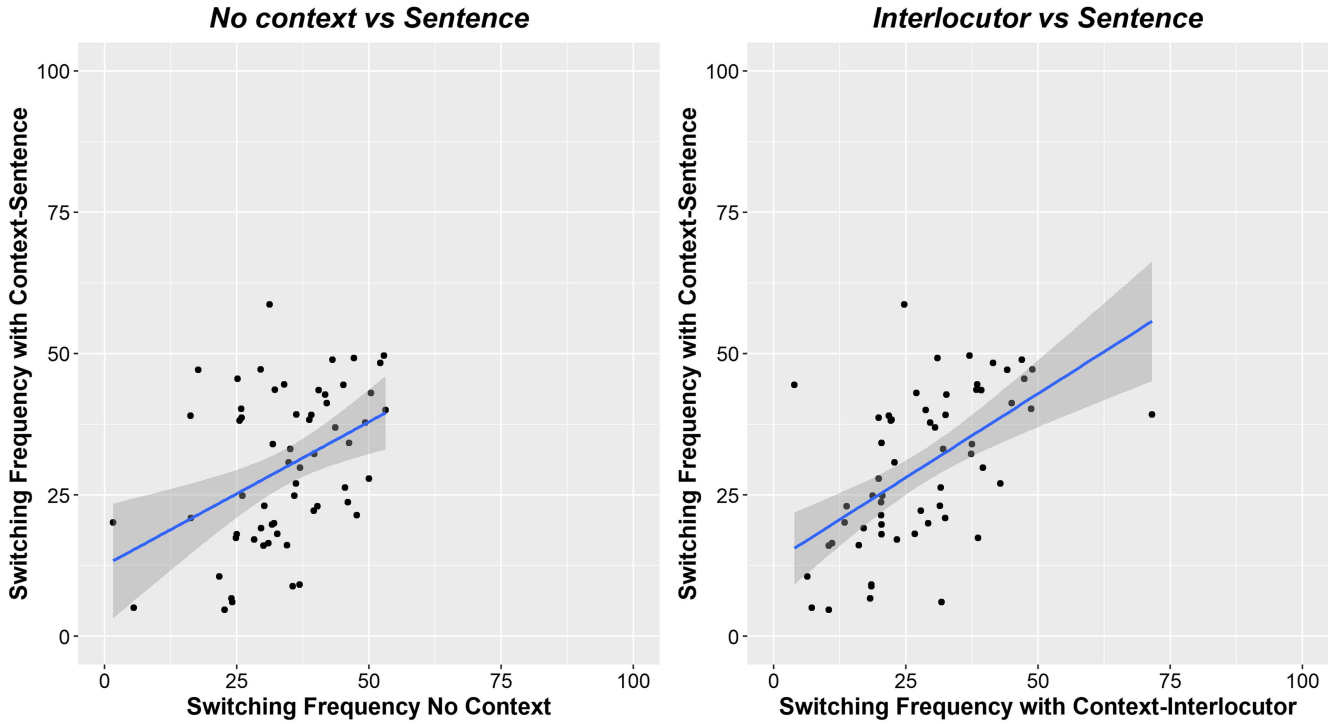
Exploratory Analyses With Interlocutor

Our main focus was on the global effect of an interlocutor on a participant's overall switching behavior. However, previous research has also shown more short-lived, local effects (e.g., Fricke & Kootstra, 2016; Kootstra et al., 2020), with the interlocutor's immediately preceding sentence influencing a bilingual's behavior. We therefore conducted two exploratory analyses.

The first examined the influence of the interlocutor's preceding sentence on the bilingual's own switching frequency. We examined whether bilinguals were more likely to switch if the interlocutor had just switched in their preceding utterance. This model converged with participant and item intercepts and the participant slope for language. Participants were more likely to switch when the interlocutor had just switched in the preceding sentence ($\beta = .242$, $SE = 0.067$, $t = 3.602$, $p < .001$, see Figure 8). Furthermore, the effect of condition (overall high- or low-switching interlocutor) was no longer near significance ($\beta = .173$, $SE = 0.175$, $t = 0.989$, $p = .323$) but did interact with the preceding utterance being a switch or not ($\beta = -.292$, $SE = 0.134$, $t = -2.171$, $p = .030$). When the interlocutor's preceding utterance contained a switch, switching frequency was not influenced by the interlocutor's overall behaviour at all ($\beta = -.019$, $SE = 0.205$, $t = -0.091$, $p = .928$). In these cases, participants switched frequently in both the overall high-switching

Figure 3

Scatterplots Showing the Correlation Between the Switching Frequency in the No-Context and Context-Sentence Tasks (Left) and the Context-Sentence and Context-Interlocutor Tasks (Right)



Note. See the online article for the color version of this figure.

($M = 31.3\%$, $SD = 12.6$) and the low-switching condition ($M = 32.1\%$, $SD = 17.3$). In contrast, when the preceding utterance did not contain any switches, participants in the overall high-switching condition were somewhat more likely to switch languages themselves ($M = 29.7\%$, $SD = 13.6$) than in the low-switching condition ($M = 24.6\%$, $SD = 12.3$), although this difference was not significant ($\beta = .288$, $SE = 0.198$, $t = 1.456$, $p = .145$). These findings were related to the interlocutor's preceding utterance and not influenced by the participant's own language switching on their previous trial. Additional analyses examining RTs instead of switching frequency furthermore showed no influence of the interlocutor's preceding sentence containing a language switch or not.

The second exploratory analysis examined whether the interlocutor's language use influenced the bilingual's language choice. In both conditions, the interlocutor used English half of the time and Bulgarian the other half. However, the language they used for each given target varied across participants. We therefore analyzed whether the language the interlocutor used for a specific word influenced the bilingual's own language choice for that word. We also considered the role of the language used by the interlocutor on the preceding trial (which language they used to name the previous target, which was different from the participant's target word). Bilinguals were indeed more likely to use the language just used by the interlocutor to name the preceding target ($\beta = .339$, $SE = 0.130$, $t = 2.611$, $p = .009$). However, language choice was more strongly influenced by the language previously used by the interlocutor to name their specific target item. Participants were more likely to use Bulgarian for an item that the interlocutor had previously named in Bulgarian (and English if the interlocutor had named that item in English; $\beta = 1.275$, $SE = 0.197$, $t = 6.476$, $p < .001$, see Figure 9). Mean English language choice was 73.2% ($SD = 17.3$) when the interlocutor had named that item in English and 50.7% ($SD = 22.4$) when it had been named in Bulgarian (relative to mean English use being 59.9% for items that had not been named yet by the interlocutor). Additional analyses examining RTs instead of language choice showed no influence of the interlocutor's preceding language choice.

Table 2

RT Means (and SDs) for the Three Naming Tasks per Trial Type (Nonswitch, Switch) and the Switching Costs (RT Difference Between Switch and Nonswitch Trials)

Trial type	No-context	Context-sentence	Context-interlocutor
Nonswitch	1,162 (174)	1,124 (199)	1,084 (189)
Switch	1,288 (192)	1,295 (225)	1,249 (227)
Switching cost	129 (110)[a]	171 (158)	165 (126)

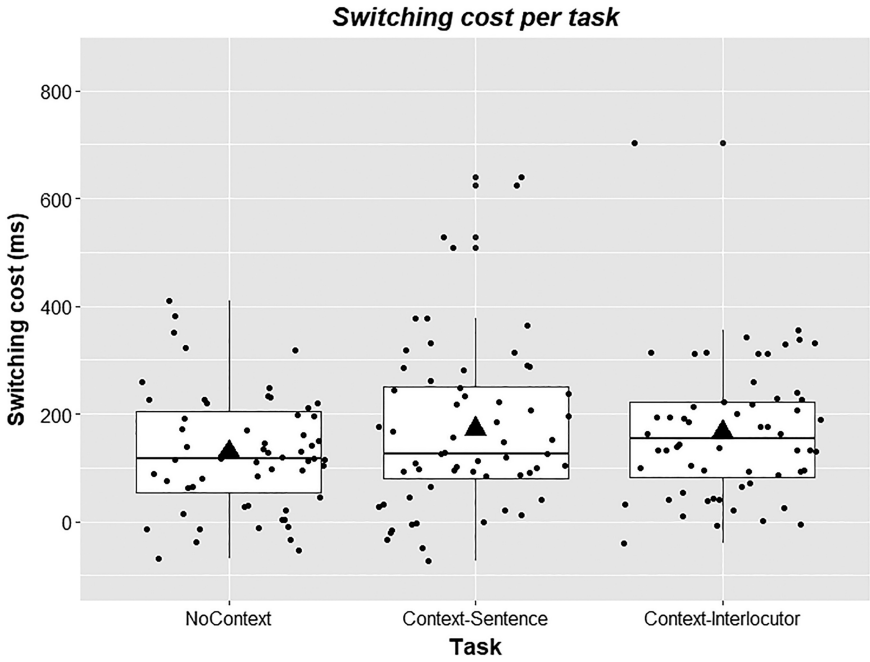
Note. RT = reaction times.
[a] One participant is not included in this switching cost as no switch trials remained after outlier removal.

Discussion

We examined how Bulgarian-English bilinguals switch languages voluntarily when naming pictures without context, in a sentence

Figure 4

Boxplots Showing the Switching Cost (RT Difference Between Switch and Nonswitch Trials) for the No-Context, Context-Sentence, and Context-Interlocutor Tasks



Note. The boxplot shows the interquartile range with the black dots representing individual participant scores. Note that the one participant with no remaining switch trials in the no-context task is not included. The median is indicated by the horizontal black line and the centers of the black triangles show the means. RT = reaction times.

context, and in a sentence context with an interlocutor. We focused on their switching frequency (how often they switched languages) and switching cost (RT difference between switch and nonswitch trials). Bilinguals switched in all contexts, which was associated with a cost in all contexts too. They furthermore switched less often and showed larger costs in the sentence context than when naming without context. Both frequency and costs correlated between the context-related comparisons of interest. Comparisons between high- and low-predictability sentences showed that bilinguals switched more often and showed somewhat smaller switching costs when sentences did not predict a specific target word. Finally, we assessed the influence of the interlocutor's switching frequency on the bilingual's own switching frequency. The global influence of the interlocutor's frequency on the bilingual's frequency was small and not significant. However, further exploratory analyses showed that the interlocutor's behavior did locally influence the bilingual's moment of switching (more frequent switching immediately after the interlocutor switched) and language choice (more frequent use of the language that the interlocutor also used for that item).

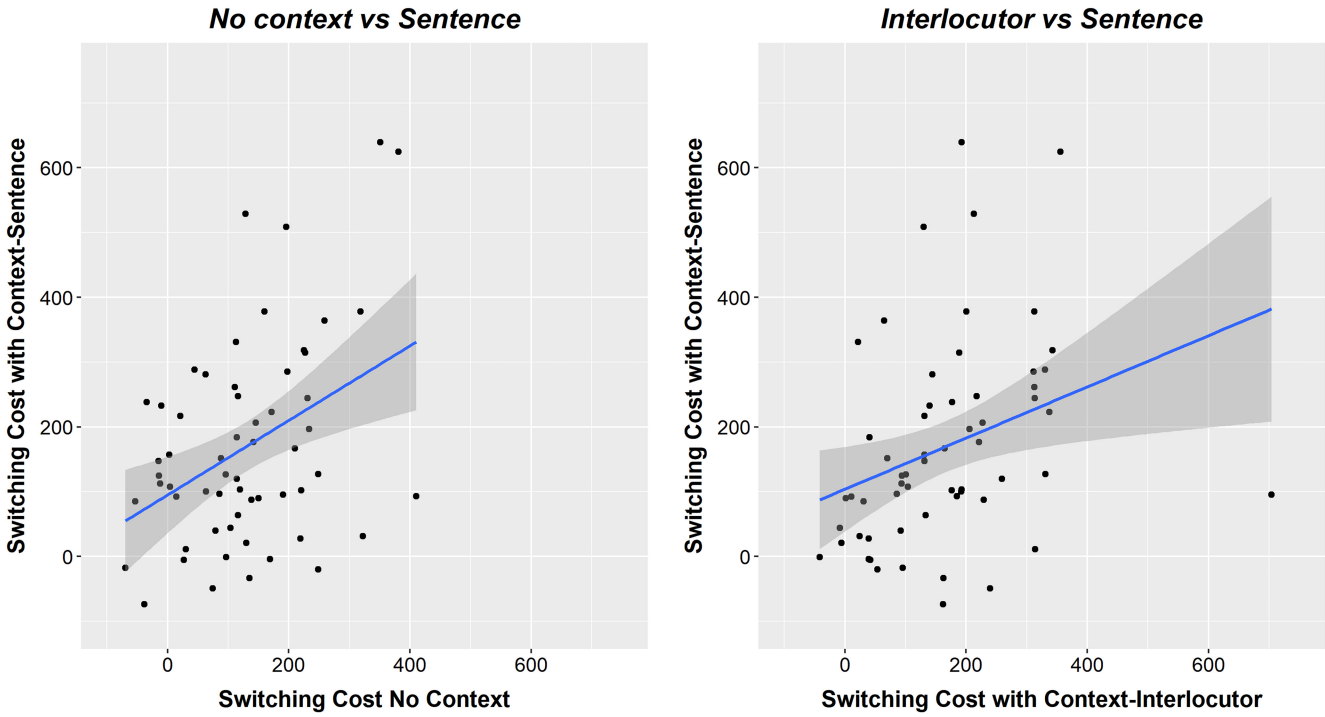
Comparison of Voluntary Switching Across Tasks

Research examining the cognitive mechanisms involved in voluntary switching has typically focused on the production of individual words without any context. Often, these studies show substantial switching costs (e.g., de Bruin et al., 2018; Gollan & Ferreira, 2009). However, previous research using other types of tasks has suggested that switching costs might not emerge in sentence context (e.g., Gullifer et al., 2013). Our first aim of the comparison of switching frequency and costs across tasks was to identify whether switching costs are potentially an artifact of less naturalistic tasks that require switching between individual words without context. That was not the case. Switching costs were observed in all three tasks, both with and without context. This is in line with recent studies using different measures (e.g., number of filler words) also showing switching costs in a sentence context (Sánchez et al., 2022), in conversations (Faroqi-Shah & Wereley, 2022), and in corpus data (Fricke et al., 2016). Given that we used naming times as a measure, however, we were also able to compare naming in context versus in isolation (as typically used in this literature) and showed through a direct comparison that participants switched regularly and showed switching costs both with and without context.

Our second aim was to compare how switching frequency and costs relate to each other across tasks with and without context. In our comparisons of interest (no-context vs. sentence-context and without interlocutor vs. with interlocutor), both frequency and switching costs correlated moderately across contexts. It is important to point out here that for the key comparison (no-context vs. sentence-context), switching frequency and costs were computed in different ways. In the no-context task, switches were identified relative to the participant's previously named picture (e.g., a trial was a switch if the previous picture had been named in Bulgarian and the current picture was named in English). This is how switching is typically defined in research using picture naming in isolation. In contrast, in our context tasks, trial type was determined relative to the language of the preceding words in the sentence, in line with how code switches are typically identified

Figure 5

Scatterplots Showing the Correlation Between the Switching Cost in the No-Context and Context-Sentence Tasks (Left) and the Context-Sentence and Context-Interlocutor Tasks (Right)



Note. Note that the participant without remaining switch trials in the no-context task is not included in the scatterplot. See the online article for the color version of this figure.

in daily-life speech. That is, a trial type was a switch if, for example, the sentence started in Bulgarian and the target word was produced in English. This was regardless of the language the participant used to name the previous target picture. Thus, despite trial type being defined in different ways, the switching costs and frequency correlated across tasks. Furthermore, internal reliability of the switching costs was comparable for the three tasks. Reliability was adequate, although a little lower than reliability observed in previous data sets of voluntary switching (e.g., de Bruin et al., 2018). It is possible that split-half reliability was lower in this study as a consequence of pictures not being repeated within the same task. Individual items can have a large impact on language choice, and potentially also the moment of switching (e.g., de Bruin & Martin, 2022; de Bruin et al., 2018). Split-half reliability with unrepeated items might therefore be more susceptible to individual item effects than analyses with repeated items.

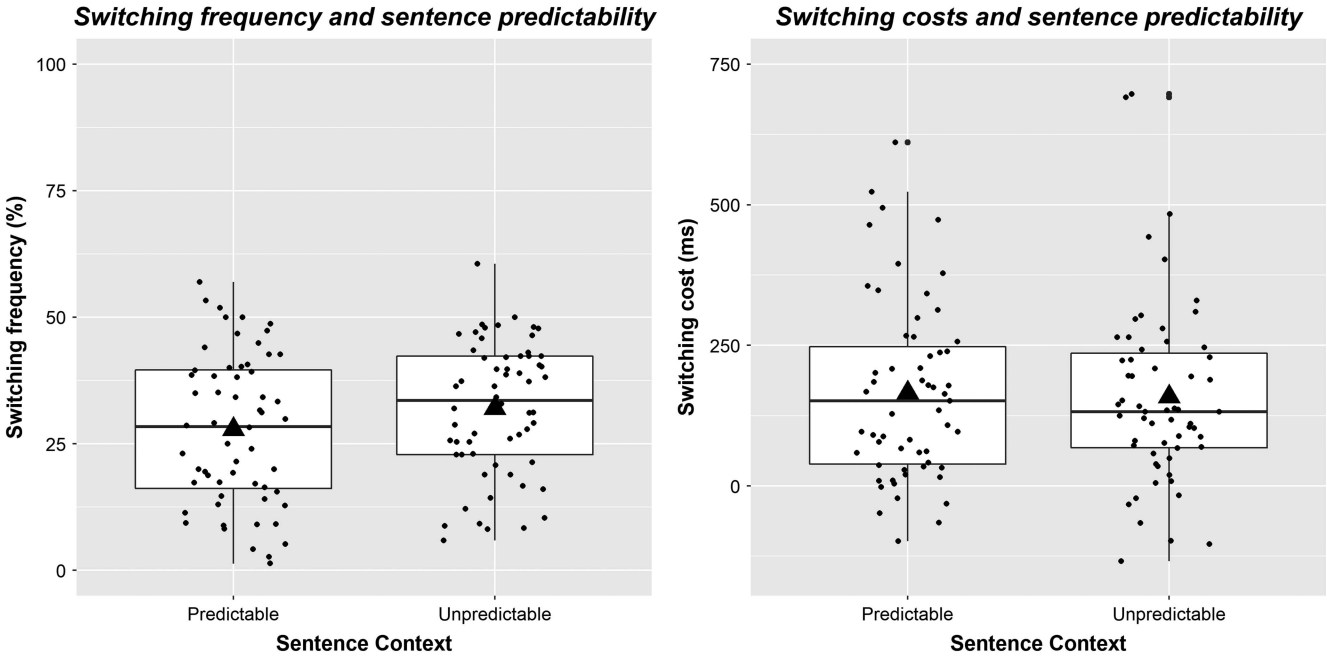
Switching costs being present in all contexts, and correlating despite the differences between tasks, suggest that bilinguals do indeed experience switching costs even when switching freely. This suggests that bilinguals experience competition between languages even when using their languages freely. Although language choice in these contexts might largely be driven by lexical access and how quickly an item can be retrieved (e.g., de Bruin et al., 2018), the other language might remain active and might require ongoing language coordination (D. W. Green & Wei, 2014) and potential language control to manage this competition (e.g., D. W. Green, 1998). The switching cost patterns across tasks suggest that these costs and potential control mechanisms might be present across contexts. Although contextual features can modulate both frequency and costs, the similarities across tasks suggest language-switching mechanisms are not fundamentally different with and without context.

However, we did also observe differences between tasks. With respect to overall differences between tasks with and without context, we hypothesized that these effects could go in two directions. On the one hand, sentence contexts might allow for more preparation time (e.g., Declerck & Philipp, 2015) and provide a more naturalistic environment facilitating language switches relative to naming in isolation. On the other hand, language coactivation might be reduced in sentence contexts (e.g., Van Hell & de Groot, 2008) and might therefore reduce language switching frequency in context relative to no-context. Our data showed that bilinguals switched more frequently without context, supporting the second hypothesis that sentence context might reduce language coactivation and therefore reduce the likelihood of switching. This hypothesis is discussed further below, when comparing predictable and unpredictable switching contexts. The overall comparison of contexts with and without an interlocutor showed no difference in switching frequency, although there was a difference when only unpredictable sentences were considered. This is likely because of the presence of the interlocutor as well as effects of sentence predictability, as discussed below.

In terms of switching costs, these were larger in the tasks measuring switching in sentence contexts. This again suggests that sentence context could have reduced language coactivation and therefore

Figure 6

Boxplots Showing the Switching Frequency (Left) and Switching Costs (Right) for the Predictable and Unpredictable Conditions in the Context-Sentence Task



Note. The boxplot shows the interquartile range with the black dots representing individual participant scores. The median is indicated by the horizontal black line and the centers of the black triangles show the means.

decreased the ease of switching relative to no-context. It also provides further evidence that switching costs are not absent in sentence contexts and that the additional preparation time sentences provide does not eliminate switching costs. This might especially apply to voluntary language switching, which is less susceptible to effects of preparation time than cued switching (de Bruin & Xu, 2023).

However, although the tasks were comparable in various ways (e.g., DVs and stimuli used), there were several differences between the tasks that could have influenced the cross-task comparison in terms of the size of the switching cost and frequency. The direct comparisons in terms of size should therefore be interpreted cautiously. First, switching was measured in different ways in sentence context tasks (relative to the sentence) versus the no-context task (relative to the previous target word). Second, in the sentence context task, in most sentences, participants also always had to use the target noun after using a verb while in the no-context task, only nouns were used. Third, in the sentence context task, participants alternated between reading a sentence and naming pictures, introducing a modality switch. These latter two task differences should apply to both nonswitch and switch trials. However, there was no overall task RT difference between the sentence context and no-context tasks. If anything, naming was slightly faster in the sentence context task, suggesting that the differences between tasks did not make the sentence task more difficult. Thus, going from reading a sentence out loud to producing a picture name does not appear to have introduced a general processing cost (across nonswitch and switch trials) relative to picture naming without context. Indeed, when the aim of the comprehension and production tasks is the same (producing words out loud) and when the comprehension task does not include additional demands (e.g., semantic classification), going from reading out loud to production of picture names might not introduce costs (cf. Li & Gollan, 2022). Furthermore, any modality switch costs might also have been reduced or eliminated in our study thanks to the short interval (500 ms) between the end of sentence reading and picture presentation. In the absence of an impact on overall RTs, it is therefore unlikely that a modality switch can explain the differences in language switching costs across tasks. However, this cannot be ruled out and the direct comparisons in terms of the size of costs should be interpreted with caution. The same applies to the comparison between contexts with and without an interlocutor, which differed in their stimulus list composition (i.e., the presence of predictable sentence contexts). Stimulus list composition in general is a variable that can influence bilingual processing (e.g., Dijkstra et al., 1998). Although these differences between the three task versions hinder the interpretation of the absolute size of the switching costs, they also make the similarities in terms of correlations and reliability across tasks more striking.

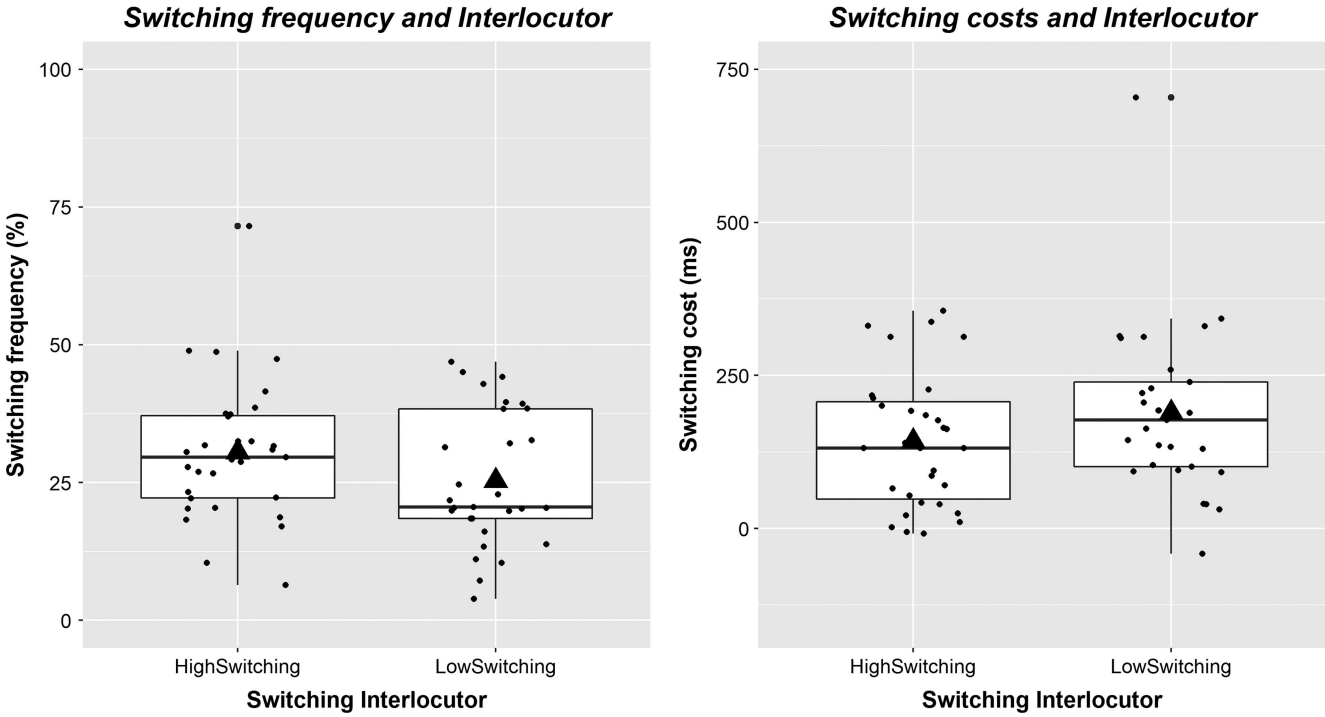
Considering these differences across tasks, further within-task comparisons provide a better opportunity to examine the impact of specific contextual factors on switching frequency and costs. Next, we therefore discuss the role of sentence predictability and the influence of the interlocutor.

Sentence Predictability

As expected, naming was faster overall in predictable than in unpredictable sentence contexts. When a word is predicted by the preceding language context, coactivation of related words in the other language might also be reduced (e.g., Lauro & Schwartz, 2017). Given that language choice and switching are related to how easily a word can be

Figure 7

Boxplots Showing the Switching Frequency (Left) and Switching Costs (Right) for the High- and Low-Switching Interlocutor Conditions in the Context-Interlocutor Task



Note. The boxplot shows the interquartile range with the black dots representing individual participant scores. The median is indicated by the horizontal black line and the centers of the black triangles show the means.

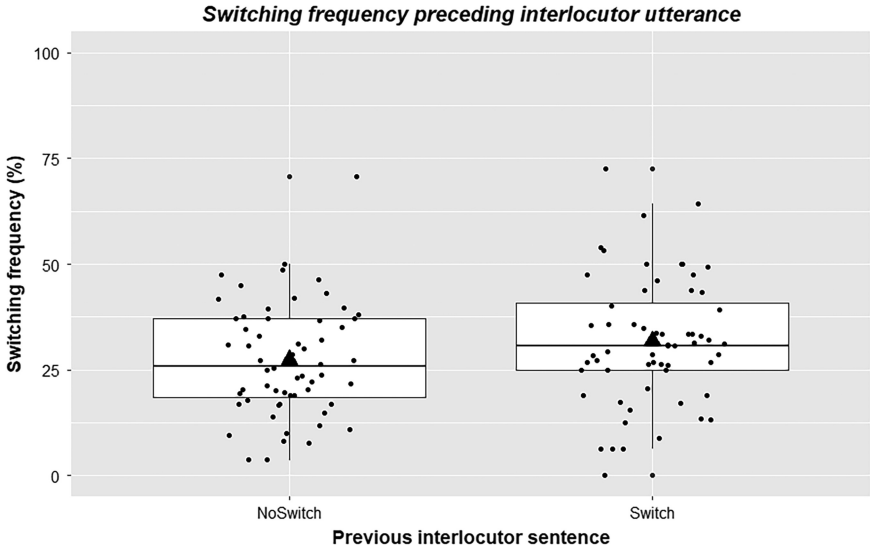
accessed in each language (e.g., de Bruin et al., 2018), we hypothesized language switching to be less frequent when a word is predictable in a given sentence context. This was indeed the case: bilinguals switched more often on a target word that was not predictable. This also aligns with corpus research (Myslín & Levy, 2015) showing that bilinguals switch languages more often when unpredictable information is coming up. These differences in switching behavior can be explained through both speaker-based and listener-based factors. On the one hand, language behavior (including comprehension processes too) might mostly be shaped by mechanisms related to the speaker and their language production (e.g., MacDonald, 2013). This can include how easily certain words or structures are retrieved or planned, with a preference for easy-to-retrieve choices that increase production fluency. In the case of unpredictable sentence contexts, this can include speakers being more likely to switch because coactivation of the "other" language is higher than in predictable contexts. On the other hand, switching behaviour might also be more listener-oriented, with a speaker switching languages to "warn" the listener that the upcoming information is less predictable or expected (e.g., Myslín & Levy, 2015). Although no listener was present in this task, a bilingual's ongoing real-life experience of adjusting their language behavior in this way (i.e., switching more often when information is not predictable) could still have shaped their language behavior more generally, including in contexts without direct interaction.

Our findings regarding switching frequency are thus compatible with both accounts. However, compared to the no-context task, predictable sentences appeared to reduce switching frequency (rather than unpredictable sentences increasing switching), suggesting this pattern is more easily explained through coactivation of translation equivalents in the other language being reduced in predictable contexts. Furthermore, the finding that switching costs were also somewhat larger in predictable sentence contexts (although the difference with unpredictable contexts was very small), is more easily explained through language coactivation being reduced in predictable contexts. Switching to the other language (e.g., because the bilingual prefers the language for that specific target item or because they prefer a specific language more globally) might take more time if the coactivation of that other language is lower, and activation and interference from the language currently in use is higher, in a predictable sentence context. This switching cost pattern seems less compatible with alternative accounts that explain the costs through language switching being used to manage or recover in situations where speech planning demands are high. While the switching frequency pattern is in line with this account too, the switching costs being smallest in the more "difficult to plan" unpredictable condition suggests these language switching costs did not arise purely as a consequence of participants switching mostly in utterances that were most difficult to plan.

Furthermore, larger switching costs in predictable contexts do not necessarily align with listener-focused "audience design" playing a large role, with bilinguals adjusting their switching behavior to accommodate a listener (Myslín & Levy, 2015). In this specific task, however, participants were not interacting with another bilingual. It is possible that predictability effects on switching costs differ

Figure 8

Boxplots Showing the Switching Frequency in the Context-Interlocutor Task for Sentences Immediately Preceded by an Interlocutor Utterance Without a Switch (Left) Versus Interlocutor Sentences With a Switch (Right)



Note. The boxplot shows the interquartile range with the black dots representing individual participant scores. The median is indicated by the horizontal black line and the centers of the black triangles show the means.

from the current findings when sentence predictability is manipulated in interaction with another bilingual. That is, listener-focused factors might play a larger role in interaction. Furthermore, bilinguals might switch for communicative purposes (and thus listener-based factors might be more influential) in particular when words are unexpected, while the current study only included low-constraint sentences in which the target word was simply not predicted but not very unexpected.

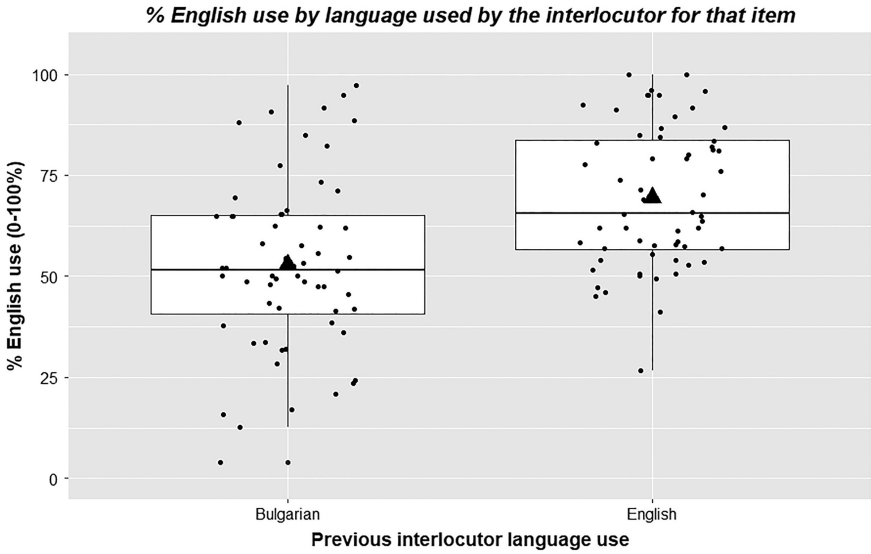
Finally, when smaller costs and higher frequency are found in the same condition, it is possible that purely the practice of switching more often in a task can explain the smaller costs. We therefore assessed whether the predictability effect in frequency correlated with the effect in costs. While there was a small correlation, this was not significant, $r(57) = -.229$, $p = .081$, suggesting that pure "practice" with switching in a task did not explain the predictability effect on switching costs. However, it should be pointed out that the switching cost effect of predictability was rather small. The main impact of sentence context appeared to be the actual switching behaviour itself, namely how frequently a bilingual switches. It is likely that any impact on naming times (switching costs) is minimal as bilinguals were switching voluntarily. Lexical access in the "other language" might be slower in predictable sentences, but if a bilingual can name that specific item much faster in the "other language," this is likely to reduce any negative effects of predictability on switching costs.

Interacting With an Interlocutor

Our main question in the final context (naming in interaction with an interlocutor) was to examine whether the interlocutor's global switching behavior influenced the bilingual's overall switching behavior. Previous studies have focused on more short-lived effects by comparing the influence of the interlocutor's preceding utterance (e.g., Kootstra et al., 2020). Here we wanted to examine whether bilinguals adjust their language behavior more globally, by comparing interactions with a high-switching or low-switching interlocutor. Language use was kept constant between conditions, with the interlocutor using English half of the time and Bulgarian the other half. There was a small effect in the expected direction, although this did not reach significance. Bilinguals switched somewhat more often overall when the interlocutor switched frequently too. This type of alignment could occur as the consequence of multiple mechanisms. High-switching contexts could increase the overall coactivation of both languages and enhance the likelihood of switching (e.g., Kootstra et al., 2020). This interpretation centers more around speaker-based benefits. However, alignment could also reflect less automatic processes focused more on the listener, for instance through social alignment, where bilinguals create a shared identity with the interlocutor by using similar language behavior. Finally, language switching is not always possible, or even viewed positively (cf. Dewaele & Wei, 2014). A low-switching interlocutor could have created an environment where participants felt that switching was perhaps not appropriate or viewed positively. Indeed, relative to the other tasks (e.g., the unpredictable Context-Sentence sentences), participants' switching behavior decreased in the presence of a low-switching interlocutor as opposed to increasing with a high-switching interlocutor. This suggests that participants perhaps evaluated this context as not being a switching environment and adjusted their behavior accordingly. However, these results should be interpreted with caution given that the effect did not reach significance. It is possible this effect is larger when interactions take place in person, although previous research has shown strong alignment even

Figure 9

Boxplots Showing Percentage English Use in the Context-Interlocutor Task for Items Previously Named by the Interlocutor in Bulgarian (Left) Versus in English (Right)



Note. The boxplot shows the interquartile range with the black dots representing individual participant scores. The median is indicated by the horizontal black line and the centers of the black triangles show the means.

when humans are interacting with computers (cf. Branigan et al., 2010). Due to COVID-19, we used an online paradigm in which participants were only introduced to the interlocutor via a recorded video in the beginning. The rest of the task used audio recordings. In-person interactions, or interactions where the interlocutor's face is visible, might increase the interlocutor's influence. Furthermore, it should be noted that this manipulation was the only between-subject manipulation in the study and power therefore was lower than for the other analyses.

While global alignment to the interlocutor had relatively little impact on the bilingual's switching, bilinguals did show stronger local alignment to the interlocutor's preceding utterance. In line with previous literature (e.g., Kootstra et al., 2020), bilinguals were more likely to switch languages if the interlocutor just switched themselves. This also interacted with the interlocutor's overall switching behaviour, which had some influence on switching frequency when the preceding utterance contained no switch, but no influence when there was a switch in the preceding utterance. Together, these findings suggest that the immediate behaviour of the interlocutor (i.e., the preceding utterance) has a larger influence on a bilingual's own behavior than their overall switching patterns. Interestingly, these patterns were observed without a "lexical boost" driven by the repetition of items. This suggests language switching alignment does not require alignment or repetition of lexical items (cf. Kootstra et al., 2020). The observed local alignment can be explained through both more automatic (speaker-based) priming mechanisms and through more mediated (potentially more listener-focused) alignment, for example for social alignment or for audience design (e.g., to help the listener process your words), with the different processes not being mutually exclusive.

Furthermore, language choice for each item was also influenced by the interlocutor, despite their overall language use being equal for each language. Language choice was influenced by the interlocutor's language in the previous trial, but even more so by the language the interlocutor used to name the same item himself previously. This is despite the often relatively large temporal distance between the interlocutor naming that item and the participant naming it. These patterns are most likely to arise through automatic priming, with the interlocutor's language use increasing the activation of that specific word in the language they used for it. Given the large number of items and the distance between interlocutor and participant naming, it is unlikely that these effects can be explained through more mediated (e.g., social) alignment, which would require the bilingual to keep track of the interlocutor's language choice for each item.

Effects of context mostly concerned the actual moment or frequency of switching, with little to no impact on the naming times or costs associated with switching. Across the different findings, this suggests effects of context do not place an additional cost on the bilingual's language processing and are likely to emerge through more automatic mechanisms such as language coactivation and priming. These findings are in line with previous research assessing effects of visual context on voluntary picture naming, which showed language-choice alignment with visual cues (flags associated with one of the languages) as well language input (de Bruin & Martin, 2022). Similar to the current study, alignment with context did not have a negative effect on naming times or switching costs either. This suggests that adjusting your language behavior to the context (in this case the interlocutor) does not increase the effort or control associated with bilingual language production. This furthermore suggests that alignment in language choice and switching might mostly take place in the form of more automatic, bottom-up mechanisms such as priming that benefit the speaker, and perhaps less through more top-down mechanisms focusing on the listener.

Context: Limitations and Future Directions

There are many contextual variables that might play a role in daily-life language production. The current study aimed to examine a limited number of specific contextual variables. However, there are many other factors that should be considered in the future. Below we summarize a few key variables that require further research. Ideally, these variables would be studied through a combination of corpus data capturing real-life conversations and experimental research that is able to study the role of specific variables in a more controlled environment.

With respect to the role of the interlocutor, our (recorded) interlocutor was another Bulgarian-English bilingual but was not known to the participants. Bilinguals' switching behavior in real-life conversations might vary depending on the identity of the interlocutor, with bilinguals found to switch more frequently with another bilingual from the same group than with people from another group (e.g., Poplack, 1983). Thus the identity of the interlocutor, and whether they are known by the participant, could influence switching behavior. Furthermore, there might be additional processing costs associated with interacting with multiple interlocutors (Peeters, 2020), which is also common in some real-life settings, and could further influence switching.

In the current study, we aimed to keep relatively high experimental control to enable us to reliably measure naming onset times. However, the experimental setup does come with several limitations. In addition to the limitations already discussed in previous paragraphs, the stimuli used embedded words in a sentence context but did not embed those sentences into a broader discourse. In daily-life conversations, sentences do not stand on their own (as was the case in the current study) but rather connect to the other sentences within a larger conversation or narrative.

Furthermore, we focused on one specific type of intrasentential switching while daily-life language switching can (depending on the bilingual) include different types of switches. This can include language switches between and within sentences. Within sentences, switches can come in the form of insertions of one word in different places of a sentence, alternating languages across multiple words, and so-called dense code-switching or congruent lexicalization using a shared language structure with words and/or morphemes from multiple languages. In the current study, we only worked with language switches that were the insertion of one noun phrase at the end of the sentence. Given that even within-sentence code switches might be processed differently depending on the type (e.g., D. W. Green & Wei, 2014), future research also needs to consider the role of different switching patterns.

In this study, we furthermore did not examine potential differences between the two languages, apart from including language as a main effect in the analyses. Some differences between languages were observed (e.g., faster responses in English than in Bulgarian). However, these differences can be related to many variables, including fundamental differences between the languages. For instance, faster naming times in English could be related to responses always starting with the same definite article ("the") while Bulgarian articles are gender-marked and placed at the end of the noun. Furthermore, although all bilinguals spoke Bulgarian and English, their language profiles varied and their first language in order of acquisition (Bulgarian) was not always the language they used most in their daily lives (with most participants living in the United Kingdom), making it difficult to compare a first versus second language. However, in some bilingual populations, code-switching patterns differ between languages (e.g., more frequent switching to one of the languages, Myslín & Levy, 2015). Furthermore, different bilingual populations also vary in their reasons for switching and the potential use of more bottom-up versus top-down mechanisms (cf. Poplack, 1988). This warrants further research also considering different bilinguals and language pairs, and potential differences between languages within a bilingual.

Conclusion

At a practical level, this study showed how language production can be studied in context, using picture naming times as a frequently used measure that is relatively easy to set up, measure, and compare across studies. Here we show that both switching frequency and switching costs can be measured reliably to assess effects of context. Our data also suggest picture naming without context can still be used as a fast(er) tool to assess voluntary switching where needed. Given the correlations between tasks and comparable split-half reliability, and considering that switching costs were observed in all contexts, when a research project aims to determine a participant's switching frequency and costs, a naming task without context appears no less suitable than a task with context. This has important practical implications, as designing a task with sentence context and/or an interlocutor poses more demands on stimulus creation and increases the time of the study. Furthermore, in many cases a task without context might be preferred because of the effects context can have on the bilingual's behaviour, which in turn can depend on the type of context used.

However, our data also show the importance of studying the role of context, given that language is rarely used in isolation and considering that various contextual factors were found to influence the bilingual's switching behavior. Ideally, therefore, future research would employ a range of measures examining context, also including more naturalistic data such as free conversations and corpus data to provide a more comprehensive understanding of language use.

In conclusion, this study shows previously observed language-switching patterns are not solely because of the use of more artificial tasks that consider word production without any context. At the same time, they highlight the need for more research on contextual effects. Combining measures of switching behavior (frequency) and naming times within the same task, as opposed to focusing on one of the two, can furthermore help to evaluate the possible mechanisms underlying effects of context. Although voluntary language choice and switching might be largely driven by lexical access (how fast a bilingual can name a word in each language, e.g., de Bruin et al., 2018), it does not take place in isolation. Language choice, and as a consequence switching, can be influenced by a range of contextual factors including visual and nonvisual language cues, surrounding language input, the topic of conversation, and information we know about the interlocutor, to name a few examples. Studying different types of context varying in their potential influence on both language behavior and naming times is crucial to increase our understanding of how and why bilinguals switch languages freely, as so many bilinguals frequently do in their daily lives.

References

Anderson, J. A., Mak, L., Keyvani Chahi, A., & Bialystok, E. (2018). The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behavior Research Methods*, *50*(1), 250–263. https://doi.org/10.3758/s13428-017-0867-9

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N. Z., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioural experiment

builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology*, *11*, Article 1699. https://doi.org/10.3389/fpsyg.2020.01699

Blanco-Elorrieta, E., & Pylkkänen, L. (2017). Bilingual language switching in the laboratory versus in the wild: The spatiotemporal dynamics of adaptive language control. *Journal of Neuroscience*, *37*(37), 9022–9036. https://doi.org/10.1523/JNEUROSCI.0553-17.2017

Blanco-Elorrieta, E., & Pylkkänen, L. (2018). Ecological validity in bilingualism research and the bilingual advantage. *Trends in Cognitive Sciences*, *22*(12), 1117–1126. https://doi.org/10.1016/j.tics.2018.10.001

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, *42*(9), 2355–2368. https://doi.org/10.1016/j.pragma.2009.12.012

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

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

de Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Frontiers in Psychology*, *8*, Article 522. https://doi.org/10.3389/fpsyg.2017.00522

de Bruin, A., & Martin, C. D. (2022). Perro or txakur? Bilingual language choice during production is influenced by personal preferences and external primes. *Cognition*, *222*, Article 104995. https://doi.org/10.1016/j.cognition.2021.104995

de Bruin, A., & McGarrigle, R. (in press). Dual-tasking while using two languages: Examining the cognitive resource demands of cued and voluntary language production in bilinguals. *Quarterly Journal of Experimental Psychology*. Advance online publication. https://doi.org/10.1177/17470218231173638

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., & Shiron, V. (2023, May 2). *Language switching in (sentence) context*. https://doi.org/10.17605/OSF.IO/NUQGA

de Bruin, A., & Xu, T. (2023). Language switching in different contexts and modalities: Response-stimulus interval influences cued-naming but not voluntary-naming or comprehension language-switching costs. *Bilingualism: Language and Cognition*, *26*(2), 402–415. https://doi.org/10.1017/S1366728922000554

Declerck, M., Grainger, J., & Hartsuiker, R. J. (2021). Proactive language control during bilingual sentence production. *International Journal of Bilingualism*, *25*(6), 1813–1824. https://doi.org/10.1177/13670069211047803

Declerck, M., Lemhöfer, K., & Grainger, J. (2017). Bilingual language interference initiates error detection: Evidence from language intrusions. *Bilingualism: Language and Cognition*, *20*(5), 1010–1016. https://doi.org/10.1017/S1366728916000845

Declerck, M., & Philipp, A. M. (2015). A sentence to remember: Instructed language switching in sentence production. *Cognition*, *137*, 166–173. https://doi.org/10.1016/j.cognition.2015.01.006

Dewaele, J. M., & Wei, L. (2014). Attitudes towards code-switching among adult mono- and multilingual language users. *Journal of Multilingual and Multicultural Development*, *35*(3), 235–251. https://doi.org/10.1080/01434632.2013.859687

Dijkstra, T., Van Hell, J. G., & Brenders, P. (2015). Sentence context effects in bilingual word recognition: Cognate status, sentence language, and semantic constraint. *Bilingualism: Language and Cognition*, *18*(4), 597–613. https://doi.org/10.1017/S1366728914000388

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*(3), 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition*, *1*(1), 51–66. https://doi.org/10.1017/S1366728998000121

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). Multipic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Elston-Güttler, K. E. (2000). *An enquiry into cross-language differences in lexical-conceptual relationships and their effect on L2 lexical processing* (Doctoral dissertation). University of Cambridge.

Faroqi-Shah, Y., & Wereley, S. (2022). Investigation of code-switching cost in conversation and self-paced reading tasks. *International Journal of Bilingualism*, *26*(3), 308–333. https://doi.org/10.1177/13670069211056438

Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language*, *91*, 181–201. https://doi.org/10.1016/j.jml.2016.04.003

Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. *Journal of Memory and Language*, *89*, 110–137. https://doi.org/10.1016/j.jml.2015.10.001

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(3), 640–665. https://doi.org/10.1037/a0014981

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

Grange, J. A. (2015). *trimr: An implementation of common response time trimming methods*. R package version 1.0.1. https://cran.r-project.org/web/packages/trimr/index.html

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. (2018). Language control and code-switching. *Languages*, *3*(2), Article 8. https://doi.org/10.3390/languages3020008

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

Green, D. W., & Wei, L. (2014). A control process model of code-switching. *Language, Cognition and Neuroscience*, *29*(4), 499–511. https://doi.org/10.1080/23273798.2014.882515

Green, P., & MacLeod, C. J. (2016). Simr: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Gullifer, J. W., Kroll, J. F., & Dussias, P. E. (2013). When language switching has no apparent cost: Lexical access in sentence context. *Frontiers in Psychology*, *4*, Article 278. https://doi.org/10.3389/fpsyg.2013.00278

Ivanov, I. P. (2012). L2 acquisition of Bulgarian clitic doubling: A test case for the Interface Hypothesis. *Second Language Research*, *28*(3), 345–368. https://doi.org/10.1177/0267658312452066

Jevtović, M., Duñabeitia, J. A., & de Bruin, A. (2020). 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

Johns, M. A., & Steuck, J. (2021). Is codeswitching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech. *Cognition*, *211*, Article 104634. https://doi.org/10.1016/j.cognition.2021.104634

Kapiley, K., & Mishra, R. K. (2019). What do I choose? Influence of interlocutor awareness on bilingual language choice during voluntary object naming. *Bilingualism: Language and Cognition*, *22*(5), 1029–1051. https://doi.org/10.1017/S1366728918000731

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

Kootstra, G. J., Dijkstra, T., & Van Hell, J. G. (2020). Interactive alignment and lexical triggering of code-switching in bilingual dialogue. *Frontiers in Psychology*, *11*, Article 1747. https://doi.org/10.3389/fpsyg.2020.01747

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2–3), 259–284. https://doi.org/10.1080/01638539809545028

Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts: A meta-analytic review. *Journal of Memory and Language*, *92*, 217–233. https://doi.org/10.1016/j.jml.2016.06.010

Li, C., & Gollan, T. H. (2022). Language-switch costs from comprehension to production might just be task-switch costs. *Bilingualism: Language and Cognition*, *25*(3), 459–470. https://doi.org/10.1017/S1366728921001061

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(2), 381–390. https://doi.org/10.1037/a0014875

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, Article 226. https://doi.org/10.3389/fpsyg.2013.00226

MacSwan, J. (2000). The architecture of the bilingual language faculty: Evidence from intrasentential code switching. *Bilingualism: Language and Cognition*, *3*(1), 37–54. https://doi.org/10.1017/S1366728900000122

Myslín, M., & Levy, R. (2015). Code-switching and predictability of meaning in discourse. *Language*, *91*(4), 871–905. https://doi.org/10.1353/LAN.2015.0068

Parsons, S. (2021). Splithalf: Robust estimates of split half reliability. *Journal of Open Source Software*, *6*(60), Article 3041. https://doi.org/10.21105/joss.03041

Peeters, D. (2020). Bilingual switching between languages and listeners: Insights from immersive virtual reality. *Cognition*, *195*, Article 104107. https://doi.org/10.1016/j.cognition.2019.104107

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*(4), 633–651. https://doi.org/10.1006/jmla.1998.2592

Poplack, S. (1983). Bilingual competence: Linguistic interference or grammatical integrity? In L. Elias-Olivares (Ed.), *Spanish in the U.S. setting: Beyond the Southwest* (pp. 107–131). National Clearinghouse for Bilingual Education.

Poplack, S. (1988). Contrasting patterns of code-switching in two communities. In M. Heller (Ed.), *Codeswitching: Anthropological and sociolinguistic perspectives* (Vol. 48, pp. 215–244). Walter de Gruyter.

Protopapas, A. (2007). Check Vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

Sánchez, L. M., Struys, E., & Declerck, M. (2022). Ecological validity and bilingual language control: Voluntary language switching between sentences. *Language, Cognition and Neuroscience*, *37*(5), 615–623. https://doi.org/10.1080/23273798.2021.2016873

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language*, *55*(2), 197–212. https://doi.org/10.1016/j.jml.2006.03.004

Tarłowski, A., Wodniecka, Z., & Marzecová, A. (2013). Language switching in the production of phrases. *Journal of Psycholinguistic Research*, *42*(2), 103–118. https://doi.org/10.1007/s10936-012-9203-9

Tomić, A., & Kaan, E. (2022). Oscillatory brain responses to processing code-switches in the presence of others. *Brain and Language*, *231*, Article 105139. https://doi.org/10.1016/j.bandl.2022.105139

Tomić, A., & Valdés Kroff, J. R. (2022). Expecting the unexpected: Code-switching as a facilitatory cue in online sentence processing. *Bilingualism: Language and Cognition*, *25*(1), 81–92. https://doi.org/10.1017/S1366728921000237

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language*, *64*(1), 88–107. https://doi.org/10.1016/j.jml.2010.08.006

Van Hell, J. G., & de Groot, A. M. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica*, *128*(3), 431–451. https://doi.org/10.1016/j.actpsy.2008.03.010

Zhang, Y., Huang, P., Song, Z., Fang, L., Shen, T., Li, Y., Gong, Q., & Xie, P. (2014). In-context language control with production tasks in bilinguals: An fMRI study. *Brain Research*, *1585*, 131–140. https://doi.org/10.1016/j.brainres.2014.08.027

Zhu, J. D., Blanco-Elorrieta, E., Sun, Y., Szakay, A., & Sowman, P. F. (2022). Natural versus forced language switching: Free selection and consistent language use eliminate significant performance costs and cognitive demands in the brain. *NeuroImage*, *247*, Article 118797. https://doi.org/10.1016/j.neuroimage.2021.118797

Received May 2, 2023
Revision received August 25, 2023
Accepted October 6, 2023 ■