Bilingualism: Language and Cognition

cambridge.org/bil

Voluntary and cued language switching in late bilingual speakers

Saskia Mooijman[1], Rob Schoonen[1], Marina B. Ruiter[1] and Ardi Roelofs[2]

[1]Centre for Language Studies, Radboud University, Nijmegen, the Netherlands and [2]Donders Centre for Cognition, Radboud University, Nijmegen, the Netherlands

Research Article

Cite this article: Mooijman, S., Schoonen, R., Ruiter, M.B., & Roelofs, A. (2024). Voluntary and cued language switching in late bilingual speakers. *Bilingualism: Language and Cognition*, **27**, 610–627. https://doi.org/10.1017/S1366728923000755

Received: 15 October 2022
Revised: 22 September 2023
Accepted: 10 October 2023
First published online: 15 November 2023

Keywords:
voluntary language switching; cued language switching; language proficiency; lexical access; late bilingualism

Corresponding author:
Saskia Mooijman
E-mail: saskia.mooijman@ru.nl

Abstract

Previous research examining the factors that determine language choice and voluntary switching mainly involved early bilinguals. Here, using picture naming, we investigated language choice and switching in late Dutch–English bilinguals. We found that naming was overall slower in cued than in voluntary switching, but switch costs occurred in both types of switching. The magnitude of switch costs differed depending on the task and language, and was moderated by L2 proficiency. Self-rated rather than objectively assessed proficiency predicted voluntary switching and ease of lexical access was associated with language choice. Between-language and within-language switch costs were not correlated. These results highlight self-rated proficiency as a reliable predictor of voluntary switching, with language modulating switch costs. As in early bilinguals, ease of lexical access was related to word-level language choice of late bilinguals.

1. Introduction

Bilinguals need to monitor and regulate their languages when they communicate. They adjust their language choice according to their interlocutors and switch between their languages when this is required. In everyday situations, language choice is predetermined when the interlocutor only knows one of the languages. In contexts where interlocutors know both languages about equally well, the bilingual speaker may feel free to choose a language. In such settings, language choice may depend on the speaker's own relative proficiency or the momentary availability of words in each language.

Previous research into the factors that determine language choice and free switching[1] mainly examined balanced bilinguals. For instance, De Bruin et al. (2018, 2020) and Jevtović et al. (2019) studied Spanish–Basque bilinguals living in the Basque Country, where many people have acquired both languages early in life and are highly proficient in both. Moreover, the languages are ever-present in daily life, allowing speakers to freely switch between languages. Importantly, the motivation to switch languages and the effort involved in switching may be different for unbalanced bilinguals. In the Netherlands, for example, speakers typically acquire English relatively late, as a second language (L2) at school. In daily life, use of Dutch is predominant whereas English is only used in specific, well-defined settings. Thus, conclusions derived from findings on early balanced bilinguals living in bilingual contexts may not generalize to late unbalanced bilinguals in largely monolingual contexts.

In the present study, we examine language switching abilities and the factors contributing to switching behavior and language choice in late Dutch–English bilinguals living in the Netherlands. We first review the relevant evidence on cued and voluntary language switching and the mechanisms of switching, and then outline the present study. Next, we describe our methods and results, and end by discussing the implications of our findings.

1.1. Cued and voluntary language switching

Language switching is typically studied using a cued switching task, in which the bilingual participant is prompted to switch based on an external cue, or an alternating runs paradigm, in which switches follow a predictable pattern (e.g., Declerck & Philipp, 2015). Switch trials are those in which a speaker produces another language than in the previous trial. In a seminal paper, Meuter and Allport (1999) found that language switching is costly: response time (RT) is longer on switch compared to repeat trials. In addition to this local, reactive switch cost, more global and proactive mixing costs have also been consistently found (e.g., Christoffels et al., 2007). Mixing costs refer to overall longer RTs in conditions requiring the use of two languages compared to one language.

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

The cost of language switching may be determined by various factors. Language proficiency has a central role in theoretical models of bilingual language processing – for example, by influencing the amount of inhibition required (Green, 1998); by the relative strength of the connections between words and concepts (Kroll & Stewart, 1994); or by the resting level activations for each language (Dijkstra et al., 2018).

Behavioral support for the role of proficiency in switch costs was provided by Bonfieni et al. (2019), who showed that switch costs in both languages were smaller in bilinguals with higher L2 proficiency. However, asymmetrical switch costs are also frequently observed, wherein switching into the more dominant language is slower and more error-prone than switching into the less dominant language (Meuter & Allport, 1999). While there is ample evidence for the existence of asymmetrical switch costs (e.g., Costa & Santesteban, 2004; Gollan et al., 2014; Philipp et al., 2007), some studies failed to find such asymmetry (e.g., Christoffels et al., 2007), or observed a reversed asymmetry (e.g., Bonfieni et al., 2019; C. Liu et al., 2019; Timmer, Christoffels, et al., 2019). In short, both magnitude and symmetry of switch costs may vary depending on proficiency.

While switch costs are a robust finding in the experimental literature, the ecological validity of cued switching paradigms has been questioned (Blanco-Elorrieta & Pylkkänen, 2018). Recent studies have started to investigate voluntary language switching, allowing participants to freely choose the language that first comes to mind. Most voluntary switching experiments still observed switch costs (De Bruin et al., 2018, 2020; Gollan et al., 2014; Gollan & Ferreira, 2009; Grunden et al., 2020; Jevtović et al., 2019; H. Liu et al., 2020; Sánchez et al., 2022; but see Blanco-Elorrieta & Pylkkänen, 2017). However, there is evidence that mixing costs are reduced (Gollan & Ferreira, 2009), or turn into benefits, such that voluntarily mixing languages leads to faster responses than single language conditions (De Bruin et al., 2018, 2020; De Bruin & Xu, 2023; Grunden et al., 2020; Jevtović et al., 2019).

When voluntary and cued switching are compared within the same task, cued switching has been found to result in slower overall responses and larger mixing and switching effects (Jevtović et al., 2019). Gollan et al. (2014) and De Bruin and Xu (2023) also reported (some) benefits in voluntary over cued language switching, whereas De Bruin et al. (2018) observed that both types of switching yielded comparable switch costs but overall faster responses for the voluntary than cued switching task.

Studies that investigate voluntary switching in balanced bilinguals who live in bilingual societies typically observe a voluntary switching rate of around 40%. Similarly high switching rates have been observed for many, though not all, late or unbalanced bilinguals (De Bruin & Xu, 2023; H. Liu et al., 2020; Sánchez et al., 2022; but see H. Liu et al., 2021). Across studies, bilinguals vary in their switching rate. Gollan and Ferreira (2009) reported that balanced bilinguals switch more frequently (35%) than unbalanced bilinguals (24%). Furthermore, the bilinguals in their study named the easiest items (regarding frequency, length, retrieval speed, and accuracy) in the non-dominant language, leaving the relatively more difficult items to be named in the stronger language. This shows that bilinguals may predominantly switch languages when items in the non-dominant language are relatively accessible.

Gollan and Ferreira (2009) presented first evidence for a relationship between language abilities and language choice. Additional support for this relation was provided by Sarkis and Montag (2021), who showed that lexical accessibility predicted code switches in a sentence production experiment. Furthermore, De Bruin et al. (2018) demonstrated that ease of lexical access, operationalized as item-level differences in naming latencies between the L1 and L2, was a predictor for language choice in the voluntary condition. Interestingly, participant's overall L2 proficiency or use, undoubtedly related to ease of lexical access, was not related to language choice or switching costs in this study.

While asymmetrical switch costs, interpreted as an index of the influence of relative language proficiency, are frequently observed in cued language switching, there is little evidence for an asymmetry in voluntary switch costs. Although larger switch costs into the dominant language were observed by H. Liu et al. (2021), the majority of studies did not find a relationship between language and magnitude of switch costs (De Bruin et al., 2018, 2020; Gollan et al., 2014; Gollan & Ferreira, 2009; Grunden et al., 2020; Jevtović et al., 2019), or found an asymmetry in the opposite direction (De Bruin & Xu, 2023; Sánchez et al., 2022).

1.2 Mechanisms of language switching

The ubiquity of switch costs in language switching suggests that top-down language control processes are involved. The extent to which bilingual language control abilities overlap with domain-general control processes has been debated in the literature. Several studies found commonalities between language switching and non-linguistic task switching, both in terms of correlations (Declerck et al., 2017; Prior & Gollan, 2011) and overlap in neural circuits (De Baene et al., 2015; De Bruin et al., 2014; Weissberger et al., 2015). Furthermore, language switching experience has been found to relate to non-linguistic task switching performance (Barbu et al., 2018; Festman & Münte, 2012; Timmer, Calabria, et al., 2019; Verreyt et al., 2016). However, other studies failed to find (complete) overlap between linguistic and non-linguistic switching (Branzi et al., 2016; Calabria et al., 2012, 2015; Klecha, 2013; Segal et al., 2019; Timmer et al., 2018; Weissberger et al., 2012). In short, the current literature is inconclusive regarding domain generality of bilingual language control abilities (see Declerck & Philipp, 2015; Jiao et al., 2022, for reviews). It has been argued that the conflicting results can, at least in part, be explained by task-related differences in response modality, stimuli type, and cues (Declerck et al., 2017; Declerck & Philipp, 2015).

Focusing on voluntary switching, De Bruin et al. (2018) found that voluntary language switching costs were related to linguistic, but not non-linguistic, inhibition. Gollan et al. (2014) compared voluntary and cued switch costs in linguistic and non-linguistic tasks and found that advantages for voluntary over cued switching may be more pronounced in non-linguistic switching than in language switching, especially when items are not repeated. The authors take this as evidence that language switching mostly relies on domain-specific mechanisms.

Importantly, switching is not exclusive to bilingual language production. Instead, every speaker needs switching skills to communicate effectively – for example, by switching between registers or syntactic constructions (i.e., within-language switching). Declerck et al. (2020) compared control mechanisms in between- and within-language switching tasks. They found evidence for overlap between control processes but also saw that the switch costs were differentially influenced by manipulations of the interval between cues and stimuli.

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

Sikora and colleagues also studied within-language switching (Sikora & Roelofs, 2018; Sikora et al., 2016a, 2016b, 2019). They designed a picture naming task in which participants produced short (e.g., *chair*) or long noun phrases (e.g., *green chair*) and were cued to switch between these phrases. Sikora et al. (2016b) found no correlation between the overall within-language switch costs and switch costs in a non-linguistic switching task. However, the results of a reaction-time distribution analysis provided evidence for the engagement of domain-general switching ability in within-language switching.

Despite disagreement in the literature about the overlap between linguistic and non-linguistic switching, the two domains appear to have similarities when task demands are kept the same. In addition, switching can be measured between and within languages, and these types of switching tend to overlap and induce switch costs. This raises the question how cued within-language switching is related to cued and voluntary between-language switching when comparable methods are used.

1.3 Summary

To recapitulate, switching in voluntary contexts remains costly, although switch costs may be reduced. Switching may be motivated by language abilities when operationalized as language dominance. Bilinguals' language proficiency, on the other hand, was not found to be related to voluntary switching and mixing costs, although this was investigated in a group of early bilinguals with a high level of proficiency in both languages. For bilinguals with more varying language abilities, the effect of proficiency on voluntary switching costs may be larger. Moreover, late bilinguals may show asymmetry in their voluntary switch costs, but the direction and prevalence of this asymmetry is not yet established. Finally, the domain generality of bilingual language switching abilities remains controversial to date, and one approach to advance is by comparing tasks with highly similar characteristics.

1.4 Research questions

We aimed to replicate and extend earlier research into voluntary switching and systematically investigated the differences and commonalities between voluntary and cued between-language switching, and cued within-language switching. Specifically, we intended to contribute to the existing literature by (1) investigating a group of late bilinguals with varying degrees of L2 proficiency living in their L1 environment, (2) focusing on the role of proficiency and ease of lexical access on language switching, and (3) investigating the extent of overlap between different types of language switching. We addressed the following research questions:

1. Do late bilinguals switch between their languages in a voluntary switching task, and can their voluntary switching behavior be explained by relative ease of lexical access and/or L2 proficiency?
2. Do voluntary and cued language switching induce similar switch costs, and can these switch costs be explained by L2 proficiency and/or more general switching abilities as measured with a cued within-language switching task?
3. To what extent do cued between-language and within-language switching abilities overlap?

1.5 Testing language production in an online setting

We examined our research questions in a web-based setting. The COVID-19 pandemic resulted in closing of labs and social distancing measures, which made it impossible to conduct in-person experiments for a considerable period of time. Using web-based tools is a way to collect behavioral data when labs are closed. Web-based tools have advantages, such as ease, efficiency, flexibility of data collection, and remote testing. Yet, there are certain risks too, especially for language production studies like the current study, in which accurate measurement of latencies is crucial. Questionable quality of speech recording and concerns about timing are common objections to web-based testing, due to instabilities of the experimental program, operating system, internet browser and internet speed (Fairs & Strijkers, 2021; He et al., 2021). Several studies have demonstrated the feasibility and accuracy of web-based language production studies. Although overall latencies are longer in web-based than in lab-based experiments (Fairs & Strijkers, 2021), frequently-observed psycholinguistic effects that rely on precise and accurate measurement of naming latencies are replicated (Stark et al., 2022; Vogt et al., 2021).

2. Method

2.1 Participants

Forty native speakers of Dutch with English as their L2 took part in this study[2]. They were all students at Dutch universities or universities of applied sciences. To ensure variation in English proficiency, we recruited half of the participants from (under)graduate programs with English as the language of instruction. The remaining participants were enrolled in studies with Dutch as the language of instruction. Dutch proficiency was not assessed, but considering that all participants were native speakers of Dutch, lived in the Netherlands full-time, and qualified for higher education, it was assumed their Dutch proficiency was at the highest level. Participants' English proficiency was verified with self-ratings and a lexical test (LexTALE, Lemhöfer & Broersma, 2012). Visual inspection of the distribution of the proficiency measures showed that this way of recruiting participants resulted in a wide range of scores on the proficiency measures, but also revealed that the distributions between the two groups overlapped sufficiently to be treated as a single group.

We encountered technical issues in the final part of the experiment for one participant and we decided to exclude this participant from the analysis. All included participants had (corrected to) normal hearing and vision. Each participant read and signed an informed consent form prior to participation. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional Ethics committee (2019-5035) on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. The participants' characteristics are presented in Table 1.

2.2 Materials

Participants completed a questionnaire about demographic and language variables. The questions were based on the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian et al., 2007) and the Language History Questionnaire (LHQ 2.0, Li et al., 2014). The questionnaire consisted of 30 questions and was administered using the Qualtrics platform (Qualtrics, 2005). Self-rated proficiency was assessed by using a sliding scale ranging

Table 1. Characteristics of the participants.

Characteristic	Statistic			
Sex	*n*	%		
Female	32	82.1		
Male	7	17.9		
	M	*SD*	*Min*	*Max*
Age (years)	21.9	2.8	18.4	29.0
LexTALE English (%)	81.1	12.6	61.3	100.0
Daily language switching: 1 (*never*) – 5 (*very often*)	3.9	0.9	2	5
Age of acquisition English	9.9	1.9	4	14
Years usage English	10.9	3.0	6	18
Self-rated proficiency Dutch (0–100)				
Comprehension	98.0	3.7	90	100
Production	97.8	4.7	80	100
Reading	98.5	4.2	80	100
Writing	95.0	6.8	79	100
Self-rated proficiency English (0-100)				
Comprehension	84.6	13.9	41	100
Production	77.9	17.2	20	100
Reading	87.0	12.8	50	100
Writing	78.8	14.6	40	100
Frequency usage Dutch: 1 *(never)* – 5 *(daily)*				
Home	5.0	0.0	5	5
Family	4.9	0.4	3	5
Friends	4.8	0.6	2	5
Study	3.8	1.5	1	5
Frequency usage English: 1 *(never)* – 5 *(daily)*				
Home	2.4	1.5	1	5
Family	1.8	1.2	1	5
Friends	3.3	1.5	1	5
Study	4.3	0.9	2	5

from 0 ("poor") to 100 ("excellent"). The questionnaire ended with the LexTALE, a 60-item word recognition test for advanced learners of English that has been shown to correlate well with general English proficiency (Lemhöfer & Broersma, 2012).

In all naming tasks, the same thirty 8×8 cm colored line drawings from the MultiPic database were used (Duñabeitia et al., 2018). Pictures were named in Dutch or English, depending on the task instruction. The target words for the pictures were one- or two-syllabic non-cognate words (Appendix A). All target words were frequent (SUBTLEX log10 frequency ≥ 2.0), acquired early (age of acquisition ≤ 7.0), highly prevalent (rating $\geq 1.7/2$) and concrete (rating $\geq 4.3/5$). Word variables and ratings were based on various databases (Brysbaert et al., 2014, 2019; Brysbaert & New, 2009; Keuleers et al., 2010, 2015; Kuperman et al., 2012). See Appendix B for mean values of the variables of the stimuli.

2.3 Design and procedure

The study involved an online testing procedure that lasted approximately 45 minutes per participant. The experimental tasks were administered in a peer-to-peer video call in Zoom (Zoom Video Communications Inc., 2012). This allowed the experimenter to monitor the test session in a way that was similar to lab-based settings. The experimental materials were shown using PowerPoint via screen sharing.

Experiments started with a familiarization task in which participants saw each picture with the two printed target words (in English and Dutch) and they were asked to read both words out loud. The four experimental tasks (further explained below) were presented in a fixed order: (1) picture naming in Dutch and English blocks, with the block order counterbalanced, (2) voluntary picture naming in Dutch or English, (3) cued picture naming in Dutch or English, and (4) cued within-language switching between phrase types in Dutch. Participants started with the single-language task to measure naming in each language separately. The voluntary switching task was administered before the cued between-language switching task to avoid priming switching behavior. The last task in the protocol was the cued within-language switch task, because this task only required naming in Dutch and it was a relatively new task compared to the more established between-language switching tasks. The tasks were not repeated to prevent fatigue, which was expected to be particularly probable in a web-based setting.

Each naming task started with an instruction and practice items, and consisted of 60 trials divided over blocks separated by short breaks. Instructions were given verbally by the experimenter and shown on the screen in Dutch, English, or a mix of both languages. Speed and accuracy of naming was emphasized in all instructions. The target pictures were presented twice in all tasks, randomized using Mix (Van Casteren & Davis, 2006), with a constraint that the repetition of each item did not follow within at least 10 trials. Two versions of the experiment were created to control for any effect of starting language of the first task, with a different starting language and a different randomization order of trials in each version. The versions were counterbalanced between participants. An overview of the experimental tasks and blocks is presented in Appendix C.

In the single-language task, participants named pictures in blocks of Dutch and English. The task consisted of four blocks of 15 trials each. The voluntary language switching task consisted of two blocks with 30 trials each, and participants were instructed to name the pictures in the language that first came to mind.

The two cued switching tasks (between-language and within-language) had the same design. Both tasks consisted of two blocks and involved cued and predictable switching in a SWITCH–REPEAT–SWITCH–REPEAT order, with a cue presented preceding and simultaneous with each item. Using ALTERNATING RUNS in a cued switching experiment has been found to elicit robust and reliable switch costs (De Bruin et al., 2020; Declerck et al., 2015; Jackson et al., 2001; Rogers & Monsell, 1995). Furthermore, a predictable switch pattern eliminates having to manipulate potentially confounding

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

factors associated with unpredictable cued switching, such as run length (Zheng et al., 2018) and preparation time effects (see Jost et al., 2013; Kiesel et al., 2010; Koch et al., 2018, for discussions). Additionally, the need to rely solely on the cue signalling an upcoming switch has been argued to obscure the distinction between cue-encoding processes and task switching (Logan & Bundesen, 2003, 2004; Schneider & Logan, 2005). Therefore, we considered the alternating switching paradigm a reliable and pure measure of switch costs.

The cued between-language switching task involved participants naming pictures in either Dutch or English based on a visually presented country flag serving as a reminder. The pictures were separated across the two versions of the experiment, such that participants named an item in either English or Dutch, not both, to avoid potential interference effects resulting from translation equivalents (Kleinman & Gollan, 2018). Items were about equally divided based on word length, frequency, and visual complexity of the picture.

The design of the cued within-language switching task involved participants naming color or size properties of the depicted object, together with the target word, indicated by a color bar or ruler. All line drawings were edited such that they were in red or blue, and big (14 × 14 cm) or small (6 × 6 cm). The target language for this task was Dutch.

The trial structures of the four tasks are illustrated in Figure 1. Pictures were shown for maximally 3000 ms but when items were named before the end of the trial, the experimenter manually initiated the next trial at the offset of the participant's answer. Every trial was preceded by a 500 ms interval, during which a fixation cross (single-language and free switching tasks) or the visual cue (cued switching tasks) was shown. To enable annotation of the audio files and extraction of the response latencies later, a click sound was presented at the onset of each picture.

We tried to manage possible negative artefacts of web-based experiments in several ways. First, we used a live connection and participants carried out the experiment whilst being in a videocall with the experimenter. Therefore, we believe that the level of distraction was similar to an in-person experiment. Second, RTs were manually extracted and this way, we could evaluate each individual data point before entering it into the statistical analysis and discard trials affected by glitches. Any remaining glitches were expected to occur randomly across trials and would therefore not systematically affect the results. Overall slower internet connections of specific participants were accounted for in the statistical analysis by including random intercepts for participants.

2.4 Analyses

The experiment was recorded in Zoom (Zoom Video Communications Inc., 2012) and audio files were annotated manually in Praat (Boersma & Weenink, 2022) to extract naming latencies. RTs were operationalized as the time between the onset of the click sound and the speech onset of the participant's response. A portion of the data (i.e., 720 observations) was independently coded by two raters (the first author and a research assistant) to establish the inter-rater reliability of the RT annotation. We analyzed the inter-rater reliability using a single-measurement, consistency, two-way random-effects model using 'irr' package (Gamer et al., 2012) in R, version 4.1.2 (R Core Team, 2022) in RStudio version 2021.09.1 (RStudio Team, 2023). The Intraclass Correlation Coefficient (ICC) calculation showed excellent agreement between raters (ICC = .91, 95% confidence interval = .89 – .92) and the remainder of the data was annotated by a single coder. The responses were categorized based on the (adapted) classification by De Bruin et al. (2018), but due to the low error rates, only binary accuracy (correct/

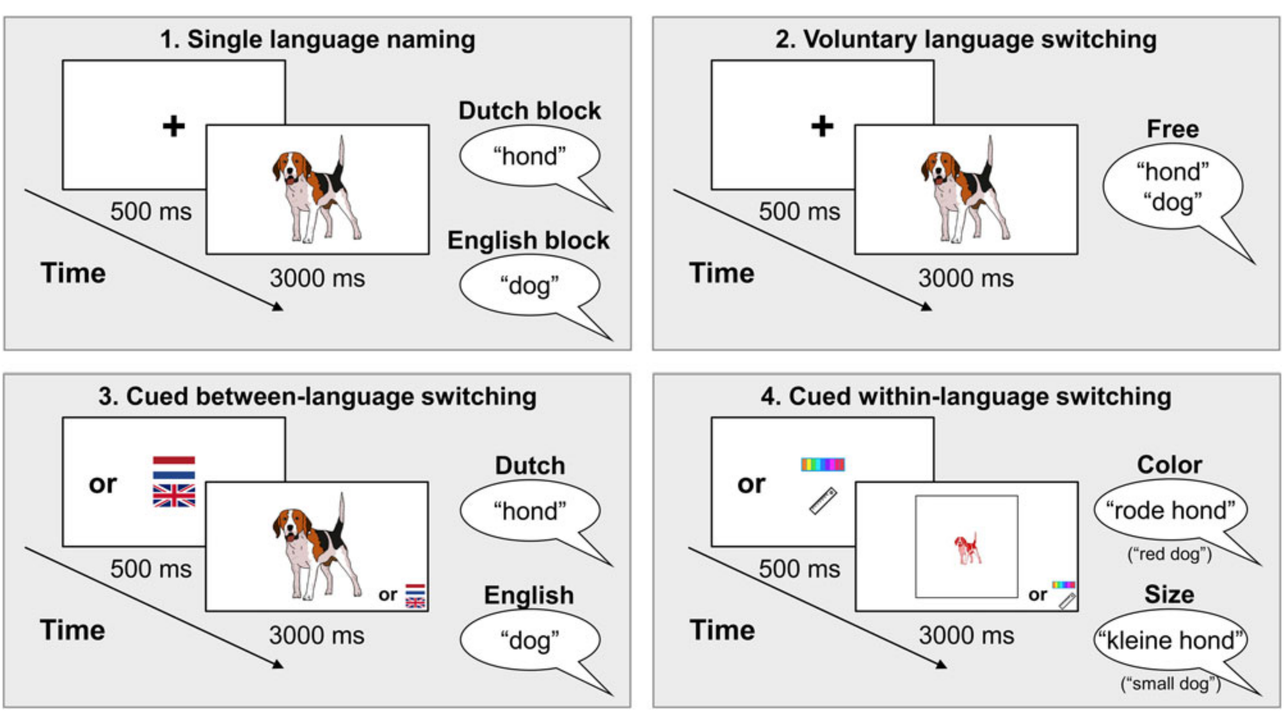

Figure 1. Illustration of the trial structures of the four naming tasks.

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

incorrect) was further analyzed (see Appendix D for the full classification scheme).

We statistically analyzed the data using packages 'lme4', 'lmerTest', 'emmeans', 'tidyverse' and 'ggplot2' (Bates et al., 2015; Lenth, 2022; Wickham, 2016; Wickham et al., 2019). We did not remove outliers, but RTs of < 500 ms or > 3000 ms and the first trials of a task or immediately after a break were excluded. For the RT analyses, incorrectly answered items were discarded. The trials preceded by a mistake were not removed because trial type (switch/repeat, L1/L2) was predetermined by the task demands or could still be established in all errors. There were missing data points ($n = 196$, 2.1%) due to technical glitches (e.g., inaudible clicks, problems with recording, or connection hiccups). In total, 689 data points (7.4%) were excluded from the RT analysis. In the accuracy analysis, we excluded 215 data points (2.3%) in total, 12 (0.1%) due to glitches and 203 (2.2%) first trials.

We ran multiple (generalized) linear mixed-effects regression models to answer our research questions. All models were fit with the theoretically informed maximal random structure that was possible without convergence issues (Barr et al., 2013) and we used the 'bobyqa' optimizer to prevent any convergence problems. To investigate factors related to voluntary switching, SWITCHING BEHAVIOR (switch: yes/no and language choice: Dutch/English) was predicted by OBJECTIVE L2 PROFICIENCY (LexTALE scores), SUBJECTIVE L2 PROFICIENCY (self-rated L2 proficiency averaged across ratings for production, comprehension, reading, and writing), LEXICAL ACCESSIBILITY (response-speed difference between Dutch and English items on the single-language task), and SWITCHING ABILITIES (switch costs on the cued within-language switching task). In these models, all predictors were scaled (centered and standardized) to address convergence warnings.

To investigate differences between different types of switching, we ran (generalized) linear mixed-effects regression models with accuracy or RTs as outcome variables. TASK (voluntary switching and cued switching), LANGUAGE (Dutch or English target items), TRIAL TYPE (switch or repeat trials), OBJECTIVE PROFICIENCY, SUBJECTIVE PROFICIENCY, and SWITCHING ABILITIES were included in the voluntary versus cued between-language switching model. Lexical accessibility was not included because this predictor was operationalized as relative naming speed and thus resembled the outcome variable (naming speed on the same items, albeit in different tasks) too much to be a meaningful predictor. Through model comparison, we saw that TRIAL NUMBER had a significant effect on RTs ($\chi^2(1) = 4.75$, $p = .029$) and we therefore included it as a covariate. In our model comparing cued between-language switching to within-language switching, we limited our predictors to TASK, TRIAL TYPE, and TRIAL NUMBER.

The continuous predictors were centered around the mean, and categorical predictors were sum-coded (-1 or +1). The interpretation of multi-level predictors was facilitated by an omnibus test and post-hoc pairwise comparisons with a Bonferroni-correction of the p-value. RTs were (natural) log transformed to reduce skewness. The model assumptions were checked visually (heterogeneity of variance and normally distributed residuals) and by inspecting Variable Inflation Factors for multicollinearity (all VIFs < 2.0).

Importantly, there was a moderate positive correlation between objective and subjective L2 proficiency ($r = .60$, $p < .001$). As this may be problematic for regression analyses, we decided to run separate models for each of the two measures of proficiency for every research question. Except for these two proficiency variables, the models were identical in terms of the remaining fixed- and random-effects structure. After fitting the models, we compared the Akaike Information Criterion (AIC) values of the two models. The main effects of objective and subjective L2 proficiency are always presented, but we only report the remaining fixed effects of the best fitting model.

3. Results

3.1 Descriptive statistics

Accuracy (in percentage correct) was highest in the free switching task ($M = 99$, $SD = 11$), followed by single-language naming across L1 and L2 ($M = 98$, $SD = 15$), cued language switching ($M = 97$, $SD = 17$), and within-language switching ($M = 90$, $SD = 30$). The RTs (in ms) showed a similar pattern. On average, RTs were shortest in free switching ($M = 968$, $SD = 280$), followed by cued switching ($M = 1037$, $SD = 309$), single-language naming ($M = 1041$, $SD = 322$), and within-language switching ($M = 1148$, $SD = 380$).

3.2 Free switching behavior

To answer the research question whether late bilinguals switch between their languages and what predicts their switching behavior, we analyzed the results of the free switching task. On average, participants switched on 41.5% of the trials (range 7-65%) and approximately half of the items (53.4%) were named in English, with considerable variation between participants (range 3–93%)[3].

Our first generalized linear mixed-effects model with switching (*yes* (1) or *no* (0)) as a dependent variable showed no significant effect of objective proficiency on switching behavior ($OR = 1.00$, $SE = 0.11$, $p = .982$). The second model, including subjective proficiency, was better fitting as indicated by a lower AIC value. This model showed that subjective proficiency was related to voluntary switching ($OR = 1.28$, $SE = 0.13$, $p = .021$). This indicates that participants who rated their own proficiency higher switched more, regardless of direction of the switches. There were no significant effects of any other predictors (Appendix E).

In the next step, we fitted the same predictors and included trial type in the model with language choice (*English* (1) or *Dutch* (0)) as outcome variable (Table 2 and Figure 2). Here too, we found that objective proficiency was not significantly related to language choice ($OR = 1.24$, $SE = 0.21$, $p = .201$, Figure 2A) and that the model including subjective proficiency better predicted language choice. There was a significant main effect of subjective proficiency, indicating that English was chosen more frequently by participants who rated their own L2 proficiency as higher ($OR = 1.87$, $SE = 0.23$, $p < .001$, Figure 2B). This factor interacted with trial type ($OR = 1.59$, $SE = 0.10$, $p < .001$), showing that the effect was driven by the repeat trials. Ease of lexical retrieval had a significant effect on language choice. Items named more quickly in English in the single-language task were subsequently named more often in English in the free switching task, and vice versa for Dutch ($OR = 0.64$, $SE = 0.04$, $p < .001$, Figure 2C). Lexical accessibility did not significantly interact with trial type ($OR = 0.97$, $SE = 0.05$, $p = .613$), and we did not find a significant effect of within-language switch costs on language choice ($OR = 0.97$, $SE = 0.11$, $p = .799$, Figure 2D).

3.3 Factors predicting voluntary and cued between-language switching

We compared switch costs in free and cued between-language switching and investigated whether both types of switching

Table 2. Switch behavior analysis: language choice on free switching task.

	Language choice (English/Dutch)			
Effect[a]	*Odds ratio*	*SE*	*95% CI*	*p-value*
Intercept	1.19	0.15	0.93 – 1.53	.174
Trial type	1.21	0.06	1.10–1.34	**<.001**
Self-rated proficiency	1.87	0.23	1.47–2.38	**<.001**
Lexical accessibility (RT Δ)	0.64	0.04	0.58 – 0.72	**<.001**
Within-language switch cost	0.97	0.11	0.78–1.21	.799
Trial type × Self-rated proficiency	1.59	0.10	1.40–1.80	**<.001**
Trial type × Lexical accessibility (RT Δ)	0.97	0.05	0.88–1.08	.613
Trial type × Within-language switch cost	1.00	0.05	0.90–1.10	.976

Note. Number of observations = 2117. [a]All predictors were scaled.

were equally predicted by proficiency and within-language switching abilities. The model output for accuracy showed no significant effect of subjective proficiency ($OR = 0.99$, $SE = 0.20$, $p = .977$) and the model including objective proficiency was a better fit. The results revealed higher accuracy on repeat trials than on switch trials ($OR = 1.59$, $SE = 0.33$, $p = .025$), indicative of an overall switch cost, and on the free task compared to the cued task ($OR = 0.56$, $SE = 0.12$, $p = .005$). Task and objective proficiency interacted ($OR = 1.03$, $SE = 0.02$, $p = .040$), indicating that participants with higher L2 proficiency had significantly higher accuracy only in cued switching. We did not find a significant effect of language or other interactions between predictors (Appendix F).

The RT models showed no significant effect of subjective proficiency ($\beta = -0.001$, $SE = 0.002$, $p = .534$), and the model including objective proficiency fitted the data better. We observed significant main effects of trial type, task, language, and trial number (Appendix G), but no significant main effect of objective proficiency ($\beta = -0.003$, $SE = 0.002$, $p = .217$) or within-language switching ability ($\beta = 0.0004$, $SE = 0.0004$, $p = .316$). A significant effect of trial number ($\beta = -0.0004$, $SE = 0.0002$, $p = .029$) indicates that participants became faster as the task progressed. The interpretation of the main effects of trial type, task, and language was complicated by the presence of a three-way interaction effect between these predictors ($\beta = 0.006$, $SE = 0.003$, $p = .029$), demonstrating that the trial type effect (i.e., switch cost) was differentially predicted by task and language. Furthermore, the predictors were part of a four-way interaction with objective proficiency ($\beta = 0.0006$, $SE = 0.0002$, $p = .011$). We first examined this four-way interaction effect by creating subsets of the data based on language.

For the English trials (Figure 3, upper panels), there was an interaction between task and trial type ($\beta = -0.008$, $SE = 0.003$, $p = .018$), revealing that the RT difference between switch and repeat trials (indicated in light grey color) was larger in the cued than voluntary condition. Furthermore, we found a significant three-way interaction effect between task, trial type, and proficiency ($\beta = 0.0006$, $SE = 0.0003$, $p = .024$). This effect was mainly driven by cued switching, since there was a significant interaction between trial type and proficiency in the cued switching condition ($\beta = 0.0009$, $SE = 0.0004$, $p = .033$), but not in the voluntary condition ($\beta = -0.0004$, $SE = 0.0004$, $p = .316$). These results show that switch costs in English were smaller in the voluntary than cued condition, and that this difference was larger for participants with lower English proficiency.

For the Dutch items (Figure 3, lower panels), we failed to find a significant two-way interaction between trial type and task ($\beta = 0.004$, $SE = 0.004$, $p = .413$). This indicates that despite overall longer RTs for switch than repeat trials ($\beta = -0.02$, $SE = 0.004$, $p < .001$), and for items in the cued compared to the voluntary condition ($\beta = 0.05$, $SE = 0.01$, $p < .001$), the magnitude of the switch cost in Dutch did not differ significantly between tasks. Despite a trend visible in the graph, the interaction between task, trial type, and proficiency was not significant for Dutch ($\beta = -0.0005$, $SE = 0.0004$, $p = .140$).

To evaluate switch cost asymmetries, we examined the interaction between task, trial type, and language by creating subsets of the data based on task. The results of the cued condition (Figure 3, left panels) confirmed that participants were slower on switch than repeat trials ($\beta = 0.02$, $SE = 0.004$, $p < .001$) and slower on Dutch items than English items ($\beta = 0.02$, $SE = 0.008$, $p = .006$), but there was no statistical evidence for a difference in switch costs between the two languages ($\beta = 0.006$, $SE = 0.004$, $p = .158$). For free switching (Figure 3, right panels), we corroborated the switch cost ($\beta = 0.02$, $SE = 0.004$, $p < .001$), but did not find a significant main effect of language ($\beta = 0.006$, $SE = 0.007$, $p = .418$). However, these predictors significantly interacted ($\beta = 0.008$, $SE = 0.004$, $p = .044$), with larger switch costs into the L1 than the L2.

To search for an explanation of this voluntary switch cost asymmetry, we inspected the strategies participants adopted in the free switching task. We evaluated the effect of SWITCHING FREQUENCY on switching costs. An interaction effect demonstrated that participants who switched more, experienced smaller voluntary switch costs ($\beta = 0.01$, $SE = 0.005$, $p = .003$). Focusing on the individual participants, they appear to have implemented different approaches: (a) Dutch as default ($N = 5$), with a majority of items in Dutch and few switches into English; (b) English as default ($N = 5$), with a majority of items in English and few switches into Dutch; (c) English default, Dutch as backup ($N = 10$), with a majority of items in English and one-time switches into Dutch; (d) Dutch default, English as backup ($N = 1$), with a

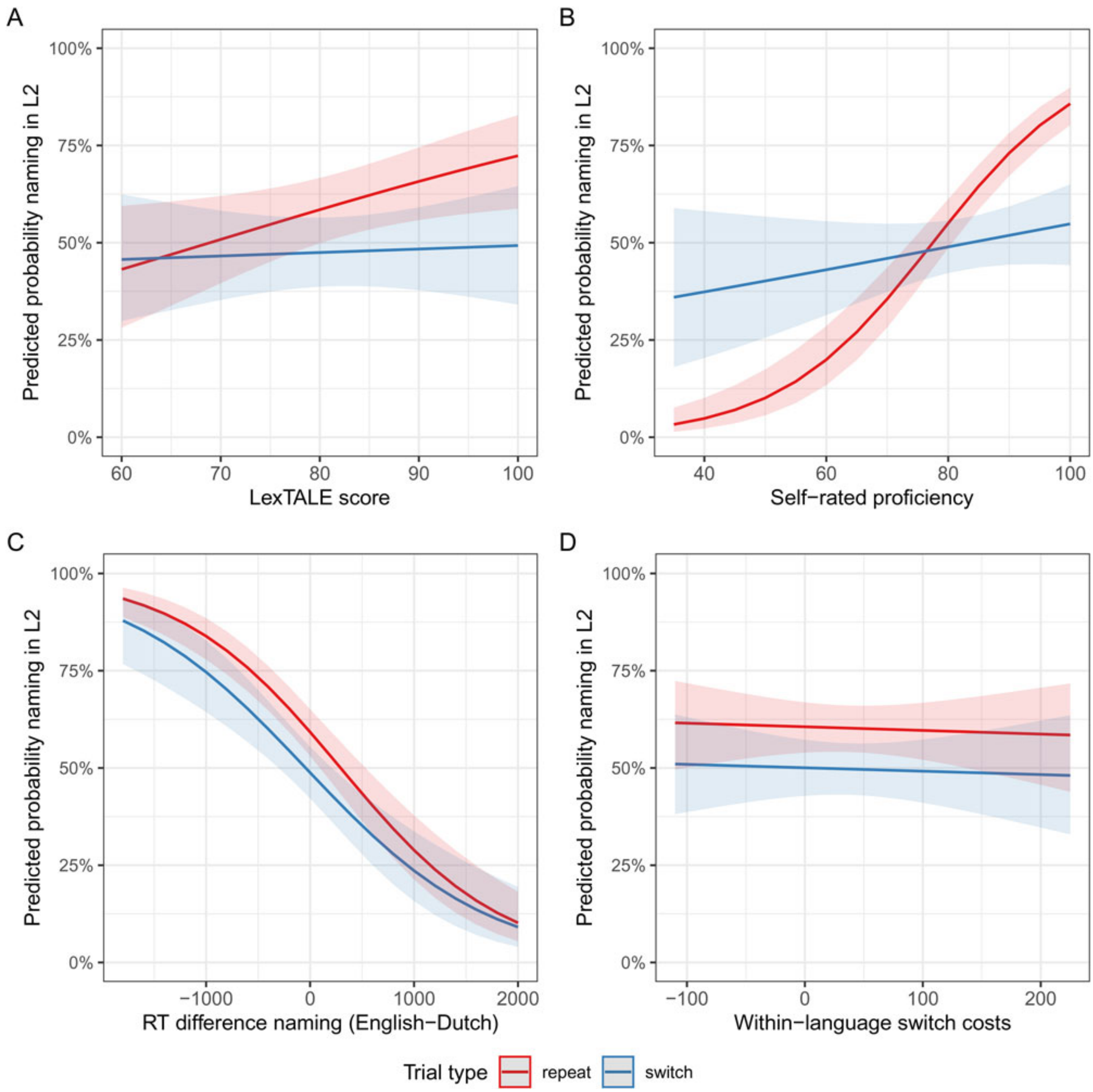

Figure 2. Predicted probabilities of language choice in switch and repeat trials as a function of objective (A) and self-rated proficiency (B), ease of lexical access (C), and within-language switch costs (D).

majority of items in Dutch and one-time switches into English; and (e) frequent switchers ($N = 18$), with many switches into both languages, and a comparable number of switch and repeat trials.

When we included INDIVIDUAL STRATEGY in a linear model with trial type and language predicting the RTs in the voluntary switching task, we observed a significant interaction effect between these predictors ($\chi^2(4) = 11.32$, $p = .023$), implying that the strategies were differentially related to the switch costs into the L1 and L2 (Figure 4). These results suggest that participants who switched infrequently showed larger switch costs, and further show that the magnitude of switch costs depended on the language they switched into. Overall, the Dutch switch costs appear larger than the English switch costs. The participant with the 'Dutch default, English as backup' approach presents as an outlier in terms of strategy, overall RTs, and switch costs.

3.4 Comparison of cued between-language and within-language switching

We compared participants' performance on the cued between-language and cued within-language switching tasks (Appendices H and I). Accuracy was higher in the between-language than the within-language switch task ($OR = 2.03$, $SE = 0.16$, $p < .001$). There were switch costs in both tasks ($OR = 1.46$, $SE = 0.11$, $p < .001$), but there was no significant interaction between trial type and task ($OR = 1.14$, $SE = 0.09$, $p = .099$).

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

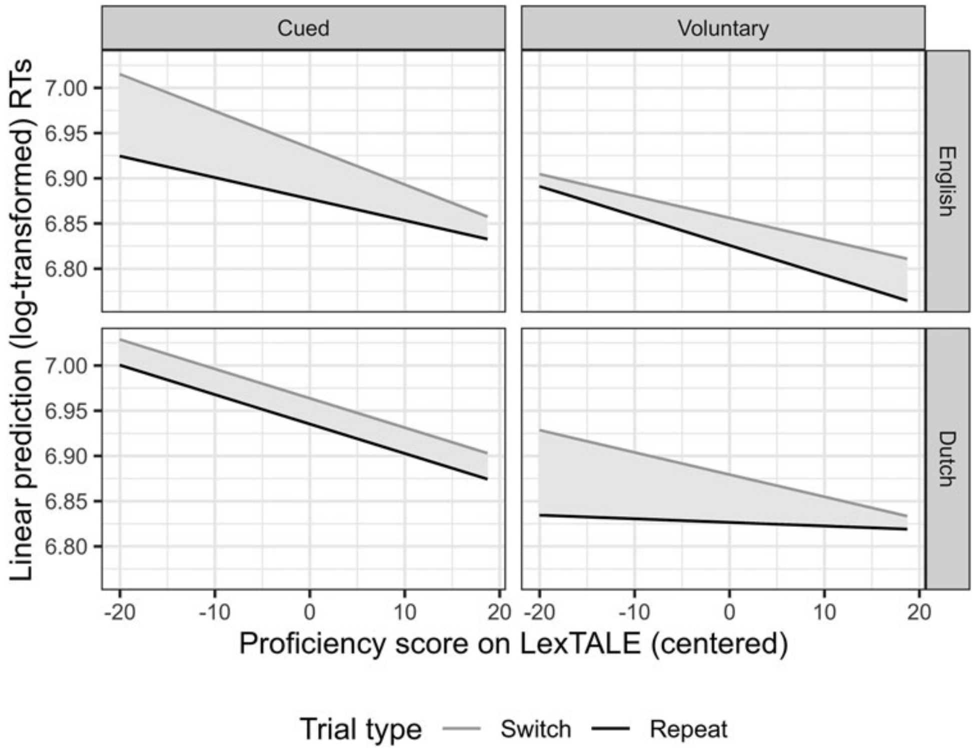

Figure 3. Linear prediction of the four-way interaction effect of task, language, trial type, and proficiency (centered scores on the LexTALE) for RTs on the switch tasks.

In correspondence with the accuracy results, RTs were longer in the within-language switch task as compared to the between-language switch task ($\beta = 0.04$, $SE = 0.003$, $p < .001$). We observed a significant switch cost across tasks ($\beta = 0.02$, $SE = 0.003$, $p < .001$), although there was no statistical evidence that these costs differed between tasks ($\beta = 0.003$, $SE = 0.003$, $p = .417$).

Finally, we carried out a correlation analysis of the switch costs and RTs on the cued between-language and within-language

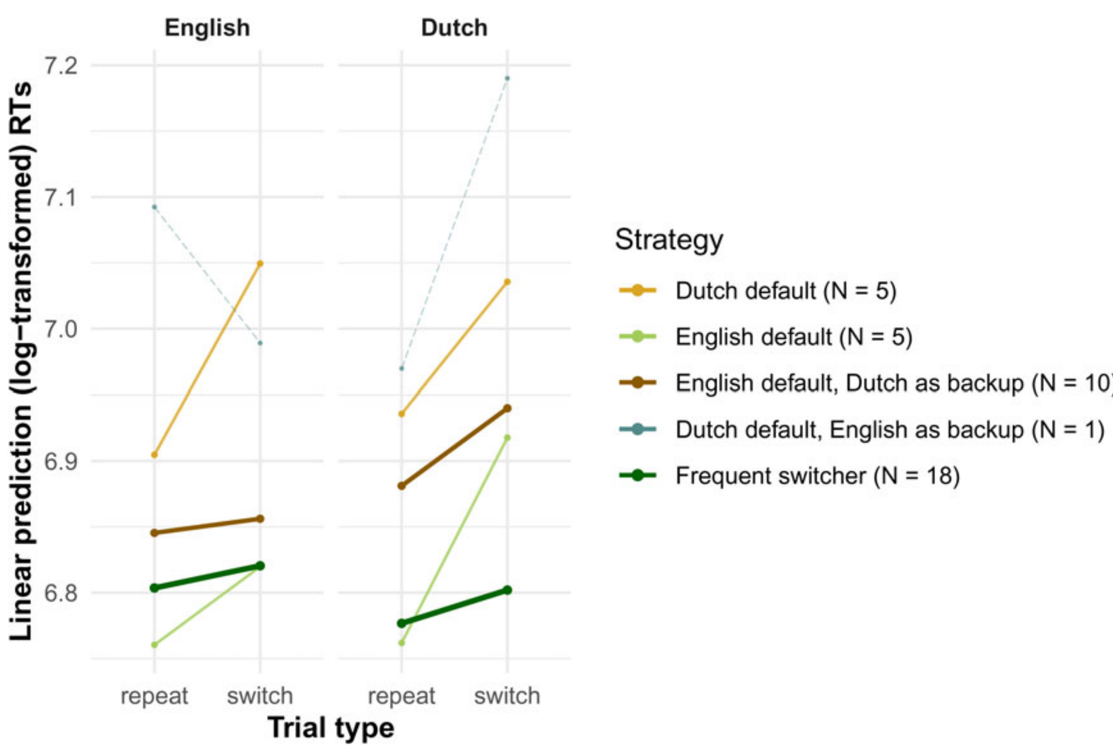

Figure 4. Visualization of the linear predictions of RTs by language, trial type, and individual strategy.

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

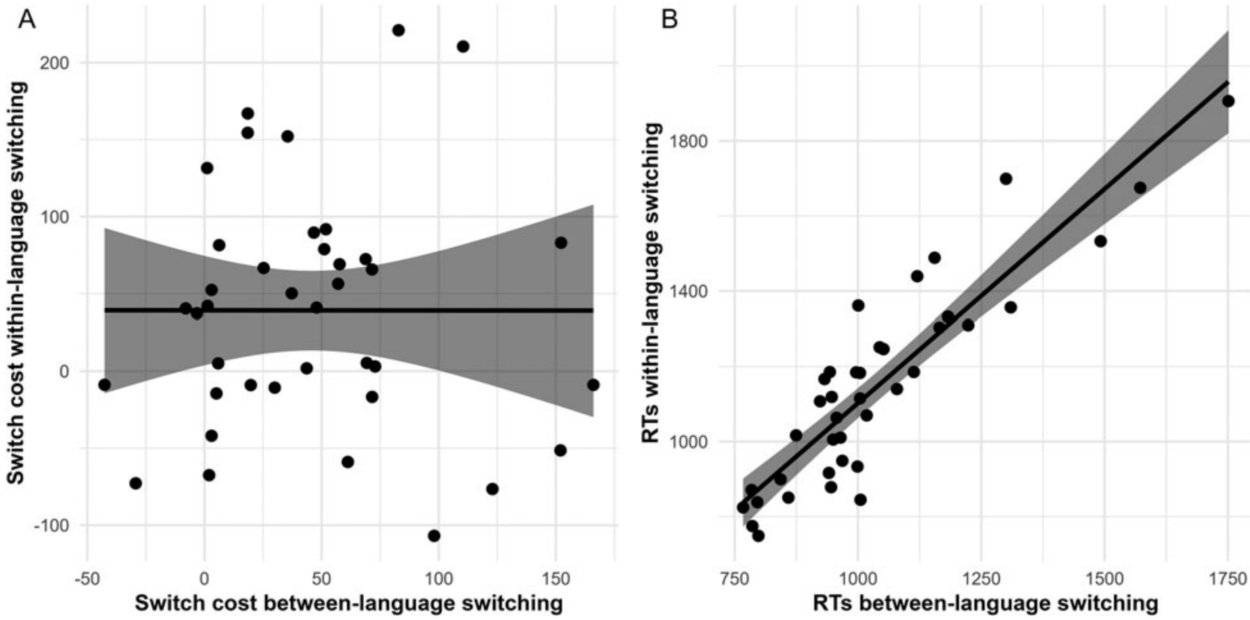

Figure 5. Correlation plots of cued between-language and within-language switching costs (A) and overall RTs (B).

switching tasks (Figure 5). There was no significant correlation of the switch costs between tasks ($r = -.0006$, $p = .997$), whereas the overall RTs showed a strong positive correlation ($r = .90$, $p < .001$).

4. Discussion

We investigated voluntary and cued language switching abilities and the factors contributing to switching behavior and language choice in late Dutch–English bilinguals. In what follows, we elaborate on our main findings.

4.1 Language switching motivations

Our results of the voluntary switching task show that participants switched on 42% of the trials, which closely matches previous research on early balanced bilinguals (Blanco-Elorrieta & Pylkkänen, 2018; De Bruin et al., 2018, 2020; Grunden et al., 2020; Jevtović et al., 2019) and unbalanced bilinguals (De Bruin & Xu, 2023; H. Liu et al., 2020; Sánchez et al., 2022). The observed switching frequency is higher than reported by Gollan and Ferreira (2009), who showed that unbalanced bilinguals switched 24% of the time. This may be due to methodological differences between the studies, as Gollan and Ferreira used more difficult target items (with regards to word frequency and length) and did not repeat stimuli. At the same time, Gollan et al. (2014) observed similar switching rates in voluntary switching tasks with and without repetition of stimuli.

The similarity between switching frequencies may be surprising given the differences in type of bilinguals between the studies. The bilinguals in our study can be regarded as less balanced and show more variability in their proficiency levels compared to most previous studies, who generally included bilinguals with a lower age of acquisition, higher proficiency, more balanced use, and who live in bilingual societies in which everyday language switching is more common (i.e., the Basque country, Catalonia, Arabic–English bilingual community, southern California) or who currently live in their L2 environment (De Bruin & Xu, 2023). The studies by H. Liu et al. (2020, 2021) are exceptions. They investigated Chinese–English bilinguals living in an L1 context with limited English proficiency and observed varying switching rates, perhaps due to differences between the two experimental designs. Sánchez et al. (2022) investigated late unbalanced bilinguals (their current language context is not given) and observed high (47%) switching rates in voluntary switching between sentences.

The relatively high switching frequency we observed is remarkable given that code-switching is rather rare in everyday conversations in the Dutch society. A likely explanation is that the number of switches is experimentally induced by a high activation level of English and Dutch as a consequence of the task instructions and experimental set-up. Both languages are explicitly made equally appropriate in the voluntary switching task, and the familiarization and single-language task directly preceded the free switching task. This experimental set-up was also used in earlier studies involving late bilinguals (De Bruin & Xu, 2023; H. Liu et al., 2020, 2021). Contrary to Gollan and Ferreira (2009), the target language of the preceding block in the single-language task did not influence the switching rates, although methodological differences, particularly pre-exposure to the pictures, could have contributed to the diverging findings. Future studies could further explore how experimentally induced voluntary switching generalizes to more naturalistic settings (Blanco-Elorrieta & Pylkkänen, 2017). However, it is also possible that switching frequency is predominantly induced by the conversational situation rather than the language situation in the society at large. Future studies may examine this possibility by manipulating the utility of voluntary switching by varying the frequency of monolingual or bilingual contextual cues (e.g., Zheng et al., 2020).

Our results of the voluntary switching task further showed that bilinguals switched more often and chose to name items in English more frequently, if they rated their own L2 proficiency as higher. This effect of proficiency was larger for language choice of the repeat than the switch trials, which indicates that participants with high self-rated proficiency were more inclined to

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

produce repeat items in English. These participants may have adopted English as their default language, and thus switched less frequently overall. Objective proficiency, operationalized as the score on the LexTALE, did not significantly predict switching behavior nor language choice. The absence of a significant main effect of objective proficiency corresponds with earlier research (De Bruin et al., 2018), whereas the relationship between self-rated proficiency and language switching and choice is a new finding, as self-rated proficiency has not yet been considered as a potential predictor for language choice.

This outcome points to a possible role for LANGUAGE ATTITUDE as an important factor for language switching in this context, complementing research showing that personal language preference of balanced bilinguals may guide language choice (De Bruin & Martin, 2022). Effects of language attitude may be even more pronounced in the Dutch–English bilinguals in our study who, in some cases, have deliberately chosen to pursue their university degree in English. Language attitude can be regarded as a top-down, socio-psychological motivation to switch languages, and has long been established as an important reason to code switch in naturalistic settings (Ritchie & Bhatia, 2012). The question of whether balanced bilinguals also choose a language according to their self-rated proficiency remains open for investigation.

Moreover, ease of lexical access predicted language choice: items that were named relatively more quickly in English than in Dutch in the single-language tasks were more likely to be named in English in the free switching task (and vice versa) – regardless of whether it concerned switch or repeat trials. In other words, participants tended to name the item in the language that was easiest to retrieve, even if that required switching languages. This outcome is in line with earlier findings of switching in a sentence production task (Sarkis & Montag, 2021) and in a picture naming task for balanced bilinguals reported by De Bruin et al. (2018), although lexical access was operationalized slightly differently in both studies. These results imply that the bilinguals in our study encounter subtle differences in lexical retrieval speed and that this affects their language choice, which results in increased efficiency of picture naming. This matches findings for early balanced bilinguals and strengthens the evidence for the role of ease of lexical access in language choice, at least in experimental settings.

4.2 Comparing cued and voluntary switching

Our analyses showed that participants were overall faster and made fewer errors in voluntary switching than cued switching, but that both tasks induced switch costs. The comparison of RTs between the two tasks further showed larger cued than voluntary switch costs for the English items, whereas an opposite trend was observed for Dutch. Diminished switch costs in voluntary switching have been observed previously (Gollan et al., 2014; Gollan & Ferreira, 2009; Jevtović et al., 2019; Zhang et al., 2015), but differences in switch costs in the two languages between task types are not yet well established. Our observation is consistent with Jevtović et al. (2019), who found that switch costs in Basque (L2), but not Spanish (L1), were larger in cued than voluntary switching.

Relatedly, we observed an asymmetrical switch cost only in voluntary switching, where participants experienced larger costs for switching into the L1 than the L2. Asymmetrical voluntary switch costs have been observed in late unbalanced bilinguals before (H. Liu et al., 2021), whereas other studies report reversed asymmetrical switch costs (De Bruin & Xu, 2023; Sánchez et al., 2022). Differences in language development and context between bilinguals might underlie the variability in findings of asymmetrical voluntary switch costs. A direct comparison between bilingual groups is warranted to investigate this further.

Asymmetrical switch costs are often explained in terms of control, whereby more inhibition is required for the more dominant language (e.g., Meuter & Allport, 1999; Philipp et al., 2007). Control demands can also result in REVERSE DOMINANCE EFFECTS, which refer to worse performance in the dominant language compared to the non-dominant language when both languages are mixed (see Declerck & Koch, 2023; Goldrick & Gollan, 2023, for reviews). Proactive control is said to function as a preventative mechanism to minimize cross-language interference resulting from language non-selective activation (see Declerck, 2020, for a review). The finding that participants in our study were overall faster to name the English than Dutch items is in line with this account. Correspondingly, many participants adopted English as their default language in the voluntary switching task. They may have exerted more inhibition over their stronger L1 than L2, resulting in longer L1 naming latencies and asymmetrical voluntary switch costs. This interpretation is supported by the observation that reverse dominance effects started to emerge only in the later blocks of the single-language condition. In other words, repeated exposure to two languages caused interference and increased the inhibitory demands, resulting in reversed dominance effects.

However, an absence of asymmetrical switch costs in the cued switching task warrants an additional explanation for the large voluntary switch costs into Dutch. We tried to clarify this pattern by focusing on the strategies participants adopted in the free switching task. This showed that participants became more efficient switchers when they switched more frequently, and appeared to have implemented different strategies that impacted switch costs in each language. Frequent and flexible switchers showed small switch costs in both languages, while participants with a clear default language showed relatively high voluntary switch costs into both languages, but more pronounced for their non-default language.

We tentatively take this as evidence for a role of ease of lexical access in explaining the voluntary language switch costs: participants with a clear default language may decide to switch to the other language only when they encounter a lexical retrieval difficulty. This is time-consuming, which is subsequently interpreted as a large switch cost. Importantly, more participants used English as their default language ($N = 15$) than Dutch ($N = 6$). This could potentially contribute to the large voluntary switch cost observed for Dutch, which, in turn, eliminated the voluntary switching advantage for Dutch. The precise mechanism underlying this effect, and why it governs switching into the L1 more prominently than switching into the L2, should be investigated in future research.

Our experimental tasks were presented in the order that was deemed optimal given the methodological constraints. However, as a result of the item repetition and fixed order in which the tasks were presented, the mixing costs of the voluntary and cued switching task could not be compared. While the finding that participants were overall faster on voluntary switching than cued switching and single-language naming can be regarded as evidence in the direction of an overall processing advantage in the voluntary condition (i.e., mixing benefits), repetition effects cannot be ruled out entirely. Because the same set of pictures was used in all tasks and the tasks were presented in the same order, participants may have become faster with each repetition of an item. Speeding up in the voluntary switching compared to the single-language task could be due to repeating stimuli, and the effect of slowing down in the cued task would likely have been larger if items had not been repeated. To adequately assess mixing costs, a single-language block should have been added after the switching tasks (De Bruin et al., 2018; Grunden et al., 2020; Jevtović et al., 2019). The effects of item repetition on bilingual picture naming are detailed in Kleinman and Gollan (2018).

4.3 Contributors to language switching abilities

Our results showed that differences in switch costs between cued and voluntary switching appeared to be moderated by L2 proficiency. More specifically, L2 proficiency and task interacted only for the English items, and played a role in cued, rather than voluntary switching into English. This could suggest that cued switching into the L2 was particularly difficult for participants with relatively low proficiency. In other words, when you MUST switch to a language of relatively low proficiency, switching takes more time. This points to a role for lexical accessibility in explaining the cued switch costs into English. This interaction was not statistically significant for the Dutch items, which may be due to the relatively smaller number of voluntary trials named in Dutch. Because these results reflect a complex interaction effect, a replication in a larger group of participants is warranted.

A potential reason for the absence of a main effect of proficiency on the naming latencies may be that the experiments did not place high demands on lexical knowledge of the participants. The highly frequent target words were repeated throughout the experiment and we familiarized participants with all items prior to the experiments. In addition, the LexTALE is a receptive vocabulary test and may therefore not directly relate to naming speed in the experiments.

In addition to proficiency, we examined the effect of more general switching abilities on cued and voluntary switching. We administered a picture naming task that required participants to make within-language switches, inspired by Sikora and colleagues (Sikora & Roelofs, 2018; Sikora et al., 2016a, 2016b, 2019). We did not find evidence that within-language switch costs predicted voluntary or cued switching behavior or abilities. A separate analysis showed that participants experienced within-language switch costs, implying that the task was successful in capturing switching abilities. We found no statistical evidence for differences or associations between the between-language and within-language switch costs. At the same time, the overall RTs on the two tasks were strongly correlated. These results could suggest that the performance on the two picture naming tasks showed considerable overlap, but that the within-language switching costs specifically may play only a minor role in between-language switching, which could provide tentative evidence for a more domain-specific nature of bilingual language switching (Branzi et al., 2016; Calabria et al., 2012, 2015; Klecha, 2013; Prior & Gollan, 2011).

However, there are two alternative explanations for the absence of significant correlations between the switch costs. Firstly, concerns have been raised about the reliability of difference scores and this could explain the discrepancy in findings between the switch costs (a difference score) and overall RTs (e.g., Draheim et al., 2019; Segal et al., 2021). Secondly, methodological differences may hinder the comparison between types of switching (Declerck et al., 2017, 2020). We tried to make our switching tasks as comparable as possible (i.e., both tasks required a verbal response, had similar cue presentation, and an alternating-runs design), but the tasks inevitably diverged in some ways. For between-language switching, the target word in the competing language was the only response alternative, whereas within-language switching had the competing target property (color or size) as the most prominent response alternative, but participants could also omit the target property altogether and produce only the bare noun. In addition, the required responses were more complex in the within-language task (adjective and noun) than the between-language switch task (bare noun). These differences between task designs are difficult to avoid but complicate drawing firm conclusions about the domain specificity of bilingual language switching.

4.4 Testing language production online

Our study confirms that collecting language production data in a web-based setting is feasible (Fairs & Strijkers, 2021; Stark et al., 2022; Vogt et al., 2021). Our data were somewhat noisier compared to data gathered in the lab, but we consider less than 3% missing data due to technical glitches acceptable. In addition, the average RTs in our data (~1000 ms) were higher than typically reported for bilingual picture naming studies (~800–900 ms), but we were nonetheless able to capture switch costs. The switch costs were in the expected direction and resembled those gathered in lab-based settings in terms of relative magnitude. Furthermore, the online data collection process was efficient, easy to carry out, and required little technical equipment. Thus, our study contributes to the evidence that a web-based administration of language experiments is a suitable method of data collection for future research.

5. Summary and conclusion

This study systematically investigated voluntary and cued language switching in late Dutch–English bilinguals, measured in a web-based setting. Our results suggest that late bilinguals behave similarly to early balanced bilinguals regarding several aspects of language switching. Their voluntary switching frequency resembled that of early balanced bilinguals, they experienced switch costs in cued as well as voluntary switching, and ease of lexical access contributed to their language choice. Moreover, our results demonstrated that self-rated proficiency rather than objective proficiency predicted voluntary switching behavior. Participants were overall slower to name pictures in cued switching than in voluntary switching. The magnitude of the switch costs for each task differed between the L1 and L2, which could partially be explained by individual approaches to voluntary switching adopted by participants. An interaction effect with proficiency revealed that switching into the L2 is particularly difficult if switching is not optional and proficiency is relatively low. Finally, there was considerable overlap in performance on between-language and within-language switching tasks, while the switch costs specifically were not significantly related. This study highlights the similarities in language switching between different types of bilinguals and provides insight into the factors that are related to voluntary language switching.

Data availability. The data that support the findings of this study are openly available at https://osf.io/gd2wv/.

Acknowledgments. The authors want to thank Rivka van den Berg for her help with annotating the data.

Competing interests. The authors declare none.

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

Notes

1 The terms 'voluntary and 'free' switching are used interchangeably, as are 'mandatory' and 'cued' switching.

2 One participant was born outside of the Netherlands, but acquired Dutch before the age of five, and we therefore decided to leave this participant in the dataset.

3 The observed variation between participants could not be explained by the two participant groups, as participants attending English instruction university showed similar switching rates compared to participants attending Dutch instruction university.

References

Barbu, C., Orban, S., Gillet, S., & Poncelet, M. (2018). The impact of language switching frequency on attentional and executive functioning in proficient bilingual adults. *Psychologica Belgica*, *58*(1), 115–127. https://doi.org/10.5334/pb.392

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3). https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/doi:10.18637/jss.v067.i01

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

Blanco-Elorrieta, E., & 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

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.2.10) [Computer software]. http://www.praat.org/

Bonfieni, M., Branigan, H. P., Pickering, M. J., & Sorace, A. (2019). Language experience modulates bilingual language control: The effect of proficiency, age of acquisition, and exposure on language switching. *Acta Psychologica*, *193*, 160–170. https://doi.org/10.1016/j.actpsy.2018.11.004

Branzi, F. M., Calabria, M., Boscarino, M. L., & Costa, A. (2016). On the overlap between bilingual language control and domain-general executive control. *Acta Psychologica*, *166*, 21–30. https://doi.org/10.1016/j.actpsy.2016.03.001

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi-org.ru.idm.oclc.org/10.3758/BRM.41.4.977

Brysbaert, M., Stevens, M., De Deyne, S., Voorspoels, W., & Storms, G. (2014). Norms of age of acquisition and concreteness for 30,000 Dutch words. *Acta Psychologica*, *150*, 80–84. https://doi.org/10.1016/j.actpsy.2014.04.010

Brysbaert, M., Mandera, P., McCormick, S. F., & Keuleers, E. (2019). Word prevalence norms for 62,000 English lemmas. *Behavior Research Methods*, *51*(2), 467–479. https://doi.org/10.3758/s13428-018-1077-9

Calabria, M., Hernandez, M., Branzi, F. M., & Costa, A. (2012). Qualitative differences between bilingual language control and executive control: Evidence from task-switching. *Frontiers in Psychology*, *2*, 399. https://doi.org/10.3389/fpsyg.2011.00399

Calabria, M., Branzi, F. M., Marne, P., Hernández, M., & Costa, A. (2015). Age-related effects over bilingual language control and executive control. *Bilingualism: Language and Cognition*, *18*(1), 65–78. https://doi.org/10.1017/S1366728913000138

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208. https://doi.org/10.1016/j.brainres.2007.01.137

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 Baene, W., Duyck, W., Brass, M., & Carreiras, M. (2015). Brain circuit for cognitive control is shared by task and language switching. *Journal of Cognitive Neuroscience*, *27*(9), 1752–1765. https://doi.org/10.1162/jocn_a_00817

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*, 104995. https://doi.org/10.1016/j.cognition.2021.104995

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

De Bruin, A., Roelofs, A., Dijkstra, T., & FitzPatrick, I. (2014). Domain-general inhibition areas of the brain are involved in language switching: FMRI evidence from trilingual speakers. *NeuroImage*, *90*, 348–359. https://doi.org/10.1016/j.neuroimage.2013.12.049

De Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2018). Voluntary language switching: When and why do bilinguals switch between their languages? *Journal of Memory and Language*, *103*, 28–43. https://doi.org/10.1016/j.jml.2018.07.005

De Bruin, A., Samuel, A. G., & Duñabeitia, J. A. (2020). Examining bilingual language switching across the lifespan in cued and voluntary switching contexts. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(8), 759–788. https://doi.org/10.1037/xhp0000746

Declerck, M. (2020). What about proactive language control? *Psychonomic Bulletin & Review*, *27*(1), 24–35. https://doi.org/10.3758/s13423-019-01654-1

Declerck, M., & Koch, I. (2023). The concept of inhibition in bilingual control. *Psychological Review*, *130*(4), 953–976. https://doi.org/10.1037/rev0000367

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic Bulletin & Review*, *22*(6), 1630–1645. https://doi.org/10.3758/s13423-015-0836-1

Declerck, M., Koch, I., & Philipp, A. M. (2015). The minimum requirements of language control: Evidence from sequential predictability effects in language switching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 377–394. https://doi.org/10.1037/xlm0000021

Declerck, M., Grainger, J., Koch, I., & Philipp, A. M. (2017). Is language control just a form of executive control? Evidence for overlapping processes in language switching and task switching. *Journal of Memory and Language*, *95*, 138–145. https://doi.org/10.1016/j.jml.2017.03.005

Declerck, M., Ivanova, I., Grainger, J., & Duñabeitia, J. A. (2020). Are similar control processes implemented during single and dual language production? Evidence from switching between speech registers and languages. *Bilingualism: Language and Cognition*, *23*(3), 694–701. https://doi.org/10.1017/S1366728919000695

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2018). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/S1366728918000287

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, *145*(5), 508–535. https://doi.org/10.1037/bul0000192

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

Fairs, A., & Strijkers, K. (2021). Can we use the internet to study speech production? Yes we can! Evidence contrasting online versus laboratory naming latencies and errors. *PLOS ONE*, *16*(10), e0258908. https://doi.org/10.1371/journal.pone.0258908

Festman, J., & Münte, T. F. (2012). Cognitive Control in Russian–German bilinguals. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00115

Gamer, M., Lemon, J., Fellows, I., & Puspendra, S. (2012). *irr. Various coefficients of interrater reliability and agreement* (0.84.1) [Computer software]. https://cran.r-project.org/web/packages/irr/irr.pdf.

Goldrick, M., & Gollan, T. H. (2023). Inhibitory control of the dominant language: Reversed language dominance is the tip of the iceberg. *Journal of Memory and Language*, *130*, 104410. https://doi.org/10.1016/j.jml.2023.104410

Gollan, T. H., & 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

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

Grunden, N., Piazza, G., García-Sánchez, C., & Calabria, M. (2020). Voluntary language switching in the context of bilingual aphasia. *Behavioral Sciences*, *10*(9), Article 9. https://doi.org/10.3390/bs10090141

He, J., Meyer, A. S., Creemers, A., & Brehm, L. (2021). Conducting language production research online: A web-based study of semantic context and name agreement effects in multi-word production. *Collabra: Psychology*, *7* (1), 29935. https://doi.org/10.1525/collabra.29935

Jackson, G. M., Swainson, R., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*(2), 169–178. https://doi.org/10.1017/S1366728901000268

Jevtović, M., Duñabeitia, J. A., & de Bruin, A. (2019). How do bilinguals switch between languages in different interactional contexts? A comparison between voluntary and mandatory language switching. *Bilingualism: Language and Cognition*, *23*(2), 401–413. https://doi.org/10.1017/S1366728919000191

Jiao, L., Meng, N., Wang, Z., Schwieter, J. W., & Liu, C. (2022). Partially shared neural mechanisms of language control and executive control in bilinguals: Meta-analytic comparisons of language and task switching studies. *Neuropsychologia*, *172*, 108273. https://doi.org/10.1016/j.neuropsychologia.2022.108273

Jost, K., De Baene, W., Koch, I., & Brass, M. (2013). A review of the role of cue processing in task switching. *Zeitschrift Für Psychologie*, *221*(1), 5–14. https://doi.org/10.1027/2151-2604/a000125

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00174

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, *68*(8), 1665–1692. https://doi.org/10.1080/17470218.2015.1022560

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin*, *136*(5), 849–874. https://doi.org/10.1037/a0019842

Klecha, A. (2013). Language and task switching in Polish–English bilinguals. *Psychology of Language and Communication*, *17*(1), 17–36. https://doi.org/10.2478/plc-2013-0002

Kleinman, D., & Gollan, T. H. (2018). Inhibition accumulates over time at multiple processing levels in bilingual language control. *Cognition*, *173*, 115–132. https://doi.org/10.1016/j.cognition.2018.01.009

Koch, I., Poljac, E., Müller, H., & Kiesel, A. (2018). Cognitive structure, flexibility, and plasticity in human multitasking—An integrative review of dual-task and task-switching research. *Psychological Bulletin*, *144*(6), 557–583. https://doi.org/10.1037/bul0000144

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lenth, R. V. (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.7.2) [Computer software]. https://CRAN.R-project.org/package=emmeans

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

Liu, C., Timmer, K., Jiao, L., Yuan, Y., & Wang, R. (2019). The influence of contextual faces on bilingual language control. *Quarterly Journal of Experimental Psychology*, *72*(9), 2313–2327. https://doi.org/10.1177/1747021819836713

Liu, H., Tong, J., de Bruin, A., Li, W., He, Y., & Li, B. (2020). Is inhibition involved in voluntary language switching? Evidence from transcranial direct current stimulation over the right dorsolateral prefrontal cortex. *International Journal of Psychophysiology*, *147*, 184–192. https://doi.org/10.1016/j.ijpsycho.2019.12.002

Liu, H., Li, W., de Bruin, A., & He, Y. (2021). Should I focus on self-language actions or should I follow others? Cross-language interference effects in voluntary and cued language switching. *Acta Psychologica*, *216*, 103308. https://doi.org/10.1016/j.actpsy.2021.103308

Logan, G. D., & Bundesen, C. (2003). Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 575–599. https://doi.org/10.1037/0096-1523.29.3.575

Logan, G. D., & Bundesen, C. (2004). Very clever homunculus: Compound stimulus strategies for the explicit task-cuing procedure. *Psychonomic Bulletin & Review*, *11*(5), 832–840. https://doi.org/10.3758/BF03196709

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*(1), 25–40. https://doi.org/10.1006/jmla.1998.2602

Philipp, A. M., Gade, M., & Koch, I. (2007). Inhibitory processes in language switching: Evidence from switching language-defined response sets. *European Journal of Cognitive Psychology*, *19*(3), 395–416. https://doi.org/10.1080/09541440600758812

Prior, A., & Gollan, T. H. (2011). Good language-switchers are good task-switchers: Evidence from Spanish–English and Mandarin–English bilinguals. *Journal of the International Neuropsychological Society*, *17*(04), 682–691. https://doi.org/10.1017/S1355617711000580

Qualtrics (August 2020). (2005). [Computer software]. Qualtrics. https://www.qualtrics.com/

R Core Team. (2022). *R: A language and environment for statistical computing.* (4.2.2) [Computer software]. https://www.R-project.org/

Ritchie, W. C., & Bhatia, T. K. (2012). Social and psychological factors in language mixing. In *The Handbook of Bilingualism and Multilingualism* (pp. 375–390). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118332382.ch15

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, *124*, 205–231. https://doi-org.ru.idm.oclc.org/10.1037/0096-3445.124.2.207

RStudio Team. (2023). *RStudio: Integrated Development for R.* (2023.03.1) [Computer software]. RStudio, PBC. http://www.rstudio.com/

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

Sarkis, J. T., & Montag, J. L. (2021). The effect of lexical accessibility on Spanish–English intra-sentential codeswitching. *Memory & Cognition*, *49* (1), 163–180. https://doi.org/10.3758/s13421-020-01069-7

Schneider, D. W., & Logan, G. D. (2005). Modeling task switching without switching tasks: A short-term priming account of explicitly cued performance. *Journal of Experimental Psychology: General*, *134*, 343–367. https://doi.org/10.1037/0096-3445.134.3.343

Segal, D., Stasenko, A., & Gollan, T. H. (2019). More evidence that a switch is not (always) a switch: Binning bilinguals reveals dissociations between task and language switching. *Journal of Experimental Psychology: General*, *148*, 501–519. https://doi.org/10.1037/xge0000515

Segal, D., Prior, A., & Gollan, T. H. (2021). Do all switches cost the same? Reliability of language switching and mixing costs. *Journal of Cognition*, *4* (1), Article 1. https://doi.org/10.5334/joc.140

Sikora, K., & Roelofs, A. (2018). Switching between spoken language-production tasks: The role of attentional inhibition and enhancement. *Language, Cognition and Neuroscience*, *33*(7), 912–922. https://doi.org/10.1080/23273798.2018.1433864

Sikora, K., Roelofs, A., & Hermans, D. (2016a). Electrophysiology of executive control in spoken noun-phrase production: Dynamics of updating, inhibiting, and shifting. *Neuropsychologia*, *84*, 44–53. https://doi.org/10.1016/j.neuropsychologia.2016.01.037

Sikora, K., Roelofs, A., Hermans, D., & Knoors, H. (2016b). Executive control in spoken noun-phrase production: Contributions of updating, inhibiting, and shifting. *Quarterly Journal of Experimental Psychology*, *69*(9), 1719–1740. https://doi.org/10.1080/17470218.2015.1093007

Sikora, K., Roelofs, A., Hermans, D., & Knoors, H. (2019). Executive control in language production by children with and without language impairment. *International Journal of Language & Communication Disorders*, *54*, 645–655. https://doi.org/10.1111/1460-6984.12470

Stark, K., van Scherpenberg, C., Obrig, H., & Abdel Rahman, R. (2022). Web-based language production experiments: Semantic interference assessment is robust for spoken and typed response modalities. *Behavior Research Methods*, *55*, 236–262. https://doi.org/10.3758/s13428-021-01768-2

Timmer, K., Calabria, M., Branzi, F. M., Baus, C., & Costa, A. (2018). On the reliability of switching costs across time and domains. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.01032

Timmer, K., Calabria, M., & Costa, A. (2019). Non-linguistic effects of language switching training. *Cognition*, *182*, 14–24. https://doi.org/10.1016/j.cognition.2018.09.001

Timmer, K., Christoffels, I. K., & Costa, A. (2019). On the flexibility of bilingual language control: The effect of language context. *Bilingualism: Language and Cognition*, *22*(3), 555–568. https://doi.org/10.1017/S1366728918000329

Van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584–589. https://doi-org.ru.idm.oclc.org/10.3758/BF03193889

Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, *19*(01), 181–190. https://doi.org/10.1017/S1366728914000352

Vogt, A., Hauber, R., Kuhlen, A. K., & Rahman, R. A. (2021). Internet-based language production research with overt articulation: Proof of concept, challenges, and practical advice. *Behavior Research Methods*, *54*, 1954–1975. https://doi.org/10.3758/s13428-021-01686-3

Weissberger, G. H., Wierenga, C. E., Bondi, M. W., & Gollan, T. H. (2012). Partially overlapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging*, *27*(4), 959–974. https://doi.org/10.1037/a0028281

Weissberger, G. H., Gollan, T. H., Bondi, M. W., Clark, L. R., & Wierenga, C. E. (2015). Language and task switching in the bilingual brain: Bilinguals are staying, not switching, experts. *Neuropsychologia*, *66*, 193–203. https://doi.org/10.1016/j.neuropsychologia.2014.10.037

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1–6. https://doi.org/10.21105/joss.01686

Zhang, Y., Wang, T., Huang, P., Li, D., Qiu, J., Shen, T., & Xie, P. (2015). Free language selection in the bilingual brain: An event-related fMRI study. *Scientific Reports*, *5*(1), 11704. https://doi.org/10.1038/srep11704

Zheng, X., Roelofs, A., & Lemhöfer, K. (2018). Language selection errors in switching: Language priming or cognitive control? *Language, Cognition and Neuroscience*, *33*(2), 139–147. https://doi.org/10.1080/23273798.2017.1363401

Zheng, X., Roelofs, A., & Lemhöfer, K. (2020). Language selection contributes to intrusion errors in speaking: Evidence from picture naming. *Bilingualism: Language and Cognition*, *23*(4), 788–800. https://doi.org/10.1017/S1366728919000683

Zoom Video Communications Inc. (2012). *Zoom* (5.10.1) [Computer software].

Appendices

Appendix A. Picture names in Dutch and English

Dutch target word	English target word
munt	coin
eend	duck
hond	dog
kip	chicken
kraan	tap
kikker	frog
mand	basket
fiets	bike
riem	belt
touw	rope
jurk	dress
spiegel	mirror
stoel	chair
ketting	chain
sleutel	key
boom	tree
mes	knife
fles	bottle
wolk	cloud
knoop	button
haai	shark
dak	roof
hek	fence
bezem	broom
slak	snail
been	leg
pijl	arrow
paard	horse
lepel	spoon
wortel	carrot

Appendix B. Mean and range of word variables of stimuli

	Dutch M (range)	English M (range)
SUBTLEX log10 frequency	2.9 (2.0–3.9)	3.1 (2.0–4.0)
Age of acquisition	5.3 (3.7–6.9)	4.6 (2.7–6.8)
Prevalence (1.5 = 93% knows word)	1.9 (1.7–2.0)	2.4 (2.2–2.6)
Concreteness (rating scale 1–5)	4.8 (4.4–5.0)	4.9 (4.3–5.0)

Appendix C. Overview of the experimental tasks and blocks

	Experiment version 1	Experiment version 2
Familiarization	Item familiarization L1 and L2	Item familiarization L1 and L2
1. Single-language naming		
Block 1	L1: Dutch (15 trials)	L2: English (15 trials)
Block 2	L2: English (15 trials)	L1: Dutch (15 trials)
Block 3	L1: Dutch (15 trials)	L2: English (15 trials)
Block 4	L2: English (15 trials)	L1: Dutch (15 trials)
2. Voluntary language switching		
Block 1	Free naming (30 trials)	Free naming (30 trials)
Block 2	Free naming (30 trials)	Free naming (30 trials)
3. Cued between-language switching		
Block 1	Cued switching (30 trials)	Cued switching (30 trials)
Block 2	Cued switching (30 trials)	Cued switching (30 trials)
4. Cued within-language switching		
Block 1	Cued switching L1 (30 trials)	Cued switching L1 (30 trials)
Block 2	Cued switching L1 (30 trials)	Cued switching L1 (30 trials)

Appendix D. The response categories with definitions and examples

Label (Accuracy)	Category	Definition/example (target word: *hond,* 'dog')
A (**1**)	Correct response	Identical to target word in target language: *hond*
B (**1**)	Hesitation	Hesitation before correct answer: *ehh… hond*
C (**0**)	No or late response	No answer within 3000 ms
D (**0**)	Selection error	Target word in wrong language: *dog* Wrong competing adjective in within-language switching task: *small* instead of *red* Wrong adjective: *blue* instead of *red*; *big* instead of small Both target and competing word produced: *dog…hond*
E (**0**)	Semantic error	Meaning-based lexical error/semantically related with target: *kat* ('cat')
F (**0**)	Phonological error	Phonological overlap (and no semantic relation) with at least 2/3 of target word: *rond*
G (**0**)	Unrelated error	Error with no phonological or semantic overlap with target: *tafel* ('table')
H (**0**)	False start	Repetition of the first syllable or phoneme: *ho- hond* Repetition of the first adjective: *klein- kleine hond* ('small- small dog') Pause between adjective and noun: *kleine…hond* ('small…dog')
I (**0**)	Wrong language Wrong word	Language intrusion and error: *cat, table.*
J (**0**)	Mix of two languages	Combination of phonemes from target word in both languages: *hog*
GLITCH	Glitch	Technical hiccup that rendered measuring RT impossible

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press

Appendix E. Model output for voluntary switching in the free switching task

	Switching (yes/no)				
Effect[a]	OR	*SE*	*LL 95% CI*	*UL 95% CI*	*p*
Intercept	0.67	0.07	0.55	0.82	**<.001**
Self-rated proficiency	1.28	0.13	1.04	1.57	**.021**
Ease of lexical access (Δ RT)	1.01	0.05	0.91	1.11	.796
Switch cost	0.95	0.10	0.78	1.16	.618

Note. Number of observations = 2117. [a]All predictors were scaled.

Appendix F. Model output for accuracy analyses of cued and free switching

	Switch model – Errors				
Predictors	*OR*	*SE*	*LL 95% CI*	*UL 95% CI*	*p*
Intercept	85.79	20.82	53.31	138.04	**<.001**
Trial type	1.59	0.33	1.06	2.39	**.025**
Task	0.56	0.12	0.37	0.84	**.005**
Language	1.28	0.27	0.85	1.92	.239
LexTALE	1.00	0.02	0.97	1.03	.866
Trial type:Task	1.04	0.22	0.69	1.56	.862
Trial type:Language	1.31	0.27	0.87	1.97	.198
Task:Language	0.70	0.15	0.47	1.06	.090
Trial type:LexTALE	0.97	0.01	0.95	1.00	.066
Task:LexTALE	1.03	0.02	1.00	1.06	**.040**
Language:LexTALE	1.00	0.02	0.97	1.029	.948
Trial type:Task:Language	0.68	0.14	0.45	1.03	.066
Trial type:Task:LexTALE	1.03	0.02	1.00	1.06	.066
Trial type:Language:LexTALE	0.98	0.02	0.95	1.01	.143
Task:Language:LexTALE	1.02	0.02	0.99	1.05	.319
Trial type:Task:Language:LexTALE	1.02	0.02	0.99	1.05	.194

Note. Number of observations = 4551.

Appendix G. Model output for RT analyses of cued and free switching

	Switch model - RTs[a]				
Effect	*Estimate*	*SE*	*LL 95% CI*	*UL 95% CI*	*p*
Intercept	6.88	0.03	6.83	6.94	**<.001**
Trial type	-0.02	0.003	-0.03	-0.02	**<.001**
Task	0.04	0.01	0.02	0.06	**<.001**
Language	-0.01	0.01	-0.03	-0.002	**.019**
LexTALE	-0.003	0.002	-0.01	0.002	.217
Switch cost	0.0004	0.0004	-0.0003	0.001	.316
Trial number	-0.0004	0.0002	-0.001	-0.00004	**.029**
Trial type:Task	-0.001	0.003	-0.006	0.004	.731
Trial type:Language	0.0003	0.003	-0.005	0.006	.919
Task:Language	-0.007	0.003	-0.01	-0.002	**.01**
Trial type:LexTALE	0.0004	0.0002	-0.0001	0.001	.107
Trial type:Switch cost	-0.00002	0.00004	-0.0001	0.0001	.613
Task:LexTALE	-0.001	0.001	-0.002	0.001	.316
Task:Switch cost	-0.00001	0.0001	-0.0002	0.0002	.933
Language:LexTALE	-0.0004	0.0004	-0.001	0.0004	.340
Language:Switch cost	0.0000	0.0001	-0.0001	0.0001	.938
Trial type:Task:Language	-0.01	0.003	-0.01	-0.001	**.029**
Trial type:Task:LexTALE	0.0001	0.0002	-0.0004	0.001	.791
Trial type:Task:Switch cost	0.00004	0.00004	-0.00003	0.0001	.297
Trial type:Language:LexTALE	-0.0002	0.0002	-0.001	0.0003	.506
Trial type:Language:Switch cost	-0.0001	0.0000	-0.0001	0.00002	.161
Task:Language:LexTALE	0.0004	0.0002	-0.0001	0.001	.110
Task:Language:Switch cost	-0.00004	0.00004	-0.0001	0.0000	.265
Trial type:Task:Language:LexTALE	0.001	0.0002	0.0001	0.0010	**.011**
Trial type:Task:Language:Switch cost	-0.00001	0.00004	-0.00001	0.00004	.771

Note. Number of observations = 4425. [a]RTs were (natural) log transformed.

Appendix H. Model output accuracy between-language and within-language switching

	Cued model – Errors				
Effect	*OR*	*SE*	*LL 95% CI*	*UL 95% CI*	*p*
Intercept	24.67	4.20	17.97	34.74	**<.001**
Trial type	1.46	0.11	1.25	1.69	**<.001**
Task	2.03	0.16	1.74	2.36	**<.001**
Trial type:Task	1.14	0.09	0.98	1.32	.099

Note. Number of observations = 4491.

Appendix I. Model output RTs between-language and within-language switching

	Cued model - RTs[a]				
Effect	*Estimates*	*SE*	*LL 95% CI*	*UL 95% CI*	*p*
Intercept	6.96	0.03	6.89	7.02	**<.001**
Trial type	-0.02	0.003	-0.02	-0.01	**<.001**
Task	-0.06	0.01	-0.08	-0.05	**<.001**
Trial number	-0.02	0.01	-0.04	-0.01	**.001**
Trial type:Task	-0.002	0.003	-0.01	0.004	.440

Note. Number of observations = 4128. [a]RTs were (natural) log transformed.

https://doi.org/10.1017/S1366728923000755 Published online by Cambridge University Press