Bilingual control: Sequential memory in language switching

Mathieu Declerck, Andrea M. Philipp and Iring Koch

Institute of Psychology, RWTH Aachen University, Aachen, Germany

This article was accepted in the Journal of Experimental Psychology: Learning, Memory, and Cognition. This article may not exactly represent the final published version. It is not the copy of record.

Authors' Note

Mathieu Declerck, Andrea M. Philipp and Iring Koch, Institute of Psychology, RWTH Aachen University, Aachen, Germany.

The authors would like to thank Johanna Josten and Aniella Thoma, for conducting part of the experiments, and Jared Linck, Allisa Melinger and two anonymous reviewers for helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Mathieu Declerck, RWTH Aachen University, Institute of Psychology, Jägerstrasse 17-19, 52066 Aachen, Germany. Email: Mathieu.Declerck@psych.rwth-aachen.de.

Abstract

To investigate bilingual language control prior language switching studies presented visual objects, which had to be named in different languages, typically indicated by a visual cue. The present study examined language switching of predictable responses by introducing a novel sequence-based language switching paradigm. In four experiments, sequential responses (i.e., weekdays, numbers or new sequences) and an alternating language sequence (L1-L1-L2-L2 etc.) were implemented, both of which were memory-based. Our data revealed switch costs, showing that a language switch is associated with worse performance compared to a language repetition, and mixing costs, which constitutes the performance difference between pure and mixed language blocks, even while producing entirely predictable responses (i.e., language and concept). Additionally, we found these switch costs with over-learned and new sequences and that switch costs were reduced with longer preparation time. The obtained data are consistent with a proactive interference account, such as the inhibitory control model.

(149 words)

Due to globalization, bilingualism has become an important topic and is gaining more and more interest, even with respect to the influence of bilingualism on non-language domains (for a review, see Bialystok, 2009). One topic in the bilingual field concerns the bilingual ability of language production in either of the two languages, which relies on language control processes. These cognitive control processes allow bilinguals to adaptively change information processing and behavior based on current goals (Monsell, 1996).

In the current study we set out to examine the influence of both language and concept predictability on language control. More specifically, response predictability was investigated during language switching, which is a popular task to investigate language control processes (e.g., Christoffels, Firk & Schiller, 2007; Green, 1998; Philipp & Koch, 2009; for reviews on language control, see Abutalebi & Green, 2007, 2008; Kroll, Bobb, Misra & Guo, 2008). Next to the empirical implications, the results of this study would also give an indication towards the functional locus of language control.

Language switching paradigms

Typical language switching studies (i.e., cued language switching; e.g., Costa & Santesteban, 2004; Declerck, Koch & Philipp, 2012; Meuter & Allport, 1999; Philipp, Gade & Koch, 2007) are characterized by their use of visual stimuli, such as pictures or digits, as a way to let participants name a pre-determined concept in their mother language (L1) or second language (L2; see top panel of Figure 1). The required language is typically indicated by presenting a visual language cue that precedes the to-be-named stimulus, or is presented simultaneously with the to-be-named stimulus.

---- Please insert Figure 1 about here ----

Whereas other language switching paradigms also use visual stimuli as a way to let bilinguals name a pre-determined concept, they implement language transitions differently from cued language switching. One approach is voluntary language switching, which requires bilingual participants to choose which language to produce on each trial (Gollan & Ferreira, 2009). This entails that no visual language cues are required. Hence, in the middle panel of Figure 1 there is no need for a visual language cue to determine the language. An alternative to this approach is the alternating language sequence (e.g., Festman, Rodriguez-Fornells & Münte, 2010; Jackson, Swainson, Mullin, Cunnington & Jackson, 2001, 2004), which uses a language sequence that changes language after every second trial (e.g., L1-L1-L2-L2 etc.). The latter set-up is similar to predictable task switching (e.g., Koch, 2003; Rogers & Monsell, 1995).

In the present article, we propose a novel language switching procedure. However, before introducing this procedure, we first describe the most relevant language switching phenomena and major models of cognitive control.

Markers of cognitive control in language switching

Language switching allows for the investigation of several markers of cognitive control, such as mixing costs and switch costs. Language mixing costs represent the performance difference between pure language responses and mixed language responses, with the typical pattern showing better performance during pure language than during mixed language responses (e.g., Christoffels, Firk & Schiller, 2007; Wang, Kuhl, Chen & Dong, 2009). This particular performance cost is considered to be a marker of sustained control processes, primarily involved in interference resolution (for reviews see Los, 1996; Kiesel, Wendt, Jost, Steinhauser, Falkenstein, Philipp & Koch, 2010).

Mixed language blocks generally consist of (at least) two languages, which results in two transitions between trials: one language succeeds the other (switch trials) or the same language has to be repeated (repetition trials). Generally it is harder to switch between languages than to repeat the same language, resulting in switch costs (e.g., Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Declerck, Koch & Philipp, 2012; Meuter & Allport, 1999; Philipp, Gade & Koch, 2007; Verhoef, Roelofs & Chwilla, 2009).

Unlike mixing costs, switch costs are a marker of transient control processes involved in interference resolution and carry-over effects (for a review see Allport & Wylie, 1999).

An additional finding with both mixing costs (e.g., Christoffels, Firk & Schiller, 2007) and switch costs (e.g., Meuter & Allport, 1999; Philipp, Gade & Koch, 2007) concerns an asymmetric performance difference across languages, with larger L1 performance costs than L2 performance costs in unbalanced bilinguals. Yet, differences in language asymmetry have also been found between switch costs and mixing costs in bilinguals (Christoffels, Firk & Schiller, 2007; Wang, Kuhl, Chen & Dong, 2009), indicating that these two markers of cognitive control are influenced differently by the first and second language. Based on these robust empirical effects in language switching, we now discuss major theoretical accounts of cognitive control processes as required in language switching.

Models of cognitive control

The current study will primarily focus on language switch costs. There are two influential models, which are derived from the task switching domain (for a review of both models in task switching, see Kiesel et al. , 2010; Vandierendonck, Liefooghe & Verbruggen, 2010), that have specified the possible underlying mechanisms of switch costs.

The first is the reconfiguration model (e.g., Rogers & Monsell, 1995). This model assumes that switch costs reflect the time needed to reconfigure the cognitive system from one task to another. An important assumption of this model is that, given substantial preparation time, switch costs should disappear. However, since Rogers and Monsell (1995) still found substantial switch costs with a long task preparation (i.e., “residual” switch costs), they assumed that reconfiguration occurs during two stages. During the first stage the task will either stay the same (repetition trial) or will need to be reconfigured to the parameters of the new task (switch trials). This first reconfiguration can be completed before the presentation of the stimulus. A second reconfiguration stage follows when the stimulus is presented and the task rules need to be activated (see also Rubenstein, Meyer & Evans, 2001).

The second account is the proactive interference model (Allport, Styles & Hsieh, 1994). In this model, activation of the previously used task or language persists and thus causes either interference with the current task (switch trials) or causes facilitation based on residual activation (repetition trials). According to Green's Inhibitory Control Model (ICM, 1998), proactive interference during bilingual language production mainly consists of persisting inhibition. Evidence for this account has been found in studies that investigated inhibition processes during language switching (e.g., Linck, Schwieter & Sunderman, 2012; Philipp & Koch, 2009).

Preparation effects in language switching

As indicated above, the reconfiguration model (Rogers & Monsell, 1995) assumes that preparation plays an important role with respect to switch costs. Several cued language switching studies have investigated the effect of preparation time on language switching (Costa & Santesteban, 2004; Macnamara, Krauthammer & Bolgar, 1968; Philipp, Gade & Koch, 2007; Verhoef, Roelofs & Chwilla, 2009). However, over these studies there was no converging pattern when it comes to the effect of preparation time on language switch costs. Whereas Costa and Santesteban (2004) and Verhoef, Roelofs and Chwilla (2009) found that switch costs were larger when preparation time decreased, the data of Philipp, Gade and Koch (2007) and Macnamara, Krauthammer and Bolgar (1968) did not confirm this effect.

Preparation of languages in language switching is also possible when the language sequence is predictable, such as during voluntary language switching (Gollan & Ferreira, 2009) and alternating language sequence (e.g., Festman, Rodriguez-Fornells & Münte, 2010; Jackson, Swainson, Mullin, Cunnington & Jackson, 2001, 2004). These language switching studies demonstrate that switch costs appear even when the language switches can be prepared due to a predictable language sequence. However, these studies did not specifically investigate the effect of language predictability on language switching (e.g. by contrasting it with performance of unpredictable switches, see Koch, 2005).

Furthermore, no study has looked into the influence of concept preparation in language switching or the combination of both language and concept preparation and thus response preparation in language switching. Response preparation is an important characteristic in language production since preplanning processes are also active during sentence production (e.g., Martin, Crowther, Knight, Tamborello & Yang, 2010; Oppermann, Jescheniak & Schriefers, 2010).

The ability to prepare upcoming responses could also have a major impact on language switching, since being able to not just prepare for the upcoming concept but also for either a language repetition or language switch should greatly diminish, or even abolish, switch costs according to the reconfiguration model (Rogers & Monsell, 1995). This is because both the first reconfiguration stage (language) and the second (concept) could be conducted in advance, so that even residual costs might be abolished. The proactive interference model (Allport, Styles & Hsieh, 1994; Green, 1998), on the other hand, makes no such claim. This entails that the proactive interference model could account for switch costs with fully predictable responses.

In the current study, we set out to investigate a novel language switching paradigm, the sequence-based language switching paradigm (SBLS). This paradigm would allow us to investigate the effect of predictable responses (both language and concept) on language control.

The sequence-based language switching paradigm

To investigate predictable responses in language switching, we chose responses from both pre-determined (i.e., weekdays and numbers) and new sequences while using an alternating language sequence (e.g., L1-L1-L2-L2-L1-L1 etc.). The task consisted of producing one response, based on a sequence of concepts and languages, after hearing an auditory response-signal. Additionally, since the SBLS paradigm relies on a sequence of concepts and an alternating language sequence, both the concepts and languages are triggered from memory, which entails that no visual stimuli or cues were used. Accordingly, in the lower panel of Figure 1, which depicts a trial with the SBLS paradigm, there are no visual triggers (i.e., cue or stimulus) to produce a response.

By using a fixed sequence of responses (e.g., weekdays or numbers), we come closer to natural speech production since language is generally produced in a sequence (e.g., Acheson & MacDonald, 2009; Dell, Burger & Svec, 1997). Furthermore, by using no visual stimuli, the SBLS paradigm allows for the investigation of endogenous, memory-based response selection. This is important since during normal bilingual speech, response selection is for the most part endogenous, memory-based, whereas during typical language switching experiments responses are exclusively exogenously triggered by visually presented pictures or digits (e.g., Costa & Santesteban, 2004; Declerck, Koch & Philipp, 2012; Meuter & Allport, 1999).

In the present study we report four experiments using the SBLS paradigm. Experiment 1 was set up to investigate whether predictable responses would have an effect on switch costs. This experiment entailed a memory-based concept sequence (i.e., weekdays) and language sequence with long and short preparation times. Smaller switch costs with a long pacing-interval, which constitutes the time between the previous response onset and the current response-signal, than with a short pacing-interval would indicate that there was a preparation benefit on switch costs due to language predictability, concept predictability or both being predictable. This is because the participants had more time to prepare for the upcoming trial using the predicable information (i.e., concept and/or language) with a long pacing-interval. Next to the predictable responses in Experiment 1, which allowed for preparation until response selection, a predictable response onset was added in Experiment 2 to investigate whether switch costs would arise when preparation was possible until response execution. Additionally, in this experiment we also investigated whether, next to switch costs, mixing costs could be elicited by predictable responses and the effect of another semantic category (i.e., numbers). Since the concept sequences in Experiments 1 and 2 were over-learned sequences (i.e., weekdays and numbers), we scrambled the items of these sequences into a new sequence in Experiment 3. This scrambled sequence, which was not over learned, was also contrasted against the over-learned standard sequence to investigate the effect of sequence novelty. Finally, in Experiment 4, we aimed at investigating whether switch costs would be elicited when the predictable responses were not part of a pre-determined sequence but a sequence of unrelated words, since responses that can be produced in a pre-determined sequence contain semantic and phonologic properties that are not common among all words. The results of these four experiments are discussed in light of the reconfiguration model (e.g., Rogers & Monsell, 1995; Rubinstein, Meyer & Evans, 2001; see Vandierendonck, Liefooghe & Verbruggen, 2010) and the proactive interference model (Allport, Styles & Hsieh, 1994; Green, 1998; see Kiesel et al., 2010).

Experiment 1

In Experiment 1, participants performed a sequential memory-based task with weekdays. The aim of the experiment was two-fold. One aim was to examine whether language switch costs would be elicited by responses that are predictable and triggered from memory (instead of being unpredictable and visually triggered).

The second aim of Experiment 1 was to investigate whether the preparation of predictable language switches and responses would have an impact on switch costs. If this was the case, longer preparation should decrease switch costs, since the participants can use the predictable information (i.e., concept and language) during a long pacing-interval to prepare for the upcoming response. This is not the case, or at least less so, with a short pacing-interval, due to reduced time to prepare for the upcoming response.

Method

Participants. 24 native German participants took part and spoke English as their second language (20 female, mean age = 22.5). Prior to the experiment they were asked to fill in a questionnaire about how old they were and when they started learning English, how many years of formal English education they had, how many other languages they know and how high they rated their own level of spoken English, with 1 being very bad and 7 being very good (see Table 1).

Apparatus and concepts. The to-be-produced concepts consisted of the seven weekdays, which the participants were required to produce from memory in the appropriate serial order (Monday to Sunday).

The trials were presented and the responses recorded using E-prime version 1.1.4.1. Speech onset of vocal responses was recorded with a voice-key and the entire experiment was recorded with a Zoom H2 Handy Portable Stereo Recorder. Errors were coded online by the experimenter in a subject file. The recorded speech files were consulted for accuracy.

Procedure. Prior to the experiment, the instructions were presented both orally and visually, with an emphasis on both speed and accuracy. This was followed by two practice blocks of 14 trials each and eight experimental blocks of 28 trials each. At the beginning of each block the participants were informed about the characteristics of the block (i.e., which language they should begin with and whether the pacing-interval would be long or short). This was followed by a fixation cross (+), presented in the centre of the screen, which stayed visible throughout the entire block.

After hearing an auditory response-signal (50 ms), the participants had to produce one of the seven weekdays from memory. They were required to produce the weekdays in the correct serial order (Monday-Sunday; see appendix for all responses), starting with the concept *Monday*, as well as in the correct language (German or English). The required language alternated after every second trial, on the basis of an alternating language sequence (i.e., L1-L1-L2-L2-L1-L1 or L2-L2-L1-L1-L2-L2). Note that repeating the weekday sequence four times in each block and the language sequence requiring a language switch after every second trial results in perfect counter-balancing of language, language sequence and serial position. That is, each weekday was named equally often in each language and equally frequent on repetition trials and on switch trials in each block.

To measure the effect of preparation time, the pacing-interval, constituting the time between the previous response onset and the current response-signal, was varied blockwise. Additionally, a random jitter of 200 ms was used so that the participants could not automate responding based on fixed timing. This resulted in blockwise short pacing-intervals (mean of 1100 ms; 900, 1100 or 1300 ms) and blockwise long pacing-intervals (mean of 2000 ms; 1800, 2000 or 2200 ms). Overall, short pacing-intervals were used during half of the blocks, while long pacing-intervals were used during the other half. Also, half of the blocks started with German and half started with English. The starting language of each block was altered after every second block, whereas the pacing-interval altered from one block to the next. The sequence of blocks was counterbalanced across participants.

Design. The dependent variables were reaction time (RT) and error rate. The within-subjects independent variables were language (German vs. English), language sequence (switch vs. repetition), and preparation time (long vs. short interval).

Results and Discussion

---- Please insert Table 2 about here ----

The first trial of each block and the error trials, which constituted the production of a wrong concept and/or production in the wrong language, were excluded from RT analyses. Furthermore, RTs in all trials were z-transformed and trials with a z-score of -2/+2 were discarded as outliers. Taking these three criteria into account, a total of 14% of the data was excluded.

An analysis of variance (ANOVA) of the RT data revealed a significant effect of language (F(1, 23) = 6.64; $p < .05$; $\eta_p^2 = .224$), with German responses (512 ms) being slower than English responses (500 ms, see Table 2), and of language sequence (F(1, 23) = 26.47; $p < .001$; $\eta_p^2 = .535$), with switch trial responses (532 ms) being slower than repetition trial responses (480 ms), indicating language switch costs of 52 ms. These data showed that substantial switch costs arise using sequential, memory-based concepts and language sequences, and thus predictable responses. This finding indicates that trial-to-trial language control processes, measured by switch costs, become necessary despite the possibility to prepare the upcoming responses. Hence, responses are not fully prepared, even though they are fully predictable. It also provides evidence that even when words are triggered from memory, instead of being visually triggered, switch costs arise.

The main effect of preparation time was not significant ($F(1, 23) = 3.06$; ns.; $\eta_p^2 = .117$), even though the data showed a trend towards slower responses with a short interval (526 ms) than with a long interval (486 ms). However, the interaction between language sequence and preparation time was significant ($F(1, 23) = 9.09$; $p < .01$; $\eta_p^2 = .283$), with larger switch costs for short intervals (76 ms) than for long intervals (37 ms). Separate t-tests revealed that switch costs were significant with both the short interval ($t(23) = 4.34$; $p < .001$) and the long interval ($t(23) = 4.43$; $p < .001$). All other interactions were not significant (Fs < 1). The reduction of switch costs with a long interval supports our hypothesis that there was a preparation benefit for the upcoming responses, since during the long pacing-intervals the participants had more time to use the predictability information (i.e., concept and/or language) to prepare for the upcoming trial.

The error data revealed no significant main effects of language ($F(1, 23) = 0.34$; ns.; $\eta_p^2 = .015$) and of preparation time ($F(1, 23) = 3.37$; ns.; $\eta_p^2 = .128$). However, preparation time did reveal a trend with more errors during short intervals (0.6%) than during long intervals (0.4%). There was a main effect of language sequence ($F(1, 23) = 11.40$; $p < .01$; $\eta_p^2 = .331$), with switch trial responses (0.7%) being more erroneous than repetition trial responses (0.4%). The interaction between language sequence and preparation time was not significant ($F(1, 23) = 1.40$; ns.; $\eta_p^2 = .058$), nor any of the other interactions (Fs < 1).

Experiment 2

The results of Experiment 1 revealed that, even with a long preparation time, switch costs could be elicited with predictable responses. However, preparation in Experiment 1 was only possible up to response selection, and not response execution, since the jitter prevented the participants from knowing when exactly to produce the next response. Since we found no hint of participants using automated strategies with respect to response onset in Experiment 1, we did not implement a jitter in Experiment 2, which would allow the participants to prepare for response execution as well.

Additionally, we implemented another marker of cognitive control, namely mixing costs, to investigate whether the effect of predictable responses with a predictable response onset elicits these performance costs. We chose mixing costs because previous studies have provided evidence that mixing costs measure a different aspect of language control than switch costs (Christoffels, Firk & Schiller, 2007; Wang, Kuhl, Chen & Dong, 2009).

Furthermore, in this experiment the participants performed a sequential memory-based task with either weekdays or numbers, in order to generalize our findings to other pre-determined sequences.

Method

Participants. 48 native German participants took part and spoke English as their second language (39 female, mean age = 22.8). A questionnaire, identical to that in Experiment 1, was given to the participants prior to the actual experiment (see Table 1).

Apparatus and concepts. The apparatus was identical to that used in Experiment 1. However, the to-be-produced concepts consisted either of the seven weekdays or numbers 1-7, which the participants were required to produce from memory in the appropriate serial order (Monday to Sunday; 1 to 7). Half the participants had to produce weekdays, whereas the other half had to produce the numbers 1-7.

Procedure. The procedure was almost identical to that used in Experiment 1. Among the differences was the addition of pure language blocks, which lead to four different block types: pure English blocks, pure German blocks and mixed blocks with the first trial being either in German or in English. Secondly, there were four practice blocks, each presented prior to the first experimental block of the corresponding block type, instead of before all experimental blocks, which was the case in Experiment 1. Each condition was presented in two blocks during the experiment, excluding the practice blocks.

The sequence of blocks consisted of two pure language blocks, one English and one German, followed by four mixed blocks and then another German and English pure language block. The starting language of the four mixed blocks was altered from one block to the next. This was counterbalanced across participants, as was the sequence of pure language blocks. Finally, the pacing-interval was fixed at 1500 ms, without a jitter.

Design. The dependent variables were RT and error rate. The between-subject variable was semantic category (weekdays vs. numbers) and the within-subjects independent variables were language (German vs. English) and, in the *switch-cost contrast*, language sequence (switch vs. repetition), whereas in the *mixing-cost contrast* this was type of block (pure vs. mixed language block).

Results and Discussion

---- Please insert Table 3 about here ----

We used identical outlier criteria and error definitions as in Experiment 1, which resulted in the exclusion of 8% of the data. Furthermore, when analyzing mixing costs, we only used the repetition trials of the mixed language blocks to have a clear distinction between switch costs and mixing costs.

Switch-cost contrast: Switch trials vs. repetition trials in mixed blocks. The corresponding ANOVA of the RT data revealed a significant main effect of language ($F(1, 46) = 11.13$; $p < .01$; $\eta_p^2 = .195$), with German responses (407 ms) being slower than English responses (396 ms, see Table 3), whereas the main effect of semantic category ($F(1, 46) = 1.48$; ns.; $\eta_p^2 = .031$) was not significant. Language sequence was also significant ($F(1, 46) =$

36.59; $p < .001$; $\eta_p^2 = .443$), with switch trial responses (414 ms) being slower than repetition trial responses (389 ms), indicating language switch costs of 25 ms. The switch costs suggest that substantial interference arises even when the response, language sequence and response onset are predictable, hence with predictable response execution.

The interaction between language and language sequence was significant ($F(1, 46) = 6.71$; $p < .05$; $\eta_p^2 = .127$), with larger switch costs for German responses (34 ms) than English responses (17 ms). Likewise, the interaction between language sequence and semantic category was significant ($F(1, 46) = 9.16$; $p < .01$; $\eta_p^2 = .166$), with larger switch costs for weekdays (37 ms) than numbers (12 ms). Separate t-tests revealed that switch costs were significant for both the weekday data ($t(23) = 5.59$; $p < .001$) and the number data ($t(23) = 2.59$; $p < .05$). The interaction between language and semantic category and the three-way interaction were not significant (Fs < 1). These findings show that asymmetrical switch costs can be found with predictable responses and that different semantic categories influence the size of switch costs. Furthermore, the data provided evidence that switch costs can be obtained with pre-determined sequences other than weekdays (i.e., numbers).

The error data revealed no significant main effect of language and semantic category (Fs < 1). There was a significant main effect of language sequence ($F(1, 46) = 15.42$; $p < .001$; $\eta_p^2 = .251$), with switch trial responses (0.5%) being more erroneous than repetition trial responses (0.2%). None of the interactions were significant (Fs < 1).

Mixing-cost contrast: Pure language blocks vs. repetitions in mixed language blocks. An ANOVA of the RT data revealed no significant main effects of language ($F(1, 46) = 2.68$; ns.; $\eta_p^2 = .055$) and semantic category ($F(1, 46) = 1.74$; ns.; $\eta_p^2 = .036$). Type of block was significant ($F(1, 46) = 12.69$; $p < .001$; $\eta_p^2 = .216$), with responses in mixed language blocks (389 ms) being slower than in pure language blocks (357 ms), indicating language mixing costs of 32 ms. That is, mixing costs can be observed even when the responses, language sequence and response onset are predictable. The interaction between type of block and semantic category ($F(1, 46) = 2.61$; ns.; $\eta_p^2 = .054$) was not significant. Similarly, all the other interactions (*F*s < 1) were not significant.

The error data revealed no significant main effects of language, semantic category and type of block (*F*s < 1). Also the interactions between type of block and language ($F(1, 46) = 2.05$; ns.; $\eta_p^2 = .043$), type of block and semantic category ($F < 1$), language and semantic category ($F < 1$), and the three-way interaction ($F(1, 46) = 1.15$; ns.; $\eta_p^2 = .024$) were not significant.

Experiment 2 demonstrates that both switch costs and mixing costs occur with the SBLS paradigm. Further, the experiment allows us to distinguish between language switch costs and language mixing costs. The data showed a difference between these two markers of cognitive control on two accounts. The first one being that, similar to the results reported by Christoffels, Firk and Schiller (2007) and Wang, Kuhl, Chen and Dong (2009), the current data reveals a difference in (a)symmetrical performance costs across languages, with asymmetrical switch costs (i.e., larger L1 performance costs than L2, see also Meuter & Allport, 1999) and symmetrical mixing costs. An explanation for the dissociation between switch costs and mixing costs across languages is presently difficult due to an unclear pattern of results across studies. Whereas Wang, Kuhl, Chen and Dong (2009) found a similar pattern as the present study, with asymmetrical switch costs and symmetrical mixing costs, the data of Christoffels, Firk and Schiller (2007) showed the opposite, with symmetrical switch costs and asymmetrical mixing costs. Yet, all three studies, including the present study, provide evidence that switch costs and mixing costs are influenced differently by languages.

The data also revealed a difference between switch costs and mixing costs due to semantic category: larger switch costs were elicited while producing weekdays than numbers. The numerical mixing cost data showed that mixing costs were larger when producing numbers than during the production of weekdays (mixing costs weekday: 18 ms and numbers: 47 ms). This difference across performance costs could be due to the phonological difference between the two semantic categories (e.g., the phonological priming effect of weekdays, which is not present between numbers), since the data of Christoffels, Firk and Schiller (2007) also revealed a similar pattern between switch costs and mixing costs due to phonology (i.e., non-cognates vs. cognates; cognates are words with a similar etymological background in two or more languages, which often co-occur with a large phonological overlap; Costa, Caramazza & Sebastián-Gallés; 2000; Declerck, Koch & Philipp, 2012).

Experiment 3

So far we have only implemented over-learned sequences (i.e., weekdays and numbers). In Experiment 3 we expanded our investigation to a scrambled sequence (i.e., weekdays or numbers in a new order) to generalize our findings to new sequences. To explore the influence of sequence novelty we investigated the difference between producing an over-learned sequence and a scrambled sequence.

Method

Participants. 24 native German participants took part and spoke English as their second language (13 female, mean age = 25.3). A questionnaire, identical to that in Experiment 1, was given to the participants prior to the actual experiment (see Table 1).

Apparatus and concepts. The apparatus was identical to those used in the previous experiments. However, the participants were instructed to produce the numbers 1-5 and weekdays Monday – Friday. Either one was used in the over-learned condition, while the other was used in the scrambled condition. This pairing of semantic category to sequence condition was counterbalanced across participants. In the over-learned condition, the normal sequential order was used, whereas in the scrambled condition a mixed sequence of the same concepts was used (see appendix for all response sequences).

Procedure. The procedure was identical to that used in Experiment 1, apart from some points. First, each block consisted of twenty trials and there were four experimental blocks for both the over-learned sequence and scrambled sequence condition. All four blocks, using the same sequence, were presented in succession. Secondly, there were two practice blocks to practice each sequence, of 20 trials each, which were presented before the four experimental blocks that would use that particular sequence. During the first of these practice blocks the participants had the responses in written form, in both languages and in the correct order, in front of them. In the second practice block the participants had to perform the task without the written responses. The order of the conditions, represented by four experimental blocks and two practice blocks, were counterbalanced across participants. Finally, the pacing-interval was fixed at 1500 ms, without a jitter.

Design. The dependent variables were RT and error rate. The within-subjects independent variables were language (German vs. English), language sequence (switch vs. repetition), and sequence condition (over-learned vs. scrambled sequence).

Results and Discussion

---- Please insert Table 4 about here ----

We used identical outlier criteria and error definitions as in Experiment 1, which resulted in the exclusion of 16% of the data. An ANOVA of the RT data revealed a significant effect of sequence condition ($F(1, 23) = 9.99$; $p < .01$; $\eta_p^2 = .303$), with responses in the scrambled sequence (659 ms) being slower than in the over-learned sequence (558 ms, see Table 4), whereas language was not significant ($F < 1$). However, language sequence was clearly significant ($F(1, 23) = 11.06$; $p < .01$; $\eta_p^2 = .325$), with switch trial responses (634 ms) being slower than repetition trial responses (583 ms), indicating language switch costs of 51 ms.

The interactions between sequence condition and language ($F(1, 23) = 2.02$; ns.; $\eta_p^2 = .081$), language and language sequence ($F < 1$) and the three-way interaction ($F < 1$) were not significant. Importantly, the interaction between sequence condition and language sequence was significant ($F(1, 23) = 5.53$; $p < .05$; $\eta_p^2 = .194$), with larger switch costs for responses in the scrambled sequence (69 ms) than in the over-learned sequence (32 ms). Separate t-test revealed that both the scrambled sequence responses ($t(23) = 3.19$; $p < .01$) and the over-learned sequence responses ($t(23) = 2.95$; $p < .01$) revealed significant switch costs. This provides evidence that producing predictable responses elicits switch costs in both over-learned and scrambled sequences.

Presumably, the additional strain put on working memory (e.g., Barrouillet, Bernardin, & Camos, 2004; Barrouillet, Bernardin, Portrat, Vergauwe & Camos, 2007; Portrat, Barrouillet & Camos, 2008) in the scrambled sequence caused the larger switch costs relative to the over-learned sequence. Contrary to the over-learned condition, the scrambled sequences require the novel concept sequence to be maintained in working memory, whereas the concept sequence in the over-learned condition is part of the information retrieved from long-term memory. This would imply that more working memory load was implemented during scrambled sequences. Hence, we assume that switch costs are influenced by working memory load, which is backed-up by recent task-switching studies (e.g., Liefooghe, Barrouillet, Vandierendonck & Camos, 2008).

The error data revealed a significant main effect of language sequence ($F(1, 23) = 14.80$; $p < .001$; $\eta_p^2 = .392$), with switch trials (0.9%) being more erroneous than repetition trials (0.3%), and of sequence condition ($F(1, 23) = 5.13$; $p < .05$; $\eta_p^2 = .182$), with more errors generated in the scrambled sequence (0.8%) than in the over-learned sequence (0.4%), whereas the main effect of language ($F(1, 23) = 2.67$; ns.; $\eta_p^2 = .104$) was not significant. The interaction between language sequence and sequence condition ($F(1, 46) = 1.12$; ns.; $\eta_p^2 = .046$) was not significant and neither were any of the other interactions (*F*s < 1).

Experiment 4

Since the responses used in over-learned sequences have semantic and phonologic properties that are not applicable to all words, Experiment 4 set out to investigate language switching with predictable responses using unrelated concepts in a novel sequence. In Experiment 3, the scrambled sequence was also not over learned, but the responses in

Experiment 3 were still part of pre-determined sequences. Thus, additional interference could have come into play because of sequential priming from the over-learned sequence (e.g., numerical distance priming when using numbers, Duyck, Depestel, Fias & Reynvoet, 2008). This was excluded in Experiment 4, since the concepts were unrelated.

Using unrelated words also has the benefit of not including semantic influences, like the words being part of one semantic category, and phonologic influences, like the high amount of cognates and phonological priming (every weekday in English ends with *–day* and almost every weekday in German ends with *–tag*).

To further reduce the amount of phonological priming from one trial to the next, Experiment 4 contains sequences in which every word (both German and English responses) contains at least one language-specific phoneme, which are phonemes that do not appear in the other language. This is then contrasted against sequences that contain words without any language-specific phonemes.

Method

Participants. 24 native German participants took part and spoke English as their second language (16 female, mean age = 23.0). A questionnaire, identical to that in Experiment 1, was given to the participants prior to the actual experiment (see Table 1).

Apparatus and concepts. The apparatus was identical to those used in Experiments 1-3. With respect to the concepts, there were twenty concepts which had to be produced from memory in the correct serial order. Each of these concepts was used in one of four sequences, two with a language-specific phonology word set and two with a language-unspecific phonology word set, each of which contained five concepts (see appendix for the full sets). The four sequences were set up so that there was as little phonological overlap from one trial to the next, no semantic relationship between concepts, and as few cognates as possible.

Each concept in the two language-specific phonology sequences had at least one language-specific phoneme in the German response and English response. The German-specific phonemes were: /ʦ/ (e.g.: Ka*tz*e, meaning cat); /ʏ/ (e.g., R*ü*cken, meaning back); / χ / (e.g., Tu*ch*, meaning cloth) and /n̩/ (e.g., Rach*en*, meaning throat), whereas the English-specific phonemes were: /æ/ (e.g., c*a*t); /ʌ/ (e.g., m*u*shroom) and / θ / (e.g., *th*roat).

Furthermore, we controlled the words of the four sequences on frequency (Baayen, Piepenbrock & Gulikers, 1995) and the amount of syllables over sequences. The two language-specific phonology sequences were also controlled for the amount of language-specific German and English phonemes.

Procedure. The procedure was identical to that used in Experiment 1, apart from the following features: There were sixteen experimental blocks, four for each sequence. These four blocks that used the same sequence were always presented one after the other. To get acquainted with a new sequence there were two practice blocks of twenty trials each prior to the four experimental blocks. During the first practice block the participants would have a card in front of them with the written responses on it, in both languages and in the correct order, while during the second practice block they would have to respond without the card. Finally, the pacing-interval was fixed at 1500 ms, without a jitter.

Design. The dependent variables were RT and error rate. The within-subjects independent variables were language (German vs. English), language sequence (switch vs. repetition), and phonology (language-specific vs. language-unspecific phonology).

Results and Discussion

---- Please insert Table 5 about here ----

We used identical outlier criteria and error definitions as in Experiment 1, which resulted in the exclusion of 14% of the data. An ANOVA of the RT data revealed a significant main effect of language ($F(1, 23) = 4.92$; $p < .05$; $\eta_p^2 = .176$), with German responses (720 ms) being slower than English responses (700 ms, see Table 5). Phonology ($F(1, 23) = 10.80$; $p < .01$; $\eta_p^2 = .320$) was also significant, with responses in the language-specific phonology condition (754 ms) being slower than responses in the language-unspecific phonology condition (666 ms). The main effect of language sequence was significant ($F(1, 23) = 14.09$; $p < .001$; $\eta_p^2 = .380$), with switch trial responses (745 ms) being slower than repetition trial responses (675 ms), indicating language switch costs of 70 ms. These switch costs show that even without the semantic factors, such as the influence of semantically-related words, without phonological factors, such as the high amount of cognates and the phonological priming from one trial to the next, and without sequential priming, that might have influenced Experiments 1-3, switch costs were elicited with predictable responses. Also, for the first time we demonstrate the effect in a completely novel sequence with unrelated concepts. This finding is important for future experiments that want to investigate concepts different from those used in pre-determined sequences.

All interactions had an *F*-value smaller than 1, apart from the significant two-way interaction between language and phonology ($F(1, 23) = 4.28$; $p < .05$; $\eta_p^2 = .157$), with the responses in the language-specific phonology condition eliciting faster responses in German than English (51 ms) and the responses in the language-unspecific phonology condition eliciting faster responses in English than German (11 ms). This could explain why the data in Experiments 1-3 revealed slower German responses than English responses: Both weekdays and numbers consist of a large number of cognates, which are phonologically similar in both languages and thus have a small amount of language-specific phonemes.

An ANOVA of the error data revealed no significant main effects of language ($F < 1$) and phonology ($F(1, 23) = 1.67$; ns.; $\eta_p^2 = .068$). The data did show a significant main effect of language sequence ($F(1, 23) = 7.65$; $p < .05$; $\eta_p^2 = .250$), with a higher occurrence of errors in switch trials (0.7%) than repetition trials (0.5%). All interactions had an *F*-value smaller than 1, apart from the two-way interaction between phonology and language ($F(1, 23) = 1.61$; ns.; $\eta_p^2 = .065$) and between phonology and language sequence ($F(1, 23) = 2.68$; ns.; $\eta_p^2 = .104$), both of which were not significant.

General Discussion

In the current study we set out to explore language control during the production of predictable sequences of responses. A novel paradigm, the SBLS paradigm, was implemented to investigate this effect. This paradigm consists of two memory-based features: a memory-based concept sequence and a memory-based alternating language sequence. Both of these features were not triggered exogenously through language cues or stimuli, as in most language switching studies, but endogenously.

The data of Experiment 1 showed that language switch costs emerged in the SBLS paradigm and were influenced by the pacing-interval. Specifically, switch costs were reduced with a long preparation time relative to a short preparation time. We found switch costs for both over-learned sequences (i.e., weekdays or numbers; Experiments 1-3) and new sequences (Experiments 3 and 4). Additionally, also mixing costs were found with this new paradigm (Experiment 2). Based on the pattern of results obtained in the four experiments of this study, we suggest that our novel language switching paradigm constitutes a viable tool to examine endogenous language switching.

In the following we first discuss the influence of response predictability and response preparation on language switching, which is followed by a discussion on how the results of Experiments 1-4 can be accounted for by two models of cognitive control (i.e., reconfiguration model and proactive interference model). Finally, we discuss the different strengths of cued language switching, voluntary language switching, and the SBLS paradigm.

Response predictability and response preparation

Based on the smaller switch costs due to increased preparation time in Experiments 1, we can conclude that participants used the predictability information, being it either concept predictability, language predictability or the combination of both, since longer pacing-intervals could not have caused reduced switch costs if this was not the case. So, even though participants used the predictability information to prepare for an upcoming response, switch costs were obtained. Please note that in cued language switching, the preparation time consists of the time between language cue and stimulus. That is, participants could prepare for the upcoming language, but not for a specific response. In contrast, in Experiment 1 preparation was possible up to response selection, because the concept was also predictable. Further preparation was not possible due to the jitter, which prohibited the participants to know when the actual response production should take place. In Experiments 2-4, however, the task allowed response execution to be predictable, by introducing a predictable response onset. Still, we observed substantial switch costs in these experiments.

The observation of switch costs with fully predictable responses is of theoretical importance because it provides evidence that language switch costs cannot be completely counteracted by processing stages that occur prior to response preparation (e.g., lexical selection). Yet, one could argue that in the present experiments, preparation time was not long enough to prepare a response. However, preparation time in Experiment 1 (i.e., 1500, 1700 or 1900 ms between the end of a response and the next response-signal in the long preparation time condition, since the average response duration of the weekdays was 300 ms) was substantially higher than in other language switching studies that investigated preparation time (Costa & Santesteban, 2004: 800 or 1200 ms; Philipp, Gade & Koch, 2007: 1100 ms; Verhoef, Roelofs & Chwilla, 2009: 750 or 1500 ms). Additionally, based on a meta-analysis on picture naming (Indefrey & Levelt, 2004), which provides a temporal structure of word production in picture naming, phonological encoding is scheduled between 455 to 600 ms after picture onset. Thus, we are confident that the preparation time in the present experiments was long enough to prepare the responses at least up to the level of phonological encoding.

To account for residual switch costs with predictable responses, we suppose that the participants were not able to fully prepare the responses due to the impact of response-related processes. Previous task switching studies have also provided evidence that response-related processes, such as response selection and response execution, influence switch costs (Philipp, Jolicoeur, Falkenstein & Koch, 2007; Schuch & Koch, 2003; Verbruggen, Liefooghe &

Vandierendonck, 2006). In a recent language switching study, Declerck, Koch and Philipp (2012) found that language switch costs were smaller during digit naming than picture naming. By contrasting specific characteristics of digits against pictures, it was found that the difference between digit naming and picture naming was due to the large amount of cognates in the digit stimulus-set. Hence, this study provided additional evidence for a response-related influence on language switch costs (see also Christoffels, Firk & Schiller, 2007).

Failing to fully prepare responses (i.e., concept and language) could be accounted for by response-related processes, such as articulation. It was argued that the participants could prepare up until response execution in Experiments 2-4, however, it is feasible that solely the mental representations of sounds can be prepared for, whereas the actual motor movement of the mouth, vocal cords, and tongue are not, especially since both languages implement different motor movement (e.g., due to language-specific phonemes). Evidence for the influence of articulation can be found in Experiment 4. The numerical data in Experiment 4 revealed larger L2 switch costs for responses that contained language-specific phonemes than for responses that only contained language-unspecific phonemes (see Table 5). This result reflects an execution-related locus of language control. More specifically, we assume that articulation is a locus of language control.

Along the same lines, switch costs found in the current study might be partially due to language comprehension, elicited by the participants own language production (Levelt, 1989; see Declerck & Kormos, 2012, for a discussion about the bilingual aspects of language monitoring). This is important, since in the current study language comprehension of one’s own speech might have strengthened the activation of the produced language, thus making it harder to switch to another language and increasing the priming effect if the language has to be repeated in the next trial. Both these effects would increase language switch costs and thus make it harder to fully prepare for an upcoming response.

Predictable language switching and models of cognitive control

There are two influential cognitive control models, borrowed from task switching (for a review of models in task switching, see Kiesel et al., 2010; Vandierendonck, Liefooghe & Verbruggen, 2010) that could account for our results. The first is the reconfiguration model (e.g., Rogers & Monsell, 1995), which assumes that during switch trials the task needs to be reconfigured on the level of the language and on the level of the actual stimulus (i.e., concept). Importantly, this model suggests that switch costs might be abolished, given ample preparation time and all the necessary information (i.e., language and concept). This means that in the current set-up switch costs should disappear, since the task meets both requirements. However, this was not the case in the current study. We found switch costs with predictability up to response selection and found switch costs and mixing costs with predictability up to response execution. Thus, the data does not correspond with the assumption that switch costs can be abolished with substantial preparation time and both languages and concepts being predictable. However, if articulation elicits language switch costs, then the reconfiguration model would fit the current data if the role of articulation would be taken into account. The model, as it is described by Rogers and Monsell (1995), would be able to account for preparation effects up to mental representations of sounds, which would be followed by a motor process. However, while the motor process also elicits switch costs, these switch costs would be unaffected by preparation.

The second model is the proactive interference model (Allport, Styles & Hsieh, 1994; Green, 1998). In this model, activation of the previously used task persists and thus causes either interference with the current task (switch trials) or results in residual activation and thus facilitation (repetition trials). Different from the reconfiguration model, this model does not make any claims about abolishing switch costs, which is in line with the current findings. The proactive interference model could account for switch costs with predictable responses by assuming that interference from the previous trial cannot be abolished completely by preparation, which would mean that on switch trials there would always be interference from the prior trial, which could be diminished, but not abolished (see also Koch & Allport, 2006).

To specify on which level there would be interference, which could not be abolished by preparation, in the proactive interference model, we turn to the ICM (Green, 1998). The ICM assumes that proactive interference is a process of inhibiting language tags, which are connected to the mental representations of words (i.e., lemmas; Levelt, Roelofs & Meyer, 1999). Hence, within the ICM, language tags that were previously inhibited would remain inhibited, even after substantial preparation time.

The proactive interference model could also account for switch costs due to articulation by assuming that articulation is a level on which interference plays a role. However, this is not in line with the assumptions of the ICM, which assumes that the language control process occurs earlier in the language production process (i.e., lemma level). Thus, the level(s) on which interference occurs in language switching still needs to be examined further.

The different strengths of cued language switching vs. voluntary language switching vs. sequence-based language switching

While the majority of language switching studies are cued language switching studies (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999), switch costs also appear when using the voluntary language switching paradigm (Gollan & Ferreira, 2009) or the new SBLS paradigm. This leaves future research with three options, depending on the different strengths of each paradigm, or an array of hybrid variations, such as an alternating language sequence with visually presented stimuli (e.g., Festman, Rodriguez-Fornells & Münte, 2010).

Cued, voluntary, and sequence-based language switching, in their strictest set-up, differ on several levels. First, stimuli are visually triggered when using the cued and voluntary language switching paradigm, whereas the SBLS paradigm relies on a predictable concept sequence that is memory-based.

The second difference is on the language level. While languages are visually triggered in a random fashion during the cued language switching paradigm, they are endogenously triggered in the voluntary language switching and SBLS paradigms. However, both the voluntary language switching and SBLS paradigms use different types of endogenous language triggers. In the voluntary language switching paradigm, the participants have to choose which language to respond in during each trial. The SBLS paradigm relies, identical to the concept level, on a predictable sequence that is memory-based (i.e., alternating language sequence).

These differences lead to different strengths for each of the three paradigms. The cued language switching paradigm, for one, is backed-up by an extensive amount of behavioral studies (e.g., Costa & Santesteban, 2004; Meuter & Allport, 1999; Philipp, Gade & Koch, 2007), neuro-imaging studies (e.g., Hernandez, Dapretto, Mazziotta & Bookheimer, 2001; Hernandez, Martinez & Kohnert, 2000) and event-related potential studies (e.g., Christoffels, Firk & Schiller, 2007; Verhoef, Roelofs & Chwilla, 2010), which created an extensive amount of knowledge about the effects and processes that are at play during cued language switching. This paradigm is also interesting when investigating the effect of preparation time, since there are different types of intervals (i.e., cue-to-stimulus interval and response-to-cue interval) that can be varied in cued language switching, which cannot be modified as readily in the other two paradigms.

One of the strengths of voluntary language switching is that it is closely related to actual language switching during natural speech (i.e., code switching). Consequently, this paradigm can be used to investigate this process in an experimental set-up. Furthermore, when using voluntary language switching, switch costs are not the only interesting phenomenon. Since participants can choose when to switch languages or repeat the same language, it also serves as a measure for *when* bilinguals switch to another language.

The SBLS paradigm also has several strengths. One of these strengths is its close resemblance to natural speech, due to the endogenous language retrieval and concept retrieval and the sequential nature of the task, since words are most often retrieved from memory and produced in a certain sequence. Additionally, using novel sequences, all word categories (e.g., verbs, adjectives, etc.) can be investigated using this paradigm, instead of only words that can be visually depicted as language-unspecific targets of naming responses. This could be crucial, since Experiment 2 of the current study has demonstrated that responses from different semantic categories elicit a difference in language switch costs. As the SBLS paradigm does not rely on pictures, one could, theoretically, even use whole sentences. Finally, it has to be noted that using visual cues and voluntary language switching can also be implemented in the SBLS paradigm, resulting in hybrid paradigms. This will of course change the concept of this paradigm and its strengths and constraints.

Conclusion

Language control during bilingual language selection is usually examined using stimulus naming tasks. Here we introduce a novel experimental paradigm relying on serial endogenous response generation. Our novel language switching paradigm demonstrates that both switch costs and mixing costs can be obtained using predictable responses, which indicates cognitive control during language selection. This effect might reveal a difficulty to fully prepare the upcoming concept and/or language, and thus response, which could be due to the relevance of late motor processes in language production.

References

Abutalebi, J., & Green, D. W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics, 20,* 242-275.

Abutalebi, J., & Green, D. W. (2008). Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Language and Cognitive Processes, 23*, 557-582.

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135,* 50-68.

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting attentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Perfomance XV: Conscious and nonconscious information processing* (pp. 421-452). Cambridge, MA: MIT Press.

Allport, A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan & A. M. Treisman (Eds.), *Attention, space and action: Studies in cognitive neuroscience* (pp. 273-296). Oxford: Oxford University Press.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 113,* 83-100.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33,* 570-585.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*, 3-11.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research, 1147*, 192-208.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for the model of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1283-1296.

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*, 491-511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory and Cognition, 32*, 1057-1074.

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

Declerck, M., & Kormos, J. (2012). The effect of dual task demands and proficiency on second language speech production. *Bilingualism: Language and Cognition, 15*, 782-796.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104*, 123-147.

Duyck, W., Depestel, I., Fias, W., & Reynvoet, B.(2008). Cross-lingual numerical distance priming with second-language number words in native- to third-language number word translation. *Quarterly Journal of Experimental Psychology, 61*, 1281-1290.

Festman, J., Rodriguez-Fornells, A., & Münte, T. F. (2010). Individual differences in control of language interference in late bilinguals are mainly related to general executive abilities. *Behavioral and Brain Functions, 6*, 5.

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*, 640-665.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*, 213-229.

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer (2001). Language switching and language representation in Spanish-English bilinguals: An fMRI study. *Neuroimage*, *14*, 510-520.

Hernandez, A. E., Martinez, A., & Kohnert, K. (2000). In search of the language switch : An fMRI study of picture naming in Spanish-English bilinguals. *Brain and Language, 73,* 421-431.

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92,* 101-144.

Jackson, G. M., Swainson, R., Mullin, A., Cunnington, R., & Jackson, S. R. (2001). ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition*, *4*, 169-178.

Jackson, G. M., Swainson, R., Mullin, A., Cunnington, R., & Jackson, S. R. (2004). ERP correlates of a receptive language-switching task. *The Quarterly Journal of Experimental Psychology, 57*, 223-240.

Kiesel, A., Wendt, M., Jost, K., Steinhauser, M., Falkenstein, M., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching: A review. *Psychological Bulletin, 136*, 849-874.

Koch, I. (2003). The role of external cues for endogenous advance reconfiguration in task switching. *Psychonomic Bulletin & Review, 10,* 488-492.

Koch, I. (2005). Sequential task predictability in task switching. *Psychonomic Bulletin & Review, 12*, 107-112.

Koch, I, & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory and Cognition, 34*, 433-444.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128*, 416-430.

Levelt,W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Liefooghe, B., Barrouillet, P., Vandierendonck, A., & Camos, V. (2008). Working memory costs of task switching. *Journal of Experimental Psychology: Learning, Memory and Cognition, 34*, 478-494.

Linck, J. A., Schwieter, J. W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition, 15*, 651-662.

Los, S. A. (1995). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica, 94*, 145-188.

Macnamara, J., Krauthammer, M., & Bolgar, M. (1968). Language switching in bilinguals as a function of stimulus and response uncertainty. *Journal of Experimental Psychology, 78*, 208-215.

Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P. II, & Yang, C. –L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition, 116*, 177-192.

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

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved Mysteries of the Mind (pp. 93-148)*. Hove, E. Sussex: Erlbaum (UK) Taylor and Francis.

Oppermann, F., Jescheniak, J. D., & Schriefers, H. (2010). Phonological advance planning in sentence production. *Journal of Memory and Language, 63*, 526-540.

Philipp, A. M., Jolicoeur, P., Falkenstein, M., & Koch, I. (2007). Response selection and response execution in task switching: Evidence from a go-signal paradigm. *Journal of Experimental Psychology: Learning, Memory and Cognition, 33*, 1062-1075.

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

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*, 395-416.

Portrat, S., Barrouillet, P., & Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *Journal of Experimental Psychology: Learning, Memory and Cognition, 33*, 1561-1564.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*, 207-231.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 763-797.

Schuch, S., & Koch, I. (2003). The role of response selection for inhibition of task sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 92-105.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin, 136*, 601-626.

Verbruggen, F., Liefooghe, B., & Vandierendonck, A. (2006). Selective stopping in task switching: The role of response selection and response execution. *Experimental Psychology, 53*, 48-57.

Verhoef, K. M. W., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition*, *110*, 84-99.

Verhoef, K. M. W., Roelofs, A., & Chwilla, D. J. (2010). Electrophysiological evidence for endogenous control of attention in switching between languages in overt picture naming. *Journal of Cognitive Neuroscience, 22*, 1832-1843.

Wang, Y., Kuhl, P. K., Chen, C., & Dong, Q. (2009). Sustained and transient language control in the bilingual brain. *NeuroImage, 47*, 414-422.

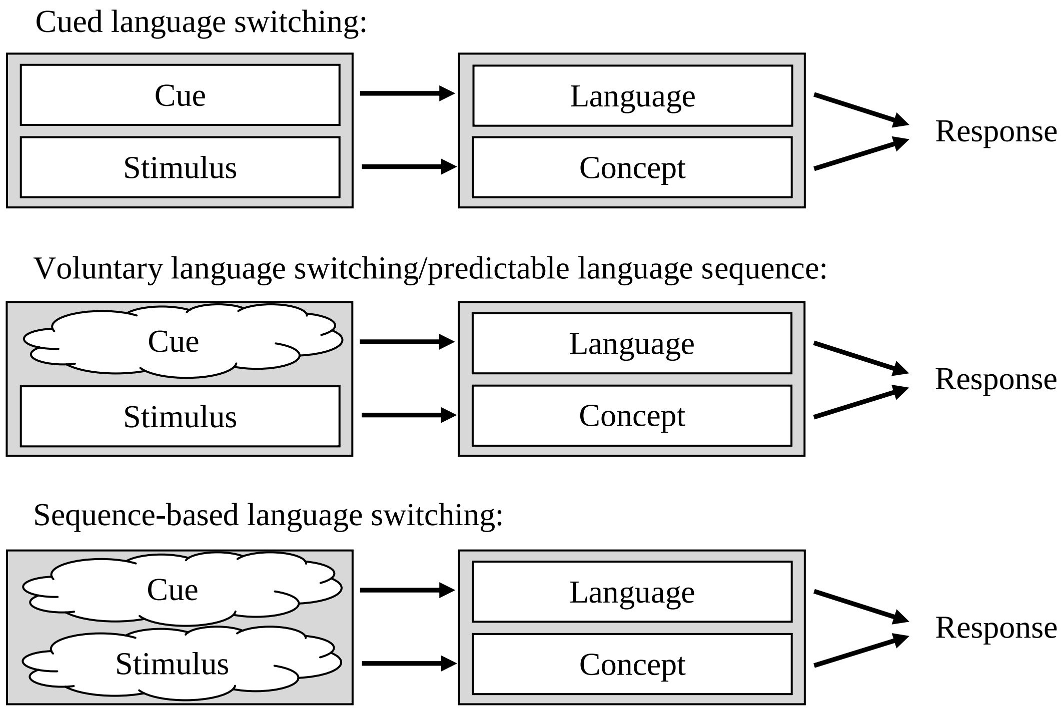

Figure 1. The top panel shows a typical progression of a trial during cued language switching: First the visually presented cue and stimulus, respectively, determine the language and concept. The combination of language and concept should then lead to a response. In the middle panel, which shows voluntary language switching/predictable language sequence, a similar progression takes place, apart from there being no exogenous language cues, but endogenous language cues. During sequence based language switching, which is depicted in the last panel, both language cue and visual stimuli are not exogenously presented, but are endogenous.

Table 1. Overview of demographic information of the participants of Experiments 1-4. The information consists of the average of English age of acquisition, the average years of formal English education, a self-rated score of spoken English from 1-7, with 1 being very bad and 7 being very good, and an average of known languages (not including the mother language).

Experiment	Age of acquisition	Formal English education	Self-rated score of spoken English	Known foreign languages
Experiment 1	9.6	8.3	4.7	2.0
Experiment 2	10.0	8.9	5.3	2.2
Experiment 3	9.9	9.4	5.0	1.8
Experiment 4	9.1	9.2	4.7	2.3

Table 2. Overall reaction time of Experiment 1 in ms (RT; SD in parenthesis) as a function of language transition (repetition vs. switch), preparation time (long vs. short interval) and language (German vs. English).

	Language			
	German		English	
	Preparation time			
	Long	Short	Long	Short
Switch trials	514 (31)	564 (40)	494 (26)	555 (38)
Repetition trials	473 (24)	496 (28)	462 (23)	490 (30)
Switch costs	41	68	32	65

Table 3. Overall reaction time of Experiment 2 in ms (RT; SD in parenthesis) as a function of language transition (repetition vs. switch vs. pure), semantic category (weekdays vs. numbers) and language (German vs. English).

	Language			
	German		English	
	Semantic categories			
	Numbers	Weekdays	Numbers	Weekdays
Switch trials	437 (26)	409 (26)	420 (25)	389 (25)
Repetition trials	419 (24)	364 (24)	413 (24)	361 (24)
Pure language trials	373 (23)	348 (23)	364 (18)	341 (18)
Switch costs	18	45	7	28
Mixing costs	46	16	49	20

Table 4. Overall reaction time of Experiment 3 in ms (RT; SD in parenthesis) as a function of language transition (repetition vs. switch), sequence condition (over-learned vs. scrambled sequence) and language (German vs. English).

	Language			
	German		English	
	Sequence condition			
	Over-learned	Scrambled	Over-learned	Scrambled
Switch trials	588 (50)	689 (38)	559 (51)	697 (43)
Repetition trials	547 (46)	619 (33)	538 (46)	628 (32)
Switch costs	41	70	21	69

Table 5. Overall reaction time of Experiment 4 in ms (RT; SD in parenthesis) as a function of language transition (repetition vs. switch), phonology (language-specific vs. language-unspecific phonology) and language (German vs. English).

	Language			
	German		English	
	Phonology			
	Language-specific	Language-unspecific	Language-specific	Language-unspecific
Switch trials	816 (81)	699 (59)	768 (74)	696 (59)
Repetition trials	741 (66)	623 (47)	688 (54)	648 (53)
Switch costs	75	76	80	48

Appendix.

Responses used in Experiments 1 and 2.

	Languages	
	German	English
Weekdays	Montag	Monday
	Dienstag	Tuesday
	Mittwoch	Wednesday
	Donnerstag	Thursday
	Freitag	Friday
	Samstag	Saturday
	Sonntag	Sunday
Numbers	eins	one
	zwei	two
	drei	three
	vier	four
	fünf	five
	sechs	six
	sieben	seven

Responses used in Experiment 3.

	Languages	
	German	English
Over-learned weekdays	Montag	Monday
	Dienstag	Tuesday
	Mittwoch	Wednesday
	Donnerstag	Thursday
	Freitag	Friday
	Samstag	Saturday
	Sonntag	Sunday
Over-learned numbers	eins	one
	zwei	two
	drei	three
	vier	four
	fünf	five
	sechs	six
	sieben	seven
Scrambled weekdays	Montag	Monday
	Freitag	Friday
	Mittwoch	Wednesday
	Dienstag	Tuesday
	Donnerstag	Thursday
Scrambled numbers	eins	one
	fünf	five
	drei	three
	zwei	two
	vier	four

Responses used in Experiment 4.

	Languages	
	German	English
Language-specific		
Stimulus-set A	Katze	cat
	Bürste	brush
	Zahn	tooth
	Prüfung	exam
	Rachen	throat
Stimulus-set B	Hauch	breath
	Pfütze	puddle
	Tuch	cloth
	Pilz	mushroom
	Rücken	back
Language-unspecific		
Stimulus-set A	Nagel	nail
	Stift	pen
	Bein	leg
	Tisch	table
	Huhn	chicken
Stimulus-set B	Kette	chain
	Glocke	bell
	Schlange	snake
	Gemälde	painting
	Ei	egg