Multilinguals' Language Control

Cumulative Ph.D. thesis

to obtain the joint degree of Doctor of Philosophy (Ph.D.)
at the University of Potsdam, University of Groningen, University of Trento,
University of Newcastle and Macquarie University

on the authority of

the President of the University of Potsdam, Prof. O. Günther, the Rector Magnificus of the University of Groningen, Prof. E. Sterken, the Rector of the University of Trento, Prof. P. Collini, the Pro-Vice-Chancellor of the University of Newcastle upon Tyne, Prof. S. Cholerton, and the Deputy Vice-Chancellor of Macquarie University, Prof. S. Pretorius and in accordance with the decision by the Doctoral Committee of the University of Potsdam.

Submitted to the Faculty of Human Sciences of the University of Potsdam

by

Michela Mosca

2017

Doctoral thesis defended on 14.07.2017

Articles of the cumulative Ph.D. thesis:

Mosca, M. & Clahsen, H. (2016). Examining language switching in bilinguals: The role of preparation time. *Bilingualism: Language and Cognition, 19*, 415-424. doi: 10.1017/S1366728915000693

Festman, J. & Mosca, M. (2016). Influence of preparation time on language control: A trilingual digit naming study. In J. W. Schwieter (Ed.) Cognitive control and consequences of multilingualism (pp. 145-170). Ontario, Canada: John Benjamins. ISBN: 9789027243720.

Mosca, M. Trilingual language switching: a flexible and strategic account (under review). *Quarterly Journal of Experimental Psychology*.

Mosca, M. & de Bot, K. Bilingual Language Switching: Production vs. Recognition (under review). *Frontiers in Psychology.*

Supervisor: Prof. Dr. Harald Clahsen (University of Potsdam)
Co-supervisor: Prof. Dr. Kees de Bot (University of Pannonia)

Preface

The present doctoral thesis was carried out under the auspices of the Erasmus Mundus Joint International Doctorate for Experimental Approaches to Language And Brain (IDEALAB) run by the University of Potsdam (GE), University of Groningen (NL), University of Trento (IT), Macquarie University (AU) and Newcastle University (UK), under grant no. <2012-0025/2013-1458-EMII EMJD>. A part of the work reported in this thesis was supported by an Alexander von Humboldt Professorship awarded to Prof. Harald Clahsen.

The laboratory work was mostly conducted at the Potsdam Research Institute for Multilingualism (PRIM) in Germany and partially at the department of Applied Linguistics at the University of Groningen in the Netherlands. Doctoral training to realize this dissertation was received from PRIM, University of Potsdam (GE), University of Groningen (NL), University of Trento (IT), Newcastle University (UK) and Macquarie University (AU).

The dissertation is written in English and is presented as a cumulative Ph.D. thesis at the University of Potsdam, Faculty of Human Sciences.

The thesis is composed of a general introduction (Chapter 1), an overview of the publications (Chapter 2) and four empirical chapters (Chapter 3 – 6), followed by a general discussion (Chapter 7) and a conclusion (Chapter 8). In the first chapter, the main research topic together with the main research questions of the dissertation are introduced. Thereafter, an overview and a summary of the manuscripts included in the dissertation are provided. The four empirical chapters present four manuscripts of which two are first-authorship publications (Chapter 3 and 6), one is a co-authorship publication (Chapter 4) and one is a single-author publication (Chapter 5). The manuscripts included in Chapters 3 and 4 have already been published, while the manuscripts in Chapters 5 and 6 are currently under revision in international peer-reviewed journals of the field. The empirical chapters are followed by a general discussion of the main findings of the four manuscripts. The thesis concludes with a chapter dedicated to the major conclusion, indications about the limitations of the present work and suggestions for future directions.

Acknowledgments

Since I was a teen, attending a high school focussing on foreign languages, I have wondered how humans manage to keep languages separate in their brain. Possible answers to this query started to come when I had the chance to dedicate myself to the deep investigation of this topic in my doctoral thesis. This doctoral thesis was possible thanks to my supervisor Prof. Harald Clahsen who has constantly guided and supported me throughout this journey with his sharp remarks and enlightening suggestions, and to my co-supervisor Prof. Kees de Bot for his unlimited support, the chances he gave me to grow academically and his invaluable advice.

My deepest gratitude goes also to Dr. Julia Festman for her precious insights and our intense discussions on a shared love called "language control".

I am, and always will be, profoundly grateful to the directors of the International Doctorate for Experimental Approaches to Language And Brain (IDEALAB), namely Prof. Roelien Bastiaanse, Prof. David Howard, Prof. Barbara Höhle, Prof. Gabriele Miceli and Prof. Lyndsey Nickels; without them, none of this would have been possible. Having the privilege of being part of the IDEALAB family was one of the most gratifying and memorable experiences of my life. Special thanks go to my colleagues from both the IDEALAB programme and the Potsdam Research Institute for Multilingualism (PRIM), most of whom are now close friends.

I am also very grateful to all the participants who took part in the experiments for tolerating exhausting sessions of testing "for the sake of science".

Writing a Ph.D. thesis is never easy. For this, I need to thank my personal hero who has inspired me with her strength and her courage in life, my mum. I am also endlessly thankful to my beloved family, whose warmth has never abandoned me despite the distance. My thanks are due also to my foster or "German" family that has unrelentingly supported and encouraged me when the way seemed to be dark and has shared with me the joy of sunny days. Special thanks goes to my older but always present friends, who have tirelessly accompanied me during this adventure. I am also thankful for the precious new friends Berlin has offered me, for experiencing the surprises of this journey along with me. Finally, I thank my partner for taking every step with me and his unconditional support and confidence in me.

Table of Contents

Summary

For several decades, researchers have tried to explain how speakers of more than one language (multilinguals) manage to keep their languages separate and to switch from one language to the other depending on the context. This ability of multilingual speakers to use the intended language, while avoiding interference from the other language(s) has recently been termed "language control".

At first, it was suggested that language control might be supported by a mental device for switching that allows speakers of two languages ("bilinguals") to go from one language to another (Penfield & Roberts, 1959), probably by turning one language *on* and the other *off* (Macnamara, Krauthammer & Bolgar, 1968). Subsequently, a multitude of studies showed that when bilinguals process one language, the other language is also activated and might compete for selection, discouraging the idea that languages could be completely deactivated.

Parallel language activation was held for both reception (Dijkstra, Timmermans & Schriefers, 2000; Blumenfeld & Marain, 2013) and production (e.g., Colomé, 2001; Poulisse & Bongaert, 1994). According to the most influential model of language control developed over the last two decades, competition from the non-intended language is solved via inhibition. In particular, the Inhibitory Control (IC) model proposed by Green (1998) puts forward that the amount of inhibition applied to the non-relevant language depends on its dominance, in that the stronger the language the greater the strength of inhibition applied to it. Within this account, the cost required to reactivate a previously inhibited language depends on the amount of inhibition previously exerted on it, that is, reactivation costs are greater for a stronger compared to a weaker language. In a nutshell, according to the IC model, language control is determined by language dominance. However, inconsistent findings within the language control literature have questioned the validity of this account (e.g., Costa & Santesteban, 2004; Costa, Santesteban & Ivanova, 2006; Verhoef, Roelofs & Chwilla, 2009).

The goal of the present dissertation is to investigate the extent to which language control in multilinguals is affected by language dominance and whether and how other factors might influence this process. Three main factors are considered in this work: (i) the time speakers have to prepare for a certain language or PREPARATION TIME, (ii) the type of languages involved in the interactional context or LANGUAGE TYPOLOGY, and (iii) the PROCESSING MODALITY, that is, whether the way languages are controlled differs between reception and production.

To investigate language control in multilinguals, the four studies presented in this dissertation made use of the language switching paradigm with late unbalanced multilinguals (one stronger native language, the "L1", plus one or two additional weaker non-native languages, the "L2" and "L3", respectively). The effect of preparation time was explored by manipulating the interval between the language cue (indicating the language to use in the next trial) and the stimulus in a bilingual picture naming task as well as in a trilingual digit naming task. The role of language typology was assessed by comparing two groups of trilinguals performing a trilingual picture naming task. For one group, the L3 was typologically closer to the L2, but typologically more distant from the L1. For the other group, the L3 was typologically more distant from both the L1 and the L2. Finally, the influence of processing modality was explored by comparing a group of bilinguals performing a bilingual lexical decision (recognition) task and a bilingual picture naming (production) task.

The results obtained in the four manuscripts, either published or in revision, indicate that language dominance alone does not suffice to explain language switching patterns. In particular, the present thesis shows that language control is profoundly affected by each of the three variables described above. Firstly, the data show that the cost of reactivating a previously suppressed language is dramatically altered by preparation time, rather than by language dominance. More precisely, results from the first two manuscripts reveal that, given ample preparation time, language switching costs for both the stronger and the weaker language can completely dissipate. This indicates not only that the cost of reactivating a language is not solely determined by its dominance, but also that language switching can be cost-free. Secondly, the present work shows that the way languages are controlled is affected by language typology. In particular, results from the third manuscript reveal that during language switching, typologically closer languages tend to interfere with one another to a greater extent than typologically more distant languages. Conflicts between languages are reduced by hampering the “disturbing” (stronger) language and/or by facilitating the “disturbed” (weaker) language. Thirdly, the present data reveal that language control is strongly modulated by processing modality. More precisely, the fourth and last manuscript shows that language control in reception seems to be modulated by language dominance, whilst language control in production appears to be regulated by a strategic controlling system aimed at optimizing performance. Overall, the findings obtained in the present dissertation indicate that language control in multilingual speakers is a much more dynamic system than previously believed and is not exclusively determined by language dominance, as predicted by the IC model (Green, 1998).

1 General Introduction

Imagine you are in a restaurant talking to a friend in language A, when the waiter approaches you in language B asking for your order. You know that in order to appropriately communicate with the waiter, you need to reply in language B, and not in language A. This ability to confine processing to the intended language, while reducing interference from the non-intended language(s) is called “language control”. The overarching goal of the present thesis is to investigate how speakers of more than one language (hereafter “multilinguals”) control their languages.

The focus of the present research is on the individual word level. More precisely, the work assesses how isolated words are selected from the multilingual mental lexicon (organization of words in the speaker’s mind) in production and recognition, and which factors might influence this process.

As a great deal of research in the last decades has systematically shown, selecting words from the mental lexicon (or “lexical selection”) is far from being an effortless process. This might explain why we experience difficulties finding the “right word” when we are tired, or to speak and write in a foreign language when we are cognitively overwhelmed (e.g., in the case of fatigue or emotional breakdown, see also Dornic, 1978). Indeed, the intention to utter a specific word (e.g., "cat") may activate not only the intended word but also semantically related words (e.g., “dog”, “mouse”), words sharing similar phonological and/or orthographic features (e.g., “cut”, “mat”) and words from a non-relevant language (e.g., “Katze”, “Hund”: German words for “cat” and “dog”). Hence, during lexical retrieval both words from the relevant and non-relevant language(s) are activated and might compete for selection (e.g., de Bot, 1992; Poulisse & Bongaert, 1994, yet see Costa & Caramazza, 1999; Costa, Miozzo & Caramazza, 1999). This has been postulated for different aspects of language processing, namely for production (Colomé, 2001; Kroll, Bobb, & Wodniecka, 2006, Poulisse & Bongaert, 1994), visual word recognition (Dijkstra & van Heuven, 1998; Grainger, 1993; Grainger & Dijkstra, 1992), spoken word recognition (Grosjean, 1988, 1997; Marian & Spivey, 2003; Shook & Marian, 2013), and for different levels of competition, such as semantic (Mägiste, 1984, 1985; Runnqvist, Strijkers, Alario & Costa, 2012) and phonological/orthographic competition (e.g., Beauvillain & Grainger, 1987; Dijkstra, Grainger, van Heuven, 1999). The present study focusses on how competition from words of the non-relevant language(s) is resolved and which factors play a role during this process.[1]

According to the most influential model of language control, the Inhibitory Control (IC) model proposed by Green (1998), the way languages are controlled is determined by their dominance level. While numerous studies have corroborated this hypothesis (e.g., de Bruin, Roelofs, Dijkstra &

---

[1] For readers interested in word competition within languages, see e.g., Dell, 1986; Levelt, Roelofs & Meyer, 1999, for language production; McClelland & Rumelhart, 1981 and McClelland & Elman, 1986, for visual and spoken word recognition, respectively; Howard, Nickels, Coltheart, & Cole-Virtue, 2006, Neely, 1991, for semantic competition and Damian & Bowers, 2003; Lupker, 1982, for phonological/orthographic competition.

Fitzpatrick, 2014; Jackson, Swainson, Mullin, Cunnington & Jackson, 2004; Linck, Schwieter & Sunderman, 2012; Macizo, Bajo & Paolieri, 2012; Meuter & Allport, 1999), many others have questioned the validity of this model (e.g., Costa et al. 1999; Costa & Santesteban, 2004; Costa, Santesteban & Ivanova, 2006; Finkbeiner, Almeida, Janssen & Caramazza, 2006; La Heij, 2005; Verhoef, Roelofs & Chwilla, 2009). The aim of the present work is to shed light on this controversy, by examining to what extent language control is affected by language dominance and investigating whether and how other variables might influence this process. Three main factors are considered in this dissertation: (i) the time that speakers have at their disposal to retrieve words in a specific language or PREPARATION TIME; (ii) the type of languages (more vs. less similar languages) involved in the interactional context or LANGUAGE TYPOLOGY, and (iii) the MODALITY in which languages need to be controlled (during production vs. recognition). This investigation will be carried out through four major steps: Firstly, the introductory chapters of this thesis will review the prevailing models of language control and the salient findings obtained by previous studies on task and language switching. At the end of this section, the overall aims and objectives of the present work will be outlined. Secondly, an overview of the four publications included in this dissertation will be given, and this will be followed by, thirdly, the four individual manuscripts that represent the central part of this dissertation. Finally, the last section provides a general discussion and conclusion of the main results obtained in this dissertation.

1.1 The IC Model

The starting point of the present work is the assumption that language control might rely on an inhibitory mechanism as put forward by Green (1986, 1993, 1998) in his Inhibitory Control (IC) model of bilingual language control. According to this view, selection of the intended language is achieved by suppressing competing words of the non-intended language. This has been postulated for both language recognition and production. Regarding language recognition, Green (1998) observes that when a visual word is presented both words from both the relevant and the non-relevant language might activate. In order to decide to which language the presented word belongs, a task schema has to be established that relates an output of the bilingual mental lexicon (e.g., L1 language tag present) to the appropriate response (e.g., press right key). A task schema can be understood as a functional control circuit that exerts control by activating and inhibiting words based on their language membership. Language membership is specified by means of language tags, in that each word in the bilingual lexicon is assumed to be associated with either an L1 or L2 language tag. Thus, when the appropriate task schema is activated (e.g., press right key for L1 response), all words belonging to the relevant language will be activated (e.g., all words with L1 language tags), while all words from the other language (e.g., all words with L2 language tags) will be suppressed. In the case of language change, words belonging to the new language (e.g., L2) must be activated, while words from the previously activated language (e.g., L1) need to be suppressed. Reactivating a previously suppressed language will require time, so a switching cost is predicted in the case of language change during visual word recognition. As for language production, the author notes that again both the target and the non-target language are potentially active and that in order to utter a word in the appropriate language, the relevant task schema has to be activated. This is achieved by linking an external cue indicating the language to be used (e.g., blue screen background) to the appropriate language (e.g., use L1 when screen background is blue). The relevant language is thus activated. At the same time, the non-relevant task schema (e.g., use L2 when screen background is green) needs to be suppressed. In this way, the non-relevant language is inhibited. Similar to recognition, in the case of language change, the previously inhibited language will require some time to reactivate, leading to language switching costs. Importantly, in both recognition and production, inhibition of the irrelevant language is supposed to be reactive, that is, more active languages (e.g., more dominant languages) require a higher degree of inhibition compared to less active languages (e.g., less dominant languages). In this framework, the cost of reactivating a language depends on the magnitude of inhibition previously applied to it, i.e., the greater the amount of inhibition the larger the reactivation cost. Following this logic, Green (1998) suggested that reactivation costs (or "switching costs") for a more and a less dominant language (for example, in the case of unbalanced bilinguals) would be *asymmetrical*, namely larger for the stronger compared to the weaker language.

A valuable way to measure language switching costs is by using a language switching paradigm. This task includes two types of trials, a *repetition trial* (stimulus is in the same language as in the preceding trial, e.g., L1-<u>L1</u>) and a *switch trial* (stimulus is in a different language than in the previous trial, e.g., L2-<u>L1</u>). Switch trials are usually responded more slowly and less accurately than repetition trials, and this difference is known as "switching costs" (e.g., Allport, Styles & Hsieh, 1994; Meiran, 1996; Roger & Monsell, 1995). In line with the assumption made by the IC model, Meuter and Allport (1999) demonstrated that switching costs for unbalanced bilinguals were asymmetrical, namely larger in the stronger L1 than in the weaker L2. Importantly, the authors also showed that when the proficiency difference between the stronger L1 and the weaker L2 was relatively small, switching costs for the two languages became comparable, i.e. *symmetrical* switching costs. Overall, these results are consistent with the idea developed within the task switching literature that elements from the previous task might have carry-over effects on the upcoming trial, causing negative priming (inhibition) in cases of task change. This is known as the "Task Set Inertia" hypothesis (Allport, et al., 1994). Consider, for example, a modified version of the Stroop paradigm, where participants either name the colour in which a colour word is written (e.g., RED for the word "green" printed in red) or read the word irrespective of the colour in which it is written (e.g., GREEN for the word "green" printed in red). According to a subsequent interpretation of the Task Set Inertia hypothesis (Wylie & Allport, 2000), when participants are requested to name the colour, the competing task, "word reading", needs to be suppressed (and vice versa). If, on the following trial, "word reading" becomes the relevant task, the inhibition previously applied to this task will interfere, causing a delay in producing the word. Therefore, within this framework, switching costs are seen as evidence of persisting inhibition of the previously irrelevant task into the current task. Furthermore, since word naming is a more automatic task than colour naming, the amount of inhibition applied to the more automatic task (word reading) is greater compared to the less automatic task (colour naming). Because of this asymmetry, persisting inhibition coming from the stronger task is greater compared to the weaker task, leading to larger switching costs when the stronger instead of the weaker task needs to be reactivated.

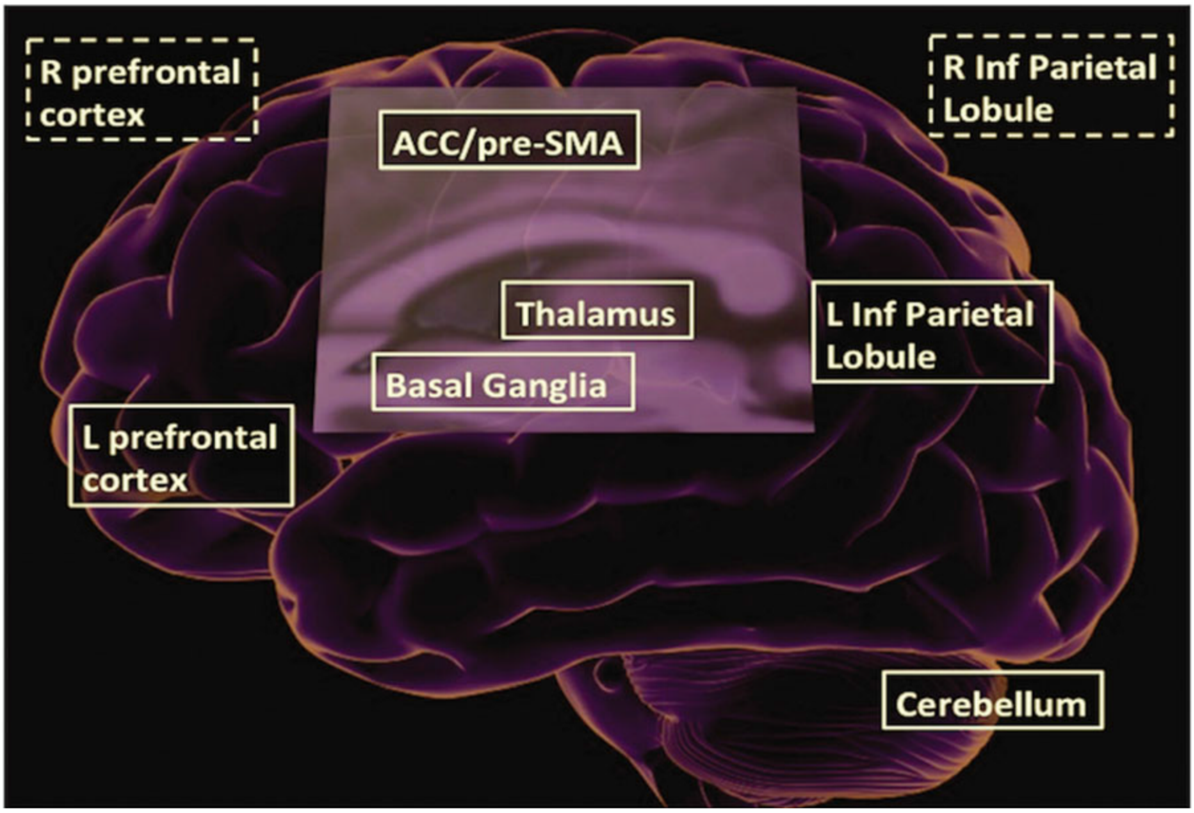
This interpretation, however, leads to the question whether switching costs might also be explained as persisting activation of the previous relevant task in the current trial. For example, one might suppose that a weaker task (e.g., naming colour), being less automatic, needs to be activated more than a stronger task (e.g., word reading). In this scenario, persisting activation from the weaker into the stronger task would be stronger than otherwise, leading to larger switching costs in the stronger than in the weaker task (see also Yeung & Monsell, 2003, for an hypothesis of persisting activation in task switch). Since both persisting inhibition and persisting activation predict larger switching costs for the stronger compared to the weaker task, it is difficult to determine the origin of the switching costs by using only two tasks. Evidence for the central role of inhibition during task switching comes, indeed, from studies investigating persisting inhibition using three tasks (e.g., Mayr & Keele, 2000; for a review see Koch, Gade, Schuch & Philipp, 2010). Specifically, performance in n-2 repetition trials (e.g., A-B-*A*) has been found to be slower and more error prone than in non-repetition trials (e.g., C-B-*A*). This effect has been interpreted as evidence for persisting inhibition of the n-2 trial (e.g., *A*-B-A) into the current trial (e.g., A-B-*A*). Crucially, these results cannot be accounted for by a persisting activation account, which predicts positive priming (facilitation) on n-2 repetition trials compared to non-repetition trials. The effect of persisting inhibition has been found both in task switching (e.g., Mayr, 2001, 2009; Philipp & Koch, 2006; Schuch & Koch, 2003; Sdoia & Ferlazzo, 2008) and language switching studies (e.g., Babcock & Vallesi, 2015; Declerck, Thoma, Koch & Philipp, 2015; Guo, Ma & Liu, 2013; Guo, Liu, Chen & Li, 2013; Philipp, Gade & Koch, 2007; Philipp & Koch, 2009).

Further evidence supporting the hypothesis that language control relies on inhibitory mechanisms comes from neurophysiological and neuroimaging research. A number of studies recording event-related potentials (ERPs) during language switching have detected a significant increased negativity peaking approximately 250-350ms after the stimulus onset (N2 component) in switch compared to repetition trials (e.g., Jackson, Swainson, Cunnington & Jackson, 2001; Khateb, Abutalebi, Michel, Pegna, Lee-Jahnke & Annoni, 2007; Misra, Guo, Bobb & Kroll, 2012; Verhoef et al., 2009). The N2 component is usually associated with response inhibition, in that it is elicited when participants are requested to withhold a response (no-go stimuli) compared to when they are asked to give a response (go stimuli) in a Go/No-go paradigm (e.g., Gemba & Sasaki, 1989; Jodo & Kayama, 1992; Pfefferbaum, Ford, Weller & Kopell, 1985; Thorpe, Fize & Marlot, 1996). Thus, the N2 found in language switching studies is usually believed to reflect general (not specific for language) inhibitory processes (e.g., Jackson et al. 2001; Christoffels et al., 2007; Liu, Rossi, Zhou & Chen, 2014; Misra et al., 2012; but see Verhoef et al., 2009). Beyond the N2, language control has been associated with a positive deflection starting around 500ms, the Late Positive Component (LPC) (e.g., van der Meij, Cuetos, Carreiras & Barber, 2011; Ng, Gonzalez & Wicha, 2014; Moreno, Federmeier & Kutas, 2002). An enhanced LPC is usually found in switch compared to repetition trials and is believed to reflect the cognitive processes of conflict resolution (Liotti, Woldorff, Perez & Mayberg, 2000; Jackson et al., 2001; Liu et al., 2014), such as linking the input to the correct lexical item after a language change (Martin, Strijkers, Santesteban, Escera, Hartsuiker & Costa, 2013). In particular, it is assumed that while the N2 is related to control in general, the LPC is linked to the consequences of this control at the level of specific lexical representations (Martin et al., 2013). More generally, the LPC is associated with those processes requiring integration or reanalysis of the information after an unexpected event (Coulson, King, & Kutas 1998; McCallum, Farmer, & Pocock 1984). In line with the IC model, some studies have shown that the modulation of these two components might be related to language dominance, with a larger N2 for L2 than L1 trials probably reflecting a greater amount of inhibition on the stronger L1 when the weaker L2 is the relevant language than vice versa (Jackson et al. 2001, yet see Christoffels et al., 2007), and a larger LPC for L2 relative to L1 trials presumably indicating that more conflict resolution processes are involved during L1 than L2 suppression (Liu et al., 2014; but see Jackson et al., 2001, 2004).

The idea that language control is mainly implemented by linguistic-independent inhibitory processes has found vast support also among neuroimaging studies. In one of the earliest studies using fMRI observing the neural aspects of language switching, Hernandez and colleagues (2000) found that the activity of the dorsolateral prefrontal cortex (DLPFC) was greater in the dual-language condition (two languages involved in the task) relative to the single-language condition (one language involved in the task). Thereafter, a number of studies provided evidence for the crucial role of the DLPFC in bilingual speakers (e.g., Hernandez, Dapretto, Mazziotta & Bookheimer, 2001; Rodriguez-Fornells, Rotte, Heinze, Noesselt & Muente, 2002; Rodriguez-Fornells, van der Lugt, Rotte, Britti, Heinze, & Muente, 2005). Generally, the DLPFC has been assumed to be related to general executive functions such as response switching and response suppression (e.g., Egner & Hirsch, 2005; Kerns, Cohen, MacDonald, Cho, Stenger & Carter, 2004; MacDonald, Cohen, Stenger & Carter, 2000). In a modified version of the Stroop task, MacDonald et al. (2000) noted that the activation of the left DLPCF was connected to general attentional processes. More precisely, the authors gave their subjects instructions whether to read a colour word (more automatic task) or name the ink colour in which a word was written (less automatic task) before each trial. Following a delay, the stimulus was presented. Authors noted that the activation of the left DLPFC was instruction-related, in that a greater activity was observed when participants were told to name the ink colour in which the word was written (less automatic task) compared to the situation where they were instructed to read the word irrespective of its ink colour (more automatic task). This finding led the authors to suggest that in order to perform less automatic tasks, such as colour naming, the DLPFC needs to increase the amount of top-down control so as to overcome more automatic responses, such as word reading. Further, the authors found that the activation of the Anterior Cingulate Cortex (ACC) was not modulated by the instructions, but that its activation was stimulus-related. In particular, they found increased ACC activity for stimuli where the word and its ink colour were incongruent (word RED printed in green ink) compared to stimuli with congruent word and colour ink (word RED printed in red ink). Based on this observation, MacDonald and colleagues (2000) suggested that the ACC is a cortical area responsible for conflicts detection and performance monitoring. Altogether, the authors proposed that while the DLPFC is mainly engaged during the preparatory period and appears to be involved in implementing and maintaining control, the ACC is activated during response execution and seems to be involved in evaluating performance and signaling to the DLPFC when the level of control needs to be enhanced (e.g., in the case of increased conflict detection). Similar to research on general cognitive control, successive studies on language switching have shown that the DLPFC is not the only area involved during language control. Indeed, language control is assumed to be implemented by a neural network comprising regions similar to those involved in non-verbal cognitive tasks, namely the DLPFC, the ACC and the caudate nuclei (e.g., Crinion et al. 2006; Abutalebi, Annoni, Seghier, Zimine, Lee-Jahnke, Lazeyras, Cappa & Khateb, 2008; Abutalebi, Brambati, Annoni, Moro, Cappa & Perani, 2007; Wang, Xue, Chen, Xue & Dong, 2007; Wang, Kuhl, Chen & Dong, 2009; Guo, Liu, Misra & Kroll, 2011). For example, in the bilingual case, the intention to speak in the less automatic L2 might be satisfied by the DLPFC, which suppresses prepotent responses of the more automatic L1, whereas conflicts arising from the simultaneous activation of incompatible responses, such as trying to say the word "cat" while simultaneously saying "gatto", might be detected by the ACC. Moreover, it has been suggested that, together with the ACC, also the head of the left caudate nucleus might be involved in the selection and inhibition of lexical alternatives (Abutalebi & Green, 2007). More broadly, according to the proposed neurocognitive models of bilingual language processing (Abutalebi & Green, 2007, 2008, 2016; Green & Abutalebi, 2013), language control is orchestrated by a network of cortical and subcortical networks, and the disruption of one or more of these neural circuits might lead to serious language control impairment, such as pathological language switching (e.g., Abutalebi, Miozzo, & Cappa, 2000; Mariën, Abutalebi, Engelborghs & De Deyn, 2005) and pathological language fixation (Aglioti, Beltramello, Girardi, & Fabbro, 1996; Aglioti & Fabbro, 1993).

Apart from the areas described above, in their recent neurocognitive model of bilingual language processing, Abutalebi and Green (2016) proposed that the neural network implementing language control includes numerous other neural substrates. More precisely, the pre-supplementary motor area (pre-SMA) is believed to be involved in initiating speech in language switching and, together with the dorsal ACC, is thought to monitor potential conflicts between languages as well as detect errors. The inferior parietal lobules (left and right, LIPL and RIPL) are associated with the maintenance of task representations, in that the LIPL is responsible for biasing selection away from the non-target language and the RIPL is in control of biasing selection towards the target language (see also Abulatebi & Green, 2008). Connected to the prefrontal cortex, the thalamus is also considered to implement language control by supporting the retrieval of relevant semantic and lexical representations. In addition to the left caudate mentioned earlier, another part of the basal ganglia seems to be relevant during bilingual speech production, namely the left putamen. Indeed, this subcortical structure appears to be specifically engaged when multilinguals use a language not mastered in a native-like manner (see also Abutalebi, Della Rosa, Castro Gonzaga, Keim, Costa & Perani, 2013). Finally, the cerebellum that is linked to all the key areas related to language control is thought to be involved in controlling verbal interferences (see also Filippi, Richardson, Dick, Leech, Green, Thomas & Price, 2011). Figure 1 illustrates the brain regions involved in language control according to Abutalebi and Green (2016).

Figure 1. Brain regions related to language control, taken from Abulatebi and Green (2016).



Altogether, evidence from the behavioural, neurophysiological and neuroimaging studies discussed above indicate that language control is likely to be implemented by means of inhibitory processes. As discussed before, according to the IC model (Green, 1998), which heavily relies on the concept of inhibition, the cost required to reactivate a previously suppressed language depends on its dominance, with a stronger language being harder to reactivate than a weaker one. In this view, reactivation or switching costs are predicted to be larger for the stronger than for the weaker language. Interestingly, in a study investigating language control in unbalanced bilinguals, Verhoef and colleagues (2009) found that switching costs were larger for the stronger than for the weaker language when speakers had less time to prepare (750ms), and that switching costs became comparable for the two languages when preparation time was longer (1500ms). This indicates that both symmetrical and asymmetrical switching costs can be found within the same group of speakers and that switching costs can be affected by the time speakers have at their disposal to prepare for the new trial.

When looking at the effect of preparation time on language control, two possible sources of switching costs have to be considered. Indeed, besides the undeniable role of (i) persisting inhibition during task switching, several authors have suggested that switching costs might also reflect (ii) the effort of reconfiguring the new task (e.g., Roger & Monsell, 1995; Mayr & Kliegl, 2000, 2003;

Meiran, 1996; Rubinstein, Meyer & Evans, 2001). Task-set reconfiguration may include processes such as identification of the new task, goal shifting, activation of the new rule and inhibition of the non-relevant task (e.g., Rubinstein et al, 2001; Monsell, 2003). Thus, as discussed by Monsell (2003), both carry-over effects from the previous trial and task-set reconfiguration are likely to be the source of switching costs.

One way to examine the effect of task-set reconfiguration or "active preparation" in language control is by manipulating the interval between the display of the language cue (indicating which language to use in the upcoming trial) and the presentation of the stimulus (Cue-Stimulus Interval, CSI). A sizable number of studies have found that longer CSI or "preparation time" leads to switching cost reduction (e.g., Costa & Santesteban, 2004; Fink & Goldrick, 2015; Guo et al., 2013b). Moreover, there is some evidence suggesting that the amount of preparation time modulates the degree of switching cost reduction, i.e. the longer the preparation the smaller the switching costs (Costa & Santesteban, 2004). These results indicate that word retrieval in a certain language can be prepared beforehand and that the degree of preparation may depend on the time at one's disposal. However, it might be argued that longer preparation time allows not only for an advanced reconfiguration of the new task but also for a passive dissipation of the persisting inhibition from the previous task. A way to distinguish between the two processes is to extend the interval between two trials. A longer interval between trials can be created by prolonging the interval between the display of the stimulus and the start of the new trial (Inter-Trial Interval, ITI) or by prolonging the interval between the response and the display of the next cue (Response-Cue Interval, RCI). The idea is that longer intervals between trials would allow for the passive decay of the interference from the previous trial, and therefore give space to the investigation of active preparation of the upcoming trial (see also Meiran, Chorev & Sapir, 2000). Unfortunately, in studies examining language control, the interval between trials was either relatively short (e.g., Declerck, Koch & Philipp, 2012; Costa & Santesteban, 2004) or left uncontrolled (e.g., Verhoef, Roelofs & Chwilla, 2009; Fink & Goldrick, 2015; Philipp et al., 2007). Because of this, it is difficult to pinpoint how strongly language switching costs are determined by the persisting inhibition coming from the previous trial and to what extent multilingual speakers may be able to prepare in advance for a language switch in a language switching context.

1.2 Switching vs. Mixing Costs

While language switching costs are usually taken as reliable indicators of the mechanisms underpinning language control (yet, see Bobb & Wodniecka, 2013), much less attention has been devoted to another kind of costs arising from language switching, that is, "mixing costs". One way to measure language mixing costs is by including in the experimental design a so-called *single*-language block, where stimuli from each language are presented separately. This is opposite to the *mixed*-language block, in which stimuli from more than one language are intermingled and from which language switching (repetition vs. switch trials) costs can be measured. Performance in the mixed-language block is usually slower and more error prone than in the single-language block (e.g., Los, 1996; Spector & Biederman, 1976). Interestingly, repetition trials of the mixed-language block are also found to be answered more slowly and less accurately than trials in the single-language block (e.g., Rubin & Meiran, 2005; Philipp, Kanilich, Koch & Schubotz, 2008). In both cases, the difference between the mixed- and the single-language block is known as *mixing costs*. However, when calculated in the latter way, mixing costs can be neatly separated from switching costs. Therefore, comparing repetition trials in the mixed-language block with trials in the single-language block has become the standard way of measuring mixing costs.

It is generally agreed that mixing and switching costs reflect two different kinds of processes (e.g., Koch, Prinz & Allport, 2005). Switching costs are supposed to reflect the temporary or “transient” control of disentangling from the previous trial and preparing for the new one. In contrast, mixing costs are believed to reflect a prolonged or “sustained” control of maintaining multiple tasks/languages active in the mixed- compared to the single-language block (e.g., Mayr, 2001; Roger & Monsell, 1995). Therefore, it is possible to distinguish between trial-specific and non trial-specific costs of language control, the former reflected in the switching costs and the latter in the mixing costs (Braver, Reynolds & Donaldson, 2003).

However, what mixing costs may reflect is still a matter of debate. Indeed, there is growing evidence that mixing costs may not be as sustained or “fixed” as believed. For example, some language switching studies found mixing costs in both L1 and L2 (e.g., Hernandez, Martinez & Kohnert, 2000; Prior & Gollan, 2001), others only in L1 but not in L2 (e.g., Christoffels, Firk & Schiller, 2007) and some others found a “mixing benefit” (faster response in the mixed- compared to the single-language block) in the weaker but not in the stronger language (e.g., Gollan & Ferreira, 2009; Hernandez et al., 2001). In this regard, it has been suggested that mixing costs might reflect adjustment of languages’ activation level in the mixed- compared to the single-language block (e.g., Christoffels et al., 2007). The idea is that, in the mixed-language block, speakers might prevent interference from the stronger into the weaker language by lowering the activation level of the stronger language (slower responses compared to the single-language block) and/or by enhancing the activation level of the weaker language (faster responses compared to the single-language block). In this perspective, mixing costs have been associated with a sustained (non trial-specific) mechanism of language control that aims at preventing interference before it occurs (see e.g. Ma, Li & Guo, 2016).

Whether mixing costs represent the effort required to keep multiple tasks active and/or the adjustment of languages' activation level in order to prevent interference, it remains difficult to make clear predictions in terms of the mixing costs pattern, given the diverse results found in the literature. Crucially, while the IC model makes clear assumptions about switching costs, no predictions are made about mixing costs. In this respect, a good start would be to expand our knowledge about the nature of mixing costs, by exploring, for example, whether mixing costs reflect non trial-specific costs as usually proposed or whether they might be susceptible to manipulation at the trial level, such as preparation time variation.

1.3 Influential Factors in Language Control

The fact that the interactional context (e.g., single vs. mixed language) can affect the way languages are controlled is not a new idea (e.g., de Groot, 1998; Grainger & Dijkstra, 1992; Grosjean, 1998, 2001). For example, according to the "adaptive control hypothesis" (Green & Abutalebi, 2013), language control is influenced by the demand of the interactional context, with a mixed-language (or "dual-language" as called by the authors) context requiring more control processes than a single-language context. As reviewed in the previous section, in the more demanding mixed-language block, enhanced controlling processes seem to be required mostly in order to avoid interference from the stronger into the weaker language. This is achieved by hampering the stronger language and/or facilitating the weaker one. Moreover, there are some indications that other context variations, such as giving speakers less or more time to prepare for the next trial, might affect the stronger and weaker languages differently. For example, in a cued language switching study comparing performance of unbalanced bilinguals in trials with and without preparation time, Fink and Goldrick (2015; Exp. 2) showed that the "without preparation" task was more costly for the stronger L1 than for the weaker L2. Interestingly, just like trials in the single- vs. the mixed-language block, also trials with vs. without preparation time could be seen as a less demanding vs. a more demanding task. This is because, while in with-preparation trials the language cue is encoded first followed by the picture to be named, in the without-preparation trials both stimuli have to be processed simultaneously, making the latter more demanding than the former. Therefore, just as in the interactional context (e.g., single- vs. mixed-language), the manipulation of preparation time (e.g., with vs. without preparation time) also shows that more demanding tasks tend to be more effortful for the stronger than for the weaker language, presumably to equalize the proficiency difference between the two languages so as to reduce conflicts.

Altogether, this pattern of results emphasizes the key role that language dominance seems to play in language control: Not only is the stronger language more difficult to reactivate than a weaker language during language switching, but it is also more hampered than the weaker counterpart when the task demand is enhanced.

In a study with unbalanced trilinguals, Philipp et al. (2007, Exp. 1) found that in each language pairing (L1-L2, L1-L3 and L2-L3) switching costs were larger for the stronger than the weaker language, supporting the idea that also in trilinguals, suppression of the non-relevant language depends on its dominance. In their study, the authors also manipulated the context in which the switching task took place. Specifically, in one block participants were given more time to prepare for the next trial, and in the other block they were given less preparation time. Lengthening preparation time is also a way to manipulate the difficulty of the task. In fact, in the case of longer preparation interval speakers have more time to reconfigure the next trial compared to shorter preparation time, making the former a less demanding task than the latter. Results showed that, compared to trials with longer preparation times, trials with shorter preparation times were more effortful for the stronger L1 than for the weaker L2 and L3. Therefore, similar to what was already found in bilingual studies, the context (less vs. more demanding) in which the language switch takes place affects the way languages are controlled. Interestingly, in Philipp et al. (2007), no difference was found between the stronger L2 and the weaker L3 in the longer vs. shorter preparation time condition. One reason for that might lie in the fact that while the L1 was a native language, both L2 and L3 were non-native languages, implying that the way languages are controlled might be determined by their status (native vs. non-native languages), rather than by their dominance (stronger vs. weaker language).

This raises the question of whether the difference usually found between the stronger L1 and the weaker L2 in bilingual studies comparing performance in more and less demanding tasks should be attributed more to the fact that often a native and a non-native language are compared in those studies, and less to the fact that a stronger and a weaker language are involved in the task. For example, it is possible that in a more demanding context, interference from the native language needs to be more strongly prevented than interference from a non-native language, and that this is achieved by lowering the activation level of the native language to a greater extent compared to a non-native language. This would lead to slower responses and more errors in the more vs. less demanding task for the native compared to a non-native language.

However, it should be noted that in Philipp et al., (2007), trilinguals' switching costs were measured in three separate experimental contexts, where only two languages at a time were used, namely L1-L2, L1-L3 and L2-L3. As proposed by Grosjean (1998, 2001), speakers are sensitive to the context and simply knowing that a certain language might become relevant will enhance the activation level of that language. Therefore, it might be supposed that in Philipp et al. (2007), trilinguals' performance was influenced by the fact that they were tested in a bilingual rather than in a trilingual context. This is because, while in a bilingual context only two languages have to be kept active, in the trilingual setting three languages have to be ready for use, making the former a less demanding task than the latter. Therefore, it might be the case that the mechanisms used to control trilinguals' languages in a bilingual context are different compared to those used in a trilingual context. Because of this, a more effective way to investigate trilingual language control is by including all three languages within a single task. Additionally, in contrast to bilingual studies, involving three languages creates an opportunity to disentangle the effect of language status (native vs. non-native language) from that of language dominance (stronger vs. weaker language) on language control. In particular, while in bilingual studies the stronger native language is usually compared to a weaker non-native language, in the trilingual case, the comparison between a stronger and a weaker language can involve a native vs. a non-native language (e.g., L1 vs. L2) as well as a non-native vs. another non-native language (L2 vs. L3).

Apart from the specifications of the interactional context (e.g., number of languages involved, time given to prepare), as Costa et al. (2006) observed, the way languages are controlled might also depend on the type of languages involved in the task. In particular, the authors suggested that the language control system should be more taxed when trying to keep apart more similar compared to less similar languages. This is because language control is a mechanism in which conflicts between potential responses need to be resolved, and conflict resolution between more similar representations is supposedly more demanding compared to less similar representations (e.g., Sigala & Logothesis, 2002). In this respect, there is growing evidence indicating that language similarity plays an important role, especially during L3 processing. According to the "language typological distance" hypothesis, language similarity determines which language is going to interfere the most during L3 processing. Specifically, it assumes that the language typologically closer to the L3 will become the main source of interference during L3 production. If the native L1 and the non-native L2 are equally close to the weaker L3, then the L2 will be the main source of interference (e.g., Cenoz, 2001, 2003; Hammarberg, 2001). In this respect, it has been suggested that during L3 production it might be easier to keep the native L1 separate than the non-native L2, because the L1 is perceived as qualitatively different from both the L2 and the L3. Moreover, when acquiring the second foreign language (L3), speakers might rely on the same strategies used to acquire the first non-native language (L2). This makes the L2 the default supplier language in the case of unsuccessful L3 processing, that is, the L2 becomes the main source of interference during L3 processing (Hammarberg, 2001). Considering this, it might be assumed that language similarity between the first and second non-native language (L2-L3) affects the way languages are controlled in multilinguals.

1.4 Alternative Specifications of the IC Model

As discussed in the previous sections, according to the IC model proposed by Green (1986, 1993, 1998), during lexical selection words from both the relevant and irrelevant languages are considered for selection, and competition between languages is solved by inhibiting the non-relevant language (but, see Finkbeiner et al., 2006; La Heij, 2005 for alternative explanations). In this view, language selection is *language non-specific*, meaning that words from both the relevant and the irrelevant language are activated and compete for selection, i.e. language coactivation *with* language competition. In contrast to this approach, Costa and Caramazza (1999) suggested that although both relevant and irrelevant languages are activated, there is no need to inhibit the irrelevant language since the controlling system can determine which lexicon to consult based on specific properties of the lexical nodes (e.g., grammatical class, language membership). In this view, language selection is *language-specific*, meaning that words from both the relevant and irrelevant languages are activated, but that only words from the intended language are considered for selection, i.e. language coactivation *without* language competition.

Interestingly, a more hybrid approach has suggested that both mechanisms might be possible, with a lexical selection process requiring inhibition of the non-relevant language in speakers with low L2 proficiency and a lexical selection process relying on specific properties of the language nodes in more advanced L2 speakers (Costa & Santesteban, 2004; Costa, Santesteban & Ivanova, 2006; Schwieter & Sunderman, 2008). Precisely, supporters of this account specified that with increasing L2 proficiency, the lexical selection process might shift from an inhibitory mechanism (language non-specific selection) to a mechanism where inhibition of the irrelevant language is not required, since only words of the relevant language are considered for selection (language-specific selection). Moreover, according to this view, once the selection process has shifted from a language non-specific (with inhibition of the non-target language) to a language-specific (without inhibition of the non-target language) mechanism, the latter will be used in other instances of language control, for example, when switching between the stronger L1 and the weaker L3 (Costa & Santesteban, 2004).

Within this approach, a way to determine whether language control relies on a language specific or non-specific mechanism is by looking at the switching cost pattern between languages. Specifically, larger costs for the stronger than for the weaker language (asymmetrical switching costs) have been taken as evidence for an inhibitory (language non-specific) mechanism of language selection. This is because, within this account, the stronger language is inhibited more than the weaker language, leading to larger reactivation costs for the former than for the latter. In contrast, a comparable amount of switching costs (symmetrical switching costs) between a stronger and a weaker language has been taken as an index of a language specific mechanism. In this framework, only words of the relevant language are considered for selection, so that words from the irrelevant language do not need to be inhibited. If inhibition does not take place, then no effects of persisting inhibition (modulated by language dominance) are expected and switching costs might reflect only the effort of reconfiguring a new task (not related to language dominance).

This language selection approach, according to which the language control mechanism is determined by L2 dominance, has been formulated after observing that while less proficient L2 speakers showed larger switching costs for the stronger L1 than for the weaker L2 as predicted by an inhibitory account of language selection, more proficient L2 speakers showed a similar amount of switching costs in the stronger L1 and the weaker L3 (Costa & Santesteban, 2004), as well as in the stronger L2 and the weaker L3 (Costa et al., 2006). However, it is worth noting that in those experiments, language switching costs involving the weaker L3 were measured in tasks where only two languages were tested at a time (i.e., L1-L3 and L2-L3 separately), instead of all three languages simultaneously. As discussed above, the context in which language switch takes place can influence the way languages are controlled. In the case of trilinguals, a dual-language setting can alter languages' activation level, in that more activation is sent to the two languages involved in the task and less to the one irrelevant for the task. Therefore, as already emphasized, a more promising way to investigate language control in trilinguals is by including all three languages in a single experimental block.

So far, language control has been discussed in relation to language production. In the following paragraph, the relatively underexplored notion of language control in recognition will be introduced and discussed in light of the IC model (Green, 1998) as well as of an alternative account specific to language recognition, namely the Bilingual Interactive Activation (BIA) model (e.g., Dijkstra & van Heuven, 1998). As a wealth of studies have demonstrated, when a word in a certain language is presented to a bilingual speaker, also words from the other language may activate and interfere with the recognition process (e.g., Dijkstra, Timmermans & Schriefers, 2000; Dijkstra, van Jaarsveld & ten Brinke, 1998; Lemhöfer & Dijkstra, 2004; Nas, 1983; van Heuven, Dijkstra & Grainger, 1998). In a series of experiments, van Heuven et al. (1998) demonstrated that the recognition speed of a visually presented word in a language can be affected by the number of similar words (called "neighbours") in another language. For example, the presentation of the English word *work* was affected by the number of Dutch words that were similar in spelling, such as *werk* (Dutch word for "work") and *vork* (Dutch word for "fork"). This indicates that during visual word recognition words from both the relevant and irrelevant languages are considered for selection and that the lexical selection system has to be able to select the appropriate word from among all possible candidates.

The ability to recognize a word in a specific language, while reducing competition from words of the other language, is known as language control in recognition.

Discovering which mechanisms underpin language control in recognition is a relatively unexplored area of research. Most of the studies on language control have, indeed, focussed on language production (e.g., Abutalebi & Green, 2007, 2008; Calabria, Hernández, Branzi, & Costa, 2012; Costa & Santesteban, 2004; Filippi, Karaminis & Thomas, 2014; Finkbeiner et al., 2006; Goldrick, Runnqvist & Costa, 2014; Gollan & Ferreira, 2009; Jackson et al., 2001; La Heij, 2005; Linck et al., 2012; Meuter & Allport, 1999), but less so on language recognition (e.g., Grainger & Beauvillain, 1987; Orfanidou & Sumner, 2005; Thomas & Allport, 2000; von Studnitz & Green, 1997; von Studnitz & Green, 2002; Wang, 2015). Moreover, language production and recognition have often been investigated separately, leaving unclear whether the two processes rely on the same or different mechanisms. Just as in language production, the IC model (e.g., Green, 1998) suggests that language control in recognition is achieved by inhibiting competing words of the other language and that the amount of inhibition applied to the stronger language is greater than to the weaker language. Therefore, as in language production, language switching costs in recognition are predicted to be larger for the stronger than for the weaker language. In this scenario, switching costs are *dominance-related*. Moreover, the way in which language control may work has been simulated with an algorithmic model of bilingual visual word recognition, the Bilingual Interactive Activation (BIA) model (Dijkstra & van Heuven, 1998; Grainger & Dijkstra, 1992; van Heuven et al., 1998). Importantly, this model was not explicitly implemented to account for language control in recognition, but rather to investigate how bilinguals recognize words from their two languages. According to this model, when a word is presented to a bilingual speaker, both words from both the target and the non-target language are activated. The activated words send activation to the respective language node (representational layer specifying language membership). More strongly activated words (e.g., words from the relevant language) will send more activation to their language node compared to less activated words (e.g., words from the irrelevant language). Language competition is solved via inhibition from the language node to the words of the other language. The amount of inhibition applied to the words of the other language depends on the strength of activation of the language node. Specifically, the stronger the activation of the language node (e.g., language node of the relevant language) the greater the inhibition applied to the words of the other language (e.g., language node of the irrelevant language), that is, “asymmetric inhibition”. Therefore, the main function of the language nodes is to collect activation from the words in the language they represent and inhibit active words from the other language. In this way, the word candidate that best matches the input will become most active and is recognized as soon as its activation level exceeds a certain threshold level. More generally, the BIA suggests that the activation of the language node reflects the amount of activity in the lexicon. Since L1 words have a higher baseline activation level than L2 words, the L1 language node is more activated than the L2 language node, and inhibition is greater on L2 than on L1 words (Dijkstra & van Heuven, 1998; van Heuven, Grainger & Dijkstra, 1998). Based on these assumptions, it can be postulated that if inhibition is asymmetric, namely greater on L2 words than on L1 words, then also the costs required to reactivate the two languages should be asymmetric. Specifically, switching costs for the more inhibited L2 words should be larger compared to the less inhibited L1 words. In this view, switching costs are expected to be *dominance-reversed* (the weaker the language the larger the costs).

Whether language control in bilingual visual word recognition can be better accounted for by the BIA model (expecting larger switching costs for the weaker than for the stronger language) or by the IC model (predicting larger switching costs for the stronger than for the weaker language) is hard to define based on previous findings. Indeed, past research on language control in bilingual recognition has not found any reliable difference in the amount of switching costs between the stronger and the weaker language (Orfanidou & Sumner, 2005; Thomas & Allport, 2000; von Studnitz & Green, 1997; Macizo et al., 2012; von Studnitz & Green, 2002; but, see Jackson et al., 2004). This has led to the question of whether different amount of switching costs for the stronger and the weaker language can be observed only in production, but not in recognition (Reynolds, Schlöffel & Peressotti, 2016). If this is the case, this would imply that language dominance affects language control in production (larger switching costs for the stronger than for the weaker language), but not in language recognition (comparable switching costs for the weaker and the stronger language). According to Reynolds and colleagues (2016), the difference usually found between the two modalities could be attributed to the fact that language production and recognition differ in terms of the specific systems required to perform the task. For example, in picture naming, after the semantic information of the picture has been retrieved, its phonological representation has to be activated, as to allow its articulation. In this sense, language production relies strongly on phonological encoding. However, during visual word recognition, most of the attention is devoted to decoding the orthographic form of the word. After the corresponding word has been found in the lexicon, activation is sent to its semantic information. In this sense, visual word recognition predominantly relies on orthographic decoding. Since phonological encoding is supposed to be a more demanding process compared to orthographic decoding (Reynolds & Besner, 2006), the former is expected to be more susceptible to interference from the other language than the latter. In this scenario, language production is more sensitive to languages' activation level than language recognition is. Even though the details of the two processes are not specified, the authors suggest that this would lead to larger switching costs for the stronger than for the weaker language in language production and to a comparable amount of switching costs in language recognition. Therefore, it has been suggested that language control in production and recognition is supported by different mechanisms. However, it should be noted that most of the studies investigating language control have looked at language production and language recognition separately, so that hardly any conclusion can be drawn as to whether or not language control in production and recognition relies on the same mechanisms. Finally, it is worth noting that most of the studies on visual word recognition were not explicitly set up to look at the effect of language dominance on language control. For example, some studies manipulated orthographic specificity (Orfanidou & Sumner, 2005; Thomas & Allport, 2000), frequency (von Studnitz & Green, 1997) or animacy (von Studnitz & Green, 2002) of the presented words. Therefore, methodological inconsistencies across studies make it difficult to infer whether language dominance modulates the way languages are controlled in visual word recognition.

1.5 Overarching Aim and Objectives of the Dissertation

This dissertation is devoted to the investigation of the assumptions made by the most popular model of language control, the Inhibitory Control (IC) model (Green, 1998). In addition to the IC model, two alternative specifications about the mechanisms underlying language control are also considered throughout the thesis. The main assumptions made by the IC model and the two alternative suggestions are summarized as follows: (1) According to the IC model, during selection words from both the relevant and irrelevant languages compete for selection. Competition is solved by inhibiting words from the irrelevant language. The amount of inhibition applied depends on the language's dominance, with the stronger language being more inhibited than the weaker language, i.e. dominance-related inhibition. This has been postulated for both (1a) language production and (1b) recognition. As an extension to the IC model for language production, (2a) the language-specific account argues that the language selection mechanism is modulated by the dominance reached in the L2. In particular, when the dominance level in the L2 is relatively low, relevant and irrelevant languages compete for selection and competition is solved via inhibition of the irrelevant language. Just as for the IC model, inhibition is dominance-related, namely greater on the stronger than on the weaker language. However, when the speaker has reached a relatively high dominance level in the L2, only words of the relevant language are considered for selection so that no inhibition of the irrelevant language is required. As an alternative suggestion to the IC model for language recognition, (2b) the BIA model argues that words of both the relevant and non-relevant languages compete for selection. Competition is solved via suppressing words from the non-relevant language. The amount of suppression depends on the language's dominance, in the sense that the stronger language inhibits words of the weaker language to a greater extent than otherwise. Hence, inhibition here is dominance-reversed. Altogether, these models suggest that language control is influenced by language dominance. However, evidence from previous studies indicates that language control might be affected by other factors than just language dominance. The overarching aim of the present dissertation is to investigate how multilingual speakers control their languages during language switching and which factors influence this mechanism. To address this issue, special attention is given to the following research questions:

1) To what extent does **preparation time** affect language control?
2) Is language control influenced by **language typology**?
3) Does the **modality** (production vs. recognition) in which languages need to be controlled influence this process?

These questions are addressed in the four publications presented in this dissertation:

Publication I examines the effect of **preparation time** on bilingual language control in production. In particular, it looks at both transient and sustained control of language, as reflected by the language switching and mixing costs, respectively. To do it, language switching and mixing costs were measured during a cued bilingual picture naming task involving preparation time manipulation (with vs. without preparation time). Participants of this study are unbalanced bilinguals with a stronger native language (L1) and a weaker non-native language (L2). This study is based on the theoretical framework offered by the IC model, according to which both relevant and non-relevant languages compete for selection and competition is solved via inhibition of the non-relevant language. In this framework, the amount of inhibition is dominance-related (the stronger the language the greater the inhibition applied to it).

Publication II looks at the effect of **preparation time** on trilingual language control in production. In particular, it examines trilingual transient and sustained control of language as reflected by trial- and block-related costs. To do this, unbalanced trilinguals are tested in a trilingual digit naming task involving preparation time manipulation between blocks (shorter vs. longer preparation time block). For each language, costs at the trial level (language switching costs) and costs at the block level (shorter vs. longer preparation time block) are measured. Participants of the study have a stronger native language (L1), an intermediately strong first non-native language (L2) and a weaker second non-native language (L3). Just as for the bilingual study described above, this study is also founded on the theoretical framework of the IC model. Specifically, it assumes that all three languages compete for selection and that selection of the relevant language is achieved by suppressing the two non-relevant languages. The amount of suppression depends on language dominance, with stronger languages being inhibited more than weaker languages.

Publication III examines whether trilingual language control is affected by **typological closeness** to the L3. With this aims, language switching costs are measured in two groups of unbalanced trilinguals performing a trilingual picture naming task. Participants have a stronger native language (L1), a relatively strong first non-native language (L2) and a weaker second non-native language (L3). For one group, the L3 is typologically closer to the L2, and for the other group the L3 is typologically more distant from both the L1 and the L2. In this study, both the IC model and its extension, the language specific account, are considered for language control in production. While according to the IC model relevant and non-relevant languages are always considered for selection, according to the language specific account the way in which languages are controlled depends on the dominance level reached in the L2. Specifically, it assumes that when the dominance level in the weaker L2 is relatively low, both the relevant and the irrelevant language are considered for selection and that competition from the irrelevant language is solved via dominance-related inhibition (as the IC model predicts). However, when the dominance level in the L2 is relatively high, only words from the relevant language are considered for selection, so that no inhibition of the irrelevant language is needed. In this framework, the only detectable cost is the one related to the effort of changing the task-goal in the case of language switch. The cost of re-setting the cognitive system for the new task does not depend on language dominance and is, therefore, predicted to be the same for the stronger L1 and the relatively strong L2.

Finally, Publication IV investigates whether language control is affected by the processing **modality** (production vs. recognition). To do this, it firstly investigates the mechanisms underpinning bilingual language control in recognition. Specifically, it examines whether language switching costs in bilingual visual word recognition depend on dominance and if they do, whether language switching costs are dominance-related (the stronger the language the greater the cost) as the IC model predicts, or dominance-reversed (the stronger the language the weaker the cost) as expected from the alternative view of language control offered by the BIA model. Secondly, this study assesses whether language control in production and recognition relies on the same or different mechanisms. To test these two issues, language switching costs are measured in a group of unbalanced bilinguals with a stronger native language (L1) and a weaker non-native language (L2), while performing a recognition (lexical decision) task and a production (picture naming) task.

A detailed overview of the four publications together with a summary of the main results obtained is provided in the next section.

2 Overview of the Publications

The following four chapters present the results obtained within this dissertation in the form of four manuscripts, either published or currently under review. An overview and a summary of the four manuscripts can be found below.

Publication I (First author; published in *Bilingualism: Language and Cognition*, 2016, doi: 10.1017/S1366728915000693)

Examining language switching in bilinguals: The role of preparation time

Authors: Michela Mosca[a,b] and Harald Clahsen[a]

Summary: This study explored the effect of preparation time on bilingual language control during language switching. To this end, late unbalanced bilinguals (L1 German – L2 English) were tested in a bilingual picture naming task involving preparation time manipulation. Two different aspects of language control were taken into consideration: (i) a temporary or TRANSIENT control and (ii) a more prolonged or SUSTAINED control. Transient control was investigated by measuring language *switching costs*, that is, the difference in performance between repetition trials (trials preceded by stimuli in the same language) and switch trials (trials preceded by stimuli in a different language). Sustained control was examined by measuring language *mixing costs*, namely the difference in performance between trials in the single-language block (where pictures had to be named in one language at a time) and the mixed-language block (where pictures had to be named in either L1 or L2). Results showed that both transient and sustained control were affected by preparation time. For transient control, the study showed that when speakers were not informed beforehand about the language to be used in the upcoming trial, language switching was costly. However, when speakers were given time to prepare, language switching became cost-free. As for sustained control, the results revealed that, compared to the single-language block, in the mixed-language block the weaker L2 was facilitated (L2 mixing benefits) in both preparation time conditions, whilst the stronger L1 was hampered when no preparation time was given (L1 mixing costs) but not when time to prepare was provided. Overall, these findings suggest that bilinguals can prepare for language switch and that during language switching languages' availability is affected by preparation time.

Personal Contribution: I contributed to conceiving the experimental set-up and I designed and programmed the experiment. I also acquired, analysed and interpreted the data. Furthermore, I wrote the first draft and the final version of the manuscript.

Co-author's Contribution: Harald Clahsen was involved in the conception of the experimental set-up as well as in editing and finalizing the manuscript.

Publication II (Co-authorship; published in *Cognitive Control and Consequences of Multilingualism*, 2016, doi: 10.1075/bpa.2)

Influence of preparation time on language control: A trilingual digit naming study

Authors: Julia Festman[c] and Michela Mosca[a,b,c]

Summary: The goal of the study was to investigate the effect of preparation time on trilingual language control. With this aim, late unbalanced trilinguals (L1 English, L2 French/German, and L3 German/French) were tested in a trilingual digit naming task featuring preparation time manipulation (shorter vs. longer preparation time). Two types of controlling mechanisms were examined: (i) a transient or “trial-related” control and (ii) a sustained or “block-related” control. To assess the temporary (or trial-related) control of languages, language switching costs were measured. This was done by comparing performance between repetition trials (digits had to be named in the same languages as in the preceding trial) and switch trials (digits had to be named in a different language compared to the preceding trial). To investigate the prolonged (or block-related) control of languages, we compared performance levels between the shorter and longer preparation time conditions. Results revealed that both types of control were affected by preparation time. With regard to the temporary or trial-related control, it was shown that when trilinguals were given less time to prepare for the upcoming trial, language switching was a costly process. In contrast, when speakers were given more time to prepare, language switching costs disappeared. As for the prolonged or block-related control, the study showed that in the case of shorter preparation time performance in each language was hampered compared to a longer preparation time condition. Importantly, the disadvantage increased along with language dominance. Altogether, these outcomes reveal that trilinguals can prepare for a language switch and that in more complex situations (such as in the case of restricted preparation time), stronger languages are made less available than weaker languages. This latter finding was interpreted as a way to strategically foster the weaker languages in a language switching context.

Personal Contribution: I analysed the data and wrote the methodological part of the final publication. I contributed to interpreting the data and in writing the introduction, discussion and conclusion of the manuscript.

Co-author’s Contribution: Julia Festman conceived, set up, programmed and ran the experiment. She was involved in the interpretation of the data and wrote the final manuscript.

Publication III (Single-author; under review with the *Quarterly Journal of Experimental Psychology*)

Trilinguals' language switching: A strategic and flexible account

Author: Michela Mosca[a,b]

Summary: The objective of the study was to determine whether trilingual language control is affected by typological closeness between languages (more vs. less similar languages). To investigate this, language switching costs were measured in two groups of late unbalanced trilinguals performing a picture naming task that involved switching in three languages (English, Italian, German). The first group (Exp. 1) were native speakers of Italian (L1), highly proficient speakers of German (L2), and learners of English (L3). The second group (Exp. 2) were native speakers of German (L1), highly proficient speakers of English (L2), and learners of Italian (L3). The study also aimed to assess whether during language switching only words from the relevant language are considered for selection (selection is *language specific*) or whether words from both the relevant and irrelevant languages compete for selection (selection is *language non-specific*).

Results revealed that, in both groups, relevant and irrelevant languages competed for selection, as indicated by the cross-language interference found in all three languages. Moreover, the study indicated that typological closeness (more vs. less similar languages) influenced the way languages were controlled in a trilingual switching context. Specifically, the study suggested that the controlling system aimed at reducing unwanted interference from the stronger into the weaker typologically closer language, by hampering the stronger "disturbing" language and/or facilitating the weaker "disturbed" one. Overall, the study puts forward the hypothesis that language control is a flexible and strategic mechanism that seeks to avoid conflicts between languages.

Personal Contribution: I conceived, set up, programmed and ran the experiment. I also acquired, analysed and interpreted the data. Finally, I wrote and edited the first draft as well as the final version of the manuscript.

Publication IV (First author; under review with *Frontiers in Psychology*)

Bilingual Language Switching: Production vs. Recognition

Authors: Michela Mosca [a,b] and Kees de Bot [d]

Summary: This study investigated the mechanisms underpinning bilingual language control in recognition and production. To do this, language switching costs were measured in a group of late unbalanced bilinguals (L1 Dutch – L2 English) performing a bilingual lexical decision task (Exp. 1) and a bilingual picture naming task (Exp. 2). Results indicated that language control in bilinguals was influenced by modality, that is, the way the languages were controlled differed between recognition and production contexts. While in recognition bilingual language control seemed to be determined by the properties of the stimuli (e.g., frequency), in production the way in which languages were controlled appeared to be modulated by speakers' unconscious strategies for minimizing conflicts between languages. Generally, we suggested that the difference between language control in recognition and production could be attributed to the fact that the two modalities are mostly supported by different controlling mechanisms. In particular, we argued that language recognition might be mainly supported by a more stable *bottom-up* mechanism, whereas language production might be primarily maintained by a more flexible *top-down* mechanism. Whilst in the former, attention cannot be conveniently shifted, in the latter, attention can be used strategically to optimize performance.

Personal Contribution: I contributed to conceiving the experimental set-up and I designed and programmed the experiment. I also acquired, analysed and interpreted the data. Furthermore, I wrote the first draft and the final version of manuscript.

Co-author's Contribution: Kees de Bot was involved in the conception of the experimental set-up as well as in editing and finalizing the manuscript.

Authors' Affiliations:

[a] Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam, 14476, Potsdam, Germany

[b] International Doctorate in Experimental Approaches to Language And Brain (IDEALAB), University of Potsdam, 14476, Potsdam, Germany; University of Groningen, 9712 EK, Groningen, the Netherlands; University of Trento, 38122, Trento, Italy; Newcastle University, NE1 7RU, Newcastle Upon Tyne, United Kingdom and Macquarie University, NSW 2109, Sydney, Australia

[c] University of Potsdam, 14476, Potsdam, Germany

[d] University of Pannonia, 8200, Veszprém, Hungary

3 Publication I

Published in: *Bilingualism: Language and Cognition* (2016)

Examining language switching in bilinguals: The role of preparation time

Michela Mosca[a,b] and Harald Clahsen[a]

[a] Potsdam Research Institute for Multilingualism (PRIM),University of Potsdam, Germany;
[b] International Doctorate in Experimental Approaches to Language And the Brain (IDEALAB), University of Potsdam, Germany; University of Groningen, the Netherlands; University of Trento, Italy; Newcastle University, UK; and Macquarie University, Australia.

Abstract

Much research on language control in bilinguals has relied on the interpretation of the costs of switching between two languages. Of the two types of costs that are linked to language control, *switching costs* are assumed to be transient in nature and modulated by trial-specific manipulations (e.g., by preparation time), while *mixing costs* are supposed to be more stable and less affected by trial-specific manipulations. The present study investigated the effect of preparation time on switching and mixing costs, revealing that both types of costs can be influenced by trial-specific manipulations.

Keywords:

Bilingual language switching, preparation time, switching costs, mixing costs, picture naming

Introduction

One of the most astonishing skills of a fluent bilingual or trilingual person is the ability to switch between different languages. The costs associated with language switching have been the centre of many studies investigating multilinguals' lexical retrieval. Two kinds of costs are usually linked to language switching tasks, costs for language SWITCHING and for language MIXING. A common technique to measure LANGUAGE SWITCHING COSTS is to compare participants' performance in a task in which they have to switch from one language to another ("switch trial") to a task in which they stay in the same language ("repetition trial"). Performance in switch trials has been found to be slower and more error-prone than in repetition trials. The reaction time (RT) difference between repetition and switch trials is called "switching costs" (e.g., Roger & Monsell, 1995). The experimental technique may also include SINGLE language blocks (in which stimuli from two or more languages are tested in separate blocks) as opposed to MIXED language blocks (in which stimuli from more than one language are intermixed). For each language, MIXING COSTS are measured as the difference in performance between trials in the single-language block and repetition trials in the mixed-language block (e.g., Rubin & Meiran, 2005). While in the single- language block only one language is active, in the mixed-language block more than one language needs to be maintained active. These two types of costs reflect different cognitive control processes (e.g., Koch, Prinz & Allport, 2005). While switching costs are believed to reflect the effort involved in configuring an upcoming task or trial, a momentary process supported by a 'transient' control mechanism, mixing costs are supposed to reflect a prolonged or 'sustained' control process of maintaining multiple languages active in the mixed compared to the single language block (Braver, Reynolds & Donaldson, 2003). In this way, we can distinguish between trial-specific versus not-trial-specific costs of language control, the former involved in switching costs and the latter in mixing costs (Braver et al., 2003). Consequently, trial-level manipulations, e.g., through different preparation times, can be expected to affect switching costs, but less so or not at all mixing costs. The current study sheds new light on these processes by investigating potential effects of different preparation times in bilingual language switching.

Several previous studies have found that switching costs are influenced by participant-level factors (e.g., language proficiency) as well by task-related factors (e.g., stimulus properties); see Bobb & Wodniecka (2013) for a review. The momentary nature of switching costs signalling transient control is confirmed by studies showing that when speakers are given a longer interval between the language cue and the stimulus ('cue-stimulus interval', CSI), switching costs decrease (e.g., Costa & Santesteban, 2004). This indicates that the earlier presentation of the cue can boost preparation for the upcoming trial (Meiran, 2000). Additionally, there is some evidence suggesting that inter-trial intervals (ITI) can also affect performance in task/language switching studies such that longer intervals between response and cue (RCI) were, for example, found to speed up reaction times (Koch & Allport, 2006; Philipp, Gade & Koch, 2007). This could be due to reduced passive interference from the previous trial, which may lead to smaller switching costs (Allport et al., 1994). In previous studies, the interval between trials (such as ITI or RCI) was either relatively short (e.g., 400 ms in Declerck, Koch & Philipp, 2012; 1150 ms in Costa & Santesteban, 2004) or left uncontrolled (i.e., variable from 1500 ms to 2300 ms in Verhoef, Roelofs & Chwilla, 2009; variable from 1000 ms to 1250 ms in Fink & Goldrick, 2015; 100 ms RCI in long CSI condition vs. 1000 ms RCI in short CSI condition in Philipp et al., 2007). From these studies, the potential effects of active preparation on transient control processes involved in language switching are hard to determine.

Moreover, while there is agreement on the beneficial effect of preparation time on switching costs, it is not clear whether or not preparation time also affects mixing costs. If mixing costs reflect a stable process of maintaining two or more languages active in the mixed-language block, we may expect that this process is not affected by any kind of task (viz. preparation time) manipulation. However, previous studies have yielded inconsistent findings regarding mixing costs in language switching. Some studies found mixing costs in both the L1 and the L2 (e.g., Prior & Gollan, 2011), others only in the L1 but not in the L2 (e.g., Christoffels, Firk & Schiller, 2007) and yet other studies obtained a “mixing benefit”, i.e., faster responses (in the L2, but not the L1) for the mixed-language than the single-language block (e.g., Hernandez, Dapretto, Mazziotta & Bookheimer, 2001). Furthermore, inconsistencies across studies could also be due to the fact that different types of bilinguals have been tested in language switching studies (e.g., early bilinguals: Prior & Gollan, 2011; late bilinguals: Christoffels et al., 2007; L2-dominant bilinguals: Hernandez et al., 2001). Finally, different time manipulations have been used in previous studies (e.g., 200 ms CSI and fixed ITI for Hernandez et al., 2001; 0 ms CSI and variable ITI in Christoffels et al., 2007, and 250 ms CSI and fixed RCI in Prior & Gollan, 2011) so that the question remains of how task manipulations, specifically trial-level differences in preparation time, affect the sustained control processes involved in language switching tasks.

The goal of the present study is to investigate the effect of preparation time on both transient and sustained control, by measuring switching and mixing costs in a bilingual picture naming task. To minimize the effect of passive interference and principally focus on that of preparation time, we compared performances in trials with and without preparation time, while using a relatively long and fixed ITI. Following the above-mentioned distinction between trial-specific vs. non trial-specific costs (i.e., switching vs. mixing costs; see Braves et al., 2003), we expect to find an effect of a trial-specific manipulation (of preparation time), specifically reduced switching (but not mixing) costs in trials with preparation time compared to trials without preparation time.

Materials

Eighteen pictures were selected from the Colorized Snodgrass and Vanderwart object set (Rossion & Pourtois, 2004) to be named in English and/or German. Pictures had a size of 197x281 pixel and were presented at the centre of a 15-inch computer screen set to 1280x800 pixel resolution. Stimuli were seen from a distance of approximately 80 cm. DMDX (Forster & Forster, 2003) was used for stimulus presentation and CheckVocal (Protopapas, 2007) for recording and measuring speech-onset latencies. See Appendix A for detailed information on the items used.

Participants

Thirty participants (11 males, 19 females, mean age: 25.6, SD: 5.26) were recruited from the student population of the University of Potsdam and tested in German and English. Participants were all university educated, right-handed and had normal or corrected to normal vision. They all gave their consent before the experiment and were paid or given course credit for their participation. All participants acquired German from birth as their sole native language (L1) and English as second language (L2) at school for a minimum of 5 years with an average age of onset of 9.55 (SD: 1.58). See Appendix B for detailed information on participants.

Procedure

Participants were seated in front of a computer screen and instructed to name each picture displayed on the computer screen either in their L1 or their L2 as quickly and accurately as possible. The language to be used was indicated by the colour of the screen background (blue=L1, red=L2). A with-preparation trial consisted of (i) a language cue (for 500 ms on red or blue background), (ii) a blank screen for (300 ms), (iii) a picture (for 1500 ms), (iv) a blank screen (for 2400 ms). A no-preparation trial entailed (i) a fixation point (for 500 ms), (ii) a blank screen (for 300 ms), (iii) a picture together with a language cue (for 1500 ms), (iv) a blank screen (for 2400 ms). Thus, both no-preparation and with-preparation trials had a constant duration of 4700 ms and different cuing time, namely CSI = 0 ms and CSI = 800 ms respectively. Moreover, independently from subjects' response speed, pictures remained on the screen for a fixed duration of 1500 ms.

Each participant completed one experimental session, which included a single followed by a mixed-language block. In the SINGLE-LANGUAGE BLOCK, participants named stimuli in the L1 and the L2 separately. Participants named a set of 36 pictures in the L1 and a set of 36 pictures in the L2, in a counterbalanced order across participants. In each language-set, the first half of the items were with-preparation trials and the second half no-preparation trials; this was also counterbalanced across participants. The presentation of the stimuli was fully randomized and each picture was seen once in each of the four conditions (L1, L2, with-preparation, no-preparation). In addition to the variables ‘Language ‘(L1 vs. L2) and ‘Presentation Type’ (with-preparation vs. no-preparation), the MIXED-LANGUAGE BLOCK also included the variable ‘Trial Type’ (no-switch vs. switch). In a no-switch trial, a given picture had to be named in the same language as the previous one and in a switch trial in a different language than the previous one. Trials were grouped such that 75% was no-switch and 25% switch trials, e.g., L1-L1-L1-L2 in which case three consecutive pictures had to be named in the L1 and one in the L2. There were 144 trials (108 no-switch and 36 switch trials) in the mixed-language block, presented half in the with-preparation and half in the no-preparation Presentation Type. The same 18 pictures as for the single-language block were used in the mixed-language block, nine for the L1 and nine for the L2, presented eight times each. Two presentation lists of pseudo-randomized trials were created of which each participant saw only one. One list had the with-preparation trials first, whereas the other started with the no-preparation trials, making cuing display predictable. Likewise, the order of the two languages (German, English) for naming pictures was also counterbalanced between the two lists. Within each list, pictures were never repeated within five trials. Furthermore, the same type of chunk pattern (e.g., L1-L1-L1-L2) did not appear more than twice in a row. Due to these precautions, participants were unable to anticipate the order of the background colour.

Prior to the experiment, participants were familiarized with the procedures using six practice trials for the single-language block and eight for the mixed one.

Data coding and analysis

The dependent variables were participants’ accuracy and picture naming response times (RT), the latter measured from the display of the target picture until speech onset. Data from four participants and two items (*Kürbis* ‘pumpkin’ in the single-language block, and *Uhr* ’watch’ in the mixed one) were excluded from any further analysis due to low accuracy rates of less than 70%. For the remaining 26 participants, RTs and accuracy scores were calculated. Prior to the RT analysis, trials with incorrect responses, hesitations and cases in which the microphone was mis-triggered (e.g., through coughs or stuttering) were excluded (5.4% of the data). Trials with RTs faster than 350 ms as well as those slower than 2,000 ms (0.26% of the data), were treated as extreme values and also removed from the RT analysis. Due to these exclusions, the total amounts of removed data were 4.7% and 5% of the L1 responses, 6.5% and 5.5% of the L2 responses, the former in the single and the latter in the mixed language block.

To analyse the data statistically, mixed-effects linear regression models were fitted to the RT data and generalized linear models with a binomial link function (Cnaan, Laird & Slasor, 1997; Guo & Zhao, 2000) to the accuracy data. See Appendix C for detailed information on data analysis.

Results

Table 1 shows mean RTs and accuracy scores for the different experimental conditions. Tables 2 and 3 present the results of the statistical analyses.

Consider first the accuracy data from Table 1 and the corresponding statistical results in Table 2. In the single-language block, accuracy rates were significantly higher for the L1 than the L2 and for the 'no-preparation' Presentation Type than for the 'with-preparation' one; see the main effects of Language and Presentation Type in Table 2. There were no further main effects or interactions. In the mixed-language block, accuracy rates were similar across conditions without any reliable main effects or interactions.

Consider next the RT data. In the single-language block naming latencies in the L1 were significantly faster than in the L2 (741 ms vs. 790 ms), while there were no differences between the two levels of Presentation Type, with and without preparation time (769 ms vs. 764 ms); see Table 1.

This contrast was confirmed by a main effect of Language, but not of Presentation Type in the single-language block; see Table 3a. These results show that in the single-language task, the Presentation Type manipulation did not affect naming latencies. This was different in the mixed-language block. While there was no reliable main effect of Language, with similar naming latencies for L1 and the L2, there were significant effects of Trial Type, with shorter RTs for no-switch than for switch trials (744 ms vs. 794 ms) and of Presentation Type, with shorter RTs for the with-preparation than for the no-preparation trials (731 ms vs. 804 ms). Most importantly, however, there was a significant interaction of Presentation Type and Trial Type in the mixed-language block ($p< .05$). To further examine this interaction, we split the data by Presentation Type; see Table 3c and Table 3d. While in the no-preparation trials, switch trials yielded significantly longer RTs than no-switch trials (760 ms vs. 855 ms, $p< .05$) in both the L1 and the L2, there was no reliable (switch vs. no-switch) contrast for the with-preparation trials (729 ms vs. 734 ms, $p= .79$), either in the L1 or in the L2. These results indicate that language-switching costs disappeared when participants were given time to prepare for the switch. Moreover, the best-fit models of both Presentation types required the exclusion of the Language and Trial Type interaction, indicating that in the no-preparation as well as in the with-preparation trials switching costs were symmetrical. Similarly, the lack of the three-way interaction for Language, Trial Type and Preparation Type in the model indicates all trials had comparable benefit from preparation time.

Table 1. Correct mean RTs (standard deviations in brackets) and accuracy rates (in percent), for L1 vs. L2, with-preparation vs. no-preparation trials, switch vs. no-switch trials, and single vs. mixed-language blocks. Switching costs for L1 and L2 (calculated as the difference between no-switch and switch trials) as well as mixing costs for L1 and L2 (calculated as the difference between single and mixed language block) are reported in italics.

	NO-PREPARATION			WITH-PREPARATION		
	L1	L2	Trial Type Mean	L1	L2	Trial Type Mean
Single-language block						
	735 ms (229)	791 ms (200)	764 ms (216)	746 ms (212)	790 ms (208)	769 ms (210)
Accuracy	96.6% (25)	93.4% (33)	95% (29)	94.2% (33)	93.7 (32)	93.8% (32)
Mixed-language block						
No-switch	791 ms (193)	729 ms (156)	760 ms (178)	731 ms (198)	727 ms (190)	729 ms (194)
Accuracy	95.9% (33)	95.7% (34)	95.8% (33)	94.8% (37)	94.8% (34)	94.8% (35)
Switch	877 ms (215)	834 ms (217)	855 ms (217)	754 ms (244)	716 ms (166)	734 ms (207)
Accuracy	95.4% (35)	92.1% (39)	93.9% (37)	93.7% (39)	94.7% (36)	94.2% (37)
Switching Costs	*86 ms*	*105 ms*		*23 ms*	*-11 ms*	
Mixing Costs	*56 ms*	*-62 ms*		*-15 ms*	*-63 ms*	
Language Mean	831 ms (208)	779 ms (195)	804 ms (203)	741 ms (221)	721 ms (178)	731 ms (200)
Accuracy	95.7% (34)	94% (37)		94.2% (38)	94.8% (35)	

Table 2. Estimated coefficients standard errors (SE) and z values from the best-fit generalized linear mixed-effects models for the accuracy data. Asterisks (*) indicate: p< .05 (*), p< .01 (**), p<. 001(***) and p< .0001 (****).

	ACCURACY		
	Estimate	E	-value
Single-language block			
Intercepts	36.16	9.097	3.97***
Language (L1 vs. L2)	8.313	2.942	2.82**
Presentation Type (no-preparation vs. with-preparation)	6.893	2.788	2.47*

Formula: Accuracy ~ Presentation Type + Language + (1 | subject) + (1| item)

Mixed-language block			
Intercepts	15.103	2.439	6.19****
Language (L1 vs. L2)	-1.828	1.303	-1.40
Trial Type (no-switch vs. switch)	0.775	0.755	1.02
Presentation Type (no-preparation vs. with-preparation)	0.266	0.730	0.36

Formula: Accuracy ~ Presentation Type + Language + Trial Type + (1+Language|subject) + (1|item)

Finally, we measured mixing costs, i.e., the difference between single-language trials and no-switch trials in the mixed-language block. Table 3 (e) reveals a main effect of Block Type (*p*< .001), with surprisingly faster RTs for the mixed-language than the single-language block (744 ms vs. 768 ms). We also found a significant interaction of Language and Block Type (*p*< .0001), with facilitation for the L2 compared to the L1 (-63 ms vs. 21 ms), as well as a significant interaction of Preparation Type and Block Type (*p*< .0001), revealing a facilitatory effect for the with-preparation trials (-42 ms) but not for the no-preparation trials (-6 ms). The three-way interaction of Language, Block Type and Presentation Type was also significant (p< .05). To examine this interaction, we split the data by Presentation Type. Results of both the no-preparation trials (Table 3f) and the with-preparation trials (Table 3g) showed a significant interaction of Language and Block Type (*p*< .0001 and p< .01 respectively). In the no-preparation trials (see Table 3h and Table 3i), we found that responses in the L1 were slower in the mixed than in the single-language block (55 ms mixing costs; *p*<.0001), whereas responses in the L2 were faster in the mixed than in the single-language block (62 ms facilitatory effect; p< .0001). In the with-preparation trials (see Table 3l and Table 3m), there was no significant effect of Block Type for the L1 (-16 ms, *p*= .10), whereas L2 responses were faster in the mixed compared to the single-language block (63 ms facilitatory effect; *p*< .0001).

Table 3. Estimated coefficients, standard errors (SE) and t values from the best-fit linear mixed effects models run on inversed-transformed RTs. Asterisks (*) indicate: p< .05 (*), p< .01 (**), p<. 001(***) and p< .0001 (****).

REACTION TIMES			
	Estimate	SE	t-value
(a) Single-language block - Overall model			
(Intercept)	1.387	0.033	41.44****
Language (L1 vs. L2)	0.053	0.022	2.42*
Presentation Type (no-preparation vs. with-preparation)	0.011	0.027	0.41
*Formula: RT~ Language*Presentation Type + (1 + Presentation Type| subject) + (1| item)*			
(b) Mixed-language block - Overall model			
(Intercept)	-1.379	0.028	-49.59****
Language (L1 vs. L2)	0.019	0.017	1.15
Trial Type (no-switch vs. switch)	-0.030	0.014	-2.11*
Presentation Type (no-preparation vs. with-preparation)	0.074	0.013	5.11****
Presentation Type*Trial Type	-0.036	0.013	-2.67*
*Formula: RT~ Language + Presentation Type*Trial Type + (1 + Language*Trial Type| subject) + (1| item)*			
(c) No-preparation			
(Intercept)	-1.302	0.029	-43.79****
Trial Type: switch vs. no-switch	- 0.069	0.017	-4.00****
Formula: RT~ Trial Type + (1 + Trial Type| subject) + (1| item)			
(d) With-preparation			
(Intercept)	-1.443	0.033	-42.95****
Trial Type: switch vs. no-switch	-0.006	0.023	0.26

Formula: RT~ Trial Type + (1 | subject) + (1| item)			
(e) Mixing costs – Overall model			
(Intercept)	1.392	0.033	41.75****
Language (L1 vs. L2)	-0.037	0.043	-0.86
Presentation Type (no-preparation vs. with-preparation	0.002	0.043	0.07
Block Type (single vs. mixed)	0.051	0.012	4.15****
Language*Presentation Type	-0.058	0.077	-0.75
Language*Block Type	0.174	0.024	7.08****
Presentation Type*Block Type	0.135	0-024	5.53****
Language*Presentation Type* Block Type	-0.120	0.048	-2.49*
*Formula: RT ~ Language * Procedure* Block + (1 + Procedure + Language | subject) + (1| item)*			
(f) Mixing costs – No preparation			
(Intercept)	738.66	16.77	44.04****
Language (L1 vs. L2)	-0.69	25.38	-0.03
Block (single vs. mixed)	12.13	9.64	1.26
Language*Block	-136.7	19.07	-7.17****
*Formula: RT ~ Language*Block Type + (1 + Language | subject) + (1| item)*			
(g) Mixing costs – With preparation			
(Intercept)	747.66	24.81	30.13****
Language (L1 vs. L2)	43.62	39.65	1.10
Block (single vs. mixed)	-63.79	18.86	-3.38**
Language*Block	-54.43	19.32	-2.81*

*Formula: RT ~ Language*Block + (1 + Block | subject) + (1| item)*			
(h) Mixing costs – No preparation L1			
(Intercept)	738.48	22.18	33.29****
Block (single vs. mixed)	82.41	15.15	5.44****
Formula: RT ~ Block + (1 | subject) + (1| item)			
(i) Mixing costs - No preparation L2			
(Intercept)	737.26	19.17	38.45****
Block (single vs. mixed)	-60.29	11.82	-5.10****
Formula: RT ~ Block + (1 | subject) + (1| item)			
(l) Mixing costs – With preparation L1			
(Intercept)	729.55	25.10	29.07****
Block (single vs. mixed)	-37.44	23.25	-1.61
Formula: RT ~ Block + (1+ Block | subject) + (1| item)			
(m) Mixing costs – With preparation L2			
(Intercept)	766.49	34.50	22.22****
Block (single vs. mixed)	-91.76	13.87	-6.61****
Formula: RT ~ Block + (1 | subject) + (1| item)			

Discussion

Investigating the role of preparation time in a bilingual picture naming task, we found symmetrical switching costs when highly proficient bilinguals had no time to prepare for the task, and no switching costs when participants were given 800 ms preparation time. Whilst symmetrical switching costs for highly proficient bilinguals have been consistently reported (e.g., Costa, Santesteban & Ivanova, 2006), complete dissipation of language switching costs is a novel finding. Table 4 presents a comparison of our findings with results from previous studies. As shown in Table 4, earlier studies have used similar or even longer preparation times, but shorter inter-stimulus or response-stimulus intervals.

Table 4. Overview of cued language switching studies. For each study information on the timing events are given: Cue-Stimulus Interval (CSI), Response-Cue Interval (RCI), Response-Stimulus Interval (RSI) and Inter-Trial Interval (ITI).

Study	Preparation	CSI	RCI	RSI	ITI	Switch costs
Costa & Santesteban (2004)	Absent	0 ms	1150 ms	1150 ms	–	Symmetric
	Short	500 ms		1650 ms		Symmetric
	Long	800 ms		1950 ms		Symmetric
Philipp et al. (2007)	Short	100 ms	1000 ms	1100 ms	_	Asymmetric
	Long	1000 ms	100 ms	1100 ms	_	Asymmetric
Verhoef et al. (2009)	Short	750 ms	Variable	Variable	1500 ms/ 2300 ms	Asymmetric
	Long	1500 ms				Symmetric
Declerck et al. (2012)	-	1000 ms	400 ms	1400 ms	_	Symmetric
Fink & Goldrick (2015)						
Exp. 1	Short	750 ms	Variable	Variable	1000 ms/ 1250 ms	Symmetric
	Long	1500 ms				Symmetric
Exp. 2	Absent	0 ms	Variable	Variable	1000 ms/ 1250 ms	Asymmetric
	Short	750 ms				Asymmetric
	Long	1500 ms				Asymmetric
This study	Absent	0 ms	Variable	Variable	2400 ms	Symmetric
	Present	800 ms				Absent

With respect to participants' accuracy of responses in the mixed-language condition, we found no effect of preparation time, language (L1 or L2) or trial type (switch vs. no-switch). This finding is in line with previous language-switching studies (e.g., Schwieter & Sunderman, 2008). As regards the response-time data in the mixed-language condition, we found a trend for L1 naming latencies to be slower than for the L2. Slower naming latencies for the L1 than for the weaker language (either L2 or L3) have been labelled a 'paradoxical language effect' (Christoffels et al., 2007; Verhoef, Roelofs & Chwilla, 2010). This effect has been attributed to the additional cost involved in globally inhibiting the L1 in a mixed-language context, to facilitate naming in the weaker language (Costa & Santesteban, 2004; Verhoef et al., 2009). However, because of methodological differences between studies (e.g., type of bilinguals, type of task and material used), the sources of the paradoxical language effect in language switching are still not fully understood.

Furthermore, we also found a preparation-time benefit in L1 no-switch trials, in line with Fink and Goldrick's (2015) findings and contra Verhoef et al.'s (2009) hypothesis that L1 repetition trials do not benefit from longer CSI. Moreover, we found mixing costs only in the L1 (for the no-preparation trials), whereas there was a mixing benefit in the L2, for both no-preparation and with-preparation trials. We suggest that these results reflect an adjustment of naming strategies depending on task-demands, in order to successfully perform the tasks. Specifically, whilst in with-preparation trials the language cue is encoded first followed by the picture to be named, in the no-preparation trials both stimuli have to be processed simultaneously, making no-preparation trials more demanding than with-preparation trials. In case of bilingual language switching, the most challenging condition, the 'worst case' in Los' (1996) terms, is naming in the L2. Consequently, the speaker might devote more attention to the weaker L2 and less attention to the stronger L1, particularly in tasks that require more attentional resources. This strategy may yield a mixing BENEFIT in the L2 and a mixing COST in the L1 for the more demanding no-preparation trials. In the less demanding with-preparation trials, however, there are no mixing costs in the L1, but still a benefit in the L2. Mixing benefits rather than mixing costs for the L2 have previously been obtained by Hernandez et al. (2001). Similar to Hernandez et al.'s (2001) study, RCI and RSI in the present study have a variable duration. We suppose that, compared to studies with fixed RCI and RSI (e.g., Prior & Gollan, 2011), unpredictable RCI and RSI might enhance the level of task uncertainty, and thus of task demand, boosting facilitation of what is unconsciously perceived as the most difficult situation, i.e., naming in the L2.

Overall, these results suggest that mixing costs are not a mere reflection of the global costs of maintaining two or more languages active, but that they rather reflect unconscious adjustments to the task. Consequently, mixing costs are also flexible in nature and can be modulated by trial-specific manipulations, such as preparation time.

To conclude, our study reveals that both transient and sustained control processes are affected by preparation time. With regard to transient control, our results show that the cognitive system is able to fully prepare for the upcoming trial, challenging the view that it is impossible to completely eliminate switching costs (e.g., Rogers & Monsell, 1995). We suggest that this is due to the relative long ITI used in the present study, which together with a preparation time of 800 ms allows for the completion of the previous task and as a result for the system to prepare for the new one. This supports the hypothesis that advanced preparation can be fulfilled before the stimulus is presented (e.g., Monsell, 2003, yet see Mayr & Kliegl, 2003). However, we acknowledge that the question of how preparation and inter-trial times affect bilingual language switching costs requires further study. In particular, the degree of overlap between passive decay and active preparation involved in modulating switching costs needs to be precisely determined. Moreover, the fact that in the present study the response-stimulus intervals were variable may have affected naming latencies and needs to be controlled for in future studies. With reference to sustained control, we found that it was also affected by a trial-specific manipulations. This undermines the idea that mixing costs are a mere reflection of the global costs of maintaining two tasks active in memory. Instead, mixing costs (like switching costs) reflect strategies speakers rely on during language switching tasks (for a review see Festman & Schwieter, 2015). We suggest that mixing costs are involuntary adjustments to a given task and are therefore affected by task-specific manipulations. Further investigation is needed to clarify not only how these strategies work but also how they are influenced by participant-level factors, specifically by bilinguals' language proficiency.

Appendix A: *Materials* (L1 German, L2 English):

Eighteen pictures were selected from the Colorized Snodgrass and Vanderwart object set (Rossion & Pourtois, 2004) to be named in English and/or German. Items were matched according to conceptual complexity, word length (letters), lemma frequency, cognateness and semantic category using the International picture naming project (IPNP) database (Bates, D'Amico, Jacobsen, Szekely, Andonova, Devescovi, Herron, Lu, Pechmann, Pleh, Wicha, Federmeier, Gerdjikova, Gutierrez, Hung, Hsu, Iyer, Kohnert, Mehotcheva, Orozco-Figueroa, Tzeng & Tzeng, 2003). One-way ANOVAs revealed that there were no statistical differences for the test words in the two languages, neither with respect to lemma frequency (English: 3.19 (SD: 1.49) vs. German: 2.88 (SD: 1.75), p=.36) nor for word length (English: 5.3 (SD: 1.5) vs. German: 5.8 (SD: 2.5), p=.53). All the chosen pictures were classified as conceptually simple (conceptual complexity variable = 1); for details see Bates et al. (2003). Moreover, pictures denoting cognates or homophones in English and German were not selected. Finally, to avoid cumulative semantic interference effects (Howard et al., 2006), we selected pictures belonging to different semantic categories.

List of the items used:

Baum, tree; Besen, broom; Blatt, leaf; Gürtel, belt; Glocke, bell; Kette, necklace; Kleid, dress; Kürbis, pumpkin; Löffel, spoon; Pfeil, arrow; Pilz, mushroom; Rad, wheel; Schmetterling, butterfly; Stuhl, chair; Tür, door; Uhr, watch; Weintraube, grapes; Zwiebel, onion

Appendix B: *Participants*

All participants were native speakers of German (L1), late learners of English (L2). Their L2 proficiency level was assessed at the beginning of each experimental session using the grammar part of the paper-based Oxford Placement Test (Allan, 2004), which yielded a mean score of 75.4% (SD: 5.1) indicating that according to the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001a) they were proficient L2 users (C1 level). Six participants knew one additional language, ten reported knowledge of two additional languages, and five spoke three additional languages. French was reported to be among these languages from 21 participants, Spanish from six, Russian and Dutch from four respectively, and finally Italian, Swedish, Norwegian, Indonesian, Korean and Chinese from 1 participant each. As for their current usage of English, most participants employ it for watching TV or listening to the radio (n=28), reading books (n=29), for work (n=27), talking to partners or family members (n=8), or communicating with friends (n=21).

Appendix C: *Data analysis*

All models were implemented with the lme4 package (Bates, Maechler, Bolker & Walker, 2014) and performed with the R software package (R Development Core Team, 2013). Models included the factors Language (L1 vs. L2), Presentation Type (with-preparation vs. no-preparation), and Trial Type (switch vs. no-switch) and Block Type (single vs. mixed). We fitted the data with crossed random factors for participants and items. Deviation contrasts were used for all fixed effects (0.5 and -0.5), so that estimates for factors reflected main effects and interactions. Intercept adjustments were included for all random factors. Slope adjustments (for the factors Language and/or Presentation Type, Language and Trial Type) were tested for inclusion through model comparisons of nested models (using AIC as a measure of model quality; e.g., Burnham & Anderson, 2004). Since our data were positively skewed, we used the Box-Cox function of the MASS package in R (Venables & Ripley, 2002) to estimate a transformation that would satisfy the assumption of normality of residuals (Kliegl, Masson & Richter, 2010). The results recommended performing an inverse transformation; all RTs were transformed accordingly prior to any further analysis (Baayen & Milin, 2010).

Acknowledgements

The research presented in this study was supported by a PhD fellowship from the Erasmus Mundus Joint Doctoral Programme (EMJD) of the European Union to MM and an Alexander-von-Humboldt-Professorship to HC. We are grateful to the members of the Potsdam Research Institute for Multilingualism for detailed and helpful comments on the present work.

4 Publication II

Published in: J. W. Schwieter (Ed.) Cognitive Control and Consequences of Multilingualism (pp. 145-170). Ontario, Canada: John Benjamins. ISBN: 9789027243720.

Influence of preparation time on language control: A trilingual digit naming study

Julia Festman[a] and Michela Mosca[a,b]

[a] University of Potsdam, Germany;

[b] International Doctorate in Experimental Approaches to Language And Brain (IDEALAB), University of Potsdam, Germany; University of Groningen, the Netherlands; University of Trento, Italy; Newcastle University, UK and Macquarie University, Australia.

Abstract

In this study, we investigated how preparation time influences speed of naming single digits in three languages under frequent switch conditions. Twenty native speakers of English (mean age 19.9 years; 3 males) with good proficiency in French and German participated in a trilingual digit naming experiment with short (150ms) or long (1000ms) preparation time (between language cue and stimulus interval, CSI) and unpredictable switch sequence. Participants responded on average significantly faster on trials with long CSI (562ms) compared to short CSI (657ms). A 2 (Short vs. Long CSI) × 3 (L1, L2, L3) ANOVA showed a significant interaction. The preparation effect was largest for L1 (137 ms) and decreased with decreasing proficiency (L2 84ms; L3 65ms). The paper describes preparation time in relation to naming speed, response accuracy, language proficiency and switching costs.

Introduction

Naming digits in different languages seems like a rather easy task, in particular when limiting the digits to be named to the most common and most frequent ones, namely 1–9. However, as already observed in previous studies on digit naming while frequently switching response language, the ease with which we name digits can be influenced by speakers' internal and/or external factors (for a review on participant and task-related factors see Bobb & Wodniecka, 2013). One important internal factor is language proficiency. Many naming studies involving languages with different levels of proficiency have shown that speakers are faster to name pictures or digits in their stronger language (e.g., Meuter & Allport, 1999). However, this proficiency-related facilitation seems to be restricted to repetition trials (when no language switch is required). When external task-related factors such as cued language switch come into play, stronger languages are usually affected to a greater extent in terms of switching costs (e.g., Philipp, Gade, & Koch, 2007). It has also been consistently reported that the more time is given to prepare for the task the better are naming performances (e.g., Guo, Liu, Chen, & Li, 2013); preparation time is thus yet another critical task-related factor. In the present study we aimed at investigating how speakers' internal and external factors determine accuracy and speed of naming digits and how they interplay with each other.

Taken together, we wanted to explore whether the facilitating modulations due to higher language proficiency (internal factor) and longer preparation time (external factor) are additive or whether one might overrule the other. We asked whether reducing preparation time requires the use of higher levels of cognitive control and whether this is qualitatively different between the languages. Therefore, we tested trilinguals with different levels of proficiency in each language while performing a digit naming task which involved switching between three languages and two preparation time conditions. This study design allows for investigating first of all how internal and external factors are related in a digit naming study. Secondly, a study design involving production in three different languages within the same blocks allows for a closer with-in subject look at language status (native vs. non-native language), which can be investigated pair-wise (native–non-native; non-native–non-native), what is only possible in trilingual processing.

Language proficiency modulates performance

Ecke (2015) stated that there is an extensive body of research demonstrating that proficiency in a language correlates with the time a speaker needs to retrieve words in a language. The pattern which emerged is that the higher the level of language proficiency, the faster the access, retrieval and production of the word. Of course, other factors modulate speed and accuracy of retrieval, such as length of the word and frequency of the word, but these generally cause no influential variation (for an example for trilingual picture naming see e.g., Festman, 2008), if items are controlled for these factors. The impact of language proficiency on speed and accuracy is almost always found when investigating pure language conditions (e.g., Philipp et al., 2007). These are suggested to reveal the language production and language control abilities for each language separately, indicating how quickly the target word can be accessed and produced and how well the processing system can be focused on the target language while avoiding cross-language interference (Festman & Schwieter, 2015).

Crucially, in order to determine language proficiency, it does not suffice to reveal speakers' age of acquisition and length of exposure to a certain language. Language proficiency is dynamic, which means that it may change over a lifetime, influenced by factors such as migration, work, family circumstances, etc. Thus, if language proficiency is a crucial factor under investigation in empirical studies, the current level of language proficiency should be at least indicated (subjective self-ratings of language proficiency) or, even better, measured (objective RTs and accuracy rates). It is not sufficient to merely determine age of acquisition and length of exposure to the respective languages. Language attrition is probably the best example for chronological order of acquisition (see Schmid & Jarvis, 2014) not necessarily being the same as order of language proficiency.

Trial conditions: Repeat or switch?

Since on repetition trials, the response language is identical with the one on the previous trial, this trial type is easier to master than a switch trial. On switch trials the response language differs from the target language on the previous trial. Consequently, the target language has to be changed, what involves – according to the Inhibitory Control (IC) Model (Green, 1998) and adapted to trilingualism (Festman, 2009) – inhibition of the previous target language, activation of the new target language and continuous inhibition of the third language, which was not directly involved in these two consecutive trials. However, other proposals suggest that the language switching cost should not be taken as evidence for inhibition (e.g., Bobb & Wodniecka, 2013; Baus, Branzi, & Costa, 2015).

Trial-type comparisons of RTs and error rates reliably show a "switching cost": switch trials have longer RTs and higher error rates than repetition trials also referred to as "non-switch", "no switch" or "stay trials"; e.g., Meiran, 1996; Monsell, 2003; Kiesel, Steinhauer, Wendt, Falkenstein, Jost, Philipp, & Koch, 2010 for review). The switching cost is calculated as the difference in mean RTs between switch and repetition trials. One way of interpreting switching costs is to take them as a reflection of transient adjustments between task configurations from trial to trial (Rogers & Monsell, 1995). In other words, such cognitive control processes are involved in updating the task-relevant representations (i.e., "task sets", following Monsell, 2003). In this scenario, switching costs are considered to indicate the difficulty of switching between tasks and the control processes involved in the selection of the appropriate task (Festman & Schwieter, 2015). In contrast to that, the task-set inertia theory suggests that switching costs represent negative priming from the previous task, namely persisting activation of the preceding task-set (e.g., Allport, Styles, & Hsieh, 1994).

Language proficiency modulates switching costs

Switching between languages can be considered a type of task-switching. If one of the languages is mastered at a higher level of proficiency than the other (e.g., in the case of unbalanced bilinguals), then this would be like switching between an easier and a more difficult task (Festman & Schwieter, 2015). Task switching experiments involving tasks of unequal difficulty have reported larger switching costs for the easier task compared to the more difficult one (e.g., Allport et al., 1994). The paradoxical asymmetry of the switching costs has been explained with the idea that while performing the weaker task the stronger one has to be more inhibited than when the stronger is performed and the weaker one is suppressed. Consequently, reactivating the stronger task takes more time (or is "*more costly*") than the reactivation of the weaker task; i.e. *asymmetrical switching costs*. The same logic was extended to language switching and described in the IC-Model. A growing number of studies with participants mastering two or three languages to different degrees have observed asymmetrical switching costs with larger costs for the stronger language (e.g., Meuter & Allport, 1999; Costa & Santesteban, 2004; Philipp et al., 2007; Linck, Schwieter & Sunderman, 2012; for a review see Bobb & Wodniecka, 2013).

Mixed-language condition: Two trial types in three languages

Why is switching between three languages so difficult and why is it not enough to do the study with two languages only? Both studies on task switching as well as on language switching provide evidence for mixed-language conditions being more difficult than single-language conditions (Weissberger, Wierenga, Bondi, & Gollan, 2012). As regards language switching studies in particular, in a single/pure condition, the participant can focus on the current target language, since language switching is not required throughout the condition. One could speculate that this is done through *sustained inhibition of the non-target language(s) throughout the entire condition* (Baus et al., 2015). In a mixed condition, however, the target language changes frequently. Sustained inhibition of non-target languages is not an effective strategy to realize frequent language switches between three languages. Thus, frequent switching requires more flexible execution of language control (Festman & Schwieter, 2015), and we therefore suggest transient inhibition as the underlying mechanism (Baus et al., 2015*)*. Regarding language switching studies, there is ample evidence for parallel language activation (e.g., Calabria, Hernández, Branzi, & Costa, 2012), but the degree of involvement of the other language is different between a single and a mixed condition. According to Rogers and Monsell (1995) the additional costs found in mixed-language conditions may be caused by high working memory demands. Extending Roger and Monsell's original reasoning to trilingual experimental settings, this would mean specifically increased working memory demands due to having to perform in three languages and switch between them.

Noteworthy, in studies investigating bilingual processing, speakers' performance in one language is compared with that in the other language (usually a native vs. a non-native language). From this comparison conclusions are drawn on how language control mechanisms might work in bilingual settings. In the present study, three languages are included in the same blocks and are therefore potentially relevant for performance in every target language throughout the entire experiment. This is a crucial difference in the study design to all other studies, in which the language pairs were manipulated blockwise.

In more detail, the inclusion of a third language to the task creates a more nuanced context, in which performance can also be compared with a focus on language status: between two non-native languages as well as between the native and one non-native language. This allows us to see whether a distinct effect is due to language status or rather to the external task-related factors that are actually under scrutiny (i.e., preparation time). Moreover, intuitively, a trilingual language condition in a mixed context is more demanding than a bilingual mixed condition, since participants have to keep two languages inhibited to some degree instead of only one (as in a bilingual setting). This kind of trilingual task setting would allow investigating if the trilingual language control mechanism is different from the bilingual one or just "delayed" (see also de Bot & Jaensch, 2015 for a review); however, this specific comparison is not part of the current study.

Task preparation and preparation time

"In task switching, the term task preparation is used to refer to processes that improve performance when participants know which task is required prior to onset of the target stimulus" (Kiesel et al., 2010; p. 853). Preparation tends to facilitate responses, as was shown in studies on task-switching with non-linguistic stimuli (Meiran, 1996; Rogers & Monsell, 1995, and Kiesel et al. 2010 for a review), as well as for bi- (Fink & Goldrick, 2015) or trilingual (Guo et al., 2013) picture naming tasks.

In task-switching studies, instructional cues presented prior to the target stimulus indicate which task should be performed. Due to the temporal sequencing of cue and target (rather than their simultaneous presentation) cue-related preparatory processes can be clearly distinguished from stimulus-based processes. It has been proposed that part of the control processes in task switching can be moved to the interval before the target stimulus is presented, allowing for a reduction of switching costs when this preparatory interval is long enough.

Meiran (2000) claims that after the presentation of a cue, a *stimulus-set bias* or configuration process begins. The switching can be anticipated and the switching cost may be diminished but cannot be eliminated entirely (see Monsell, 2003 for a review). Importantly, task switching has been found not to be reducible to cue switching alone (Altmann, 2007), since task inhibition (a process related to task switching) is independent of processes related to cue switches. Meiran (1996) and many others found that with increasing length of the CSI (=cue-stimulus interval between the cue and the target stimulus) (a) performance generally improves and (b) in many cases results in decreased switching costs (see Monsell 2003 for review). This being said about task switching, we were wondering whether in a trilingual language switching paradigm in long CSI advanced reconfiguration is fully accomplished prior to stimulus presentation and further processing is thus facilitated, whereas in short CSI, advanced reconfiguration is not fully accomplished before presentation of the stimulus and thus might be on-going when the stimulus is perceived.

In the few cued language switching studies involving preparation time manipulation, participants were given no, short or long preparation time (CSI). Costa and Santesteban (2004) used 0ms (Exp. 2), or 500ms and 800ms (Exp. 5) CSI with highly proficient Spanish-Catalan bilinguals in a picture naming task. According to the assumption that, compared to shorter CSI, longer CSI allows for a more fulfilled reconfiguration of the upcoming task, they found that bilinguals were faster to name pictures on all trials when they had more time to prepare for the stimulus following the cue (at 500ms and at 800ms). Interestingly, authors observed that also switching costs decreased with longer CSI implying that preparation time was more beneficial for switch than for repetition trials. Philipp, et al. (2007) used 100ms and 1000ms CSI in a trilingual digit naming experiment (German, English, French). RTs were faster on all trials with long CSI, and even more pronounced for repetition trials compared to switch trials. Fink and Goldrick (2015) used 500ms and 1250ms CSI in a bilingual digit naming task (digits 1–9) and also found the facilitating effect of longer preparation time for all trials. Despite the fact that all trials benefited from longer CSI compared to shorter CSI, they found that preparation time advantage was greater for L1 compared to L2 trials and for switch rather than for repetition trials (Discussion of Exp. 2). In a trilingual digit naming task with Uighur-Chinese-English trilinguals, Guo et al. (2013) revealed that all three languages were similarly affected by the CSI manipulation (100ms vs. 350ms). They argue that "at the long CSI, the target language task schema can be well established, and then help the following stage of lexical selection (…). At the short CSI, schemas of the target language may not be sufficiently accessed, and accordingly it cannot effectively regulate subsequent lexical selection. As a result, stronger inhibition of the non-target languages is required to guarantee successful lexical selection in the target language (…)" (Guo et al., 2013; p. 281). Finally, Verhoef, Roelofs, and Chwilla (2009) suggested that long preparation times (1250ms) in a cued bilingual language-switching task allowed participants to prepare their L2 responses to such degree that L1 production was no longer facilitated.

All these studies provide strong evidence for the reliability and replicability of the preparation time effect in language switching, although with some variation, probably due to language proficiency differences, task demands and timing, language pairs or triplets and their typological relatedness, etc. As suggested by Monsell (2003), enough preparation time might allow for full reconfiguration of the next relevant task-set. This "advanced reconfiguration" (to use Roger & Monsell's terms) includes processes such as "goal shifting" (Rubinstein, Meyer, & Evans, 2001), "stimulus set biasing" (Meiran, 2000), and "retrieval of task rules" (Mayr & Kliegl, 2003). In a similar way, Verhoef et al. (2009) noticed that a longer interval between cue and stimulus (1250ms) allowed for a better "readiness" for the next trial. This readiness was achieved by speakers' strategic preparation of internal intentions and goals.

In behavioural terms, local (trial-related) and global (block-related) control are reflected in two different kinds of costs: *trial-related costs* (i.e., "switching costs") reflect the difference between two different trials (such as repetition vs. switch trial), whereas *block-related costs* – in this study – reflect the difference between two blocks of unequal difficulty (such as blocks with longer CSI vs. blocks with short CSI). Following this logic, by manipulating CSI, we should be able to detect both types of costs. Note that we did not include the investigation of block-related costs referring to a comparison between pure/single and mixed language conditions.

Present Study

Predictions

To investigate how language proficiency (as our internal factor) is reflected in the standard measures of speed and accuracy, we selected participants with distinct language proficiency differences between L1, L2 and L3. We predict that participants' accuracy scores and RTs for digit naming in the three languages reflect their respective language proficiency and language control abilities with highest accuracy scores and shortest RTs in L1 and lowest accuracy scores and longest RTs in L3.

Due to the additional effort that switching between languages imposes on the processing system, we expect that – because of instantiating reconfiguration to the new target language – participants commit more errors and are generally slower on switch compared to repetition trials. Moreover, according to the switching cost paradox, we expect switching costs to increase with greater language proficiency, i.e. largest switching costs for the L1 and smallest amount of switching costs for the L3, reflecting the application of different levels of language control.

Concerning the external task-related factor, we manipulated preparation time. By defining two extreme CSIs (150ms vs. 1000ms) we expect a strong modulation of both accuracy scores and naming latencies. In short CSI, when the processing system has little time to prepare for the upcoming stimulus, we expect more errors compared to long CSI trials. For every modulation considered separately, we expect an increase of naming latencies and error rates for weaker languages, switch trials and shorter CSI compared to stronger languages, repetition trials and long CSI, respectively.

Furthermore, we predict that in the long CSI trial-related costs are reduced due to accomplished task reconfiguration in terms of target language activation, consequently better “readiness” for the upcoming stimulus, whereas in the short CSI larger trial-related costs indicate incomplete task reconfiguration that does not allow for such performance facilitation.

Finally, we expect to measure block-related costs, defined as the difference between repetition trials in Short vs. Long CSI.

Methods

Participants

Twenty native speakers of English (mean age 19.9 years; 3 male) with good proficiency in French and German participated in this study. They were recruited at the University of Exeter, Great Britain, via advertisement on the university blackboard and by email, since they were part of the Modern Languages Study Programme of the university. They were paid for participation.

Participants were screened for language proficiency before being invited to participate in the study. To avoid language-specific biases, we enrolled eleven participants with French as L2 and German as L3, and nine which had German as L2 and French as L3. Sliding contrasts on the self-ratings mean scores revealed that speakers belonging to the L2 French/ L3 German subgroup considered themselves being more fluent in English than in French ($t= 18.6$, $p< .0001$) and in French more than in German ($t= 6.9$, $p< .0001$). For the L2 German/L3 French subgroup, the analysis showed that English was perceived as the language in which they were more fluent compared to German ($t= 10.6$, $p< .0001$), but also that they were more fluent in German than in French ($t= 4.1$, $p< .01$). Since we did not find any evident difference of language proficiency between the two subgroups, we could confidently merge them into a single experimental group.

Participants filled in a short questionnaire including demographic information, study programme, the level of French and German attained, information on time spent immersed in French- or German-speaking milieu, self-ratings of spoken language proficiency for French and German and information on other language knowledge apart from the ones under investigation (see Appendix A).

Procedure

The study took place in a quiet laboratory room at the University of Exeter. Participants sat in front of a computer screen at a distance of 70cm and in front of a microphone which recorded the verbal responses. A light gray circle with 1.7cm diameter was presented in the center of the screen on white background and served as a fixation circle.

Based on the observation that cue switches cause substantial costs, Logan and Bundesen (2003) claimed that the repetition of a task-indicating cue presents a processing advantage of task-repetition trials, causing cue-encoding benefits (i.e., priming) and thus do not measure cognitive control. Consequently, some studies in the task-switching literature have used a 2:1 cue-task mapping (two cues are used to instruct the participant to perform on the same task). That way, task switches could be de-confounded from cue switches (see a recent review by Jost, De Baene, Koch & Brass, 2013). In an attempt to adapt this reasoning to the trilingual language switching paradigm in this study and to rule out cue-based contributions to specific language-switching costs, we used the 2:1 cue-language mapping, resulting in two cues for each language, thus overall six language cues, each appearing equally often. In task-switching studies, rather arbitrary cues were used, such as different shapes that appear in different colours. No a priori association exists between such cues and the task. In this study we used highly transparent and compatible cues which clearly associate pictorial cue with language. Two types of language cues were used, namely national flags of United Kingdom, France and Germany and pictures of famous buildings (Tower Bridge in London, Eiffel tower in Paris, Brandenburger Tor in Berlin) (see Figure 1). Language cues (10.5 x 5.8cm) were displayed in the center of the screen (in the background of the gray circle).

Fig. 1. Two different types of language cues: national flags and most famous monuments.

Digit stimuli (1cm of height, Arabic numbers 1–9) were presented in black in the center of the gray fixation circle; simultaneously, an auditory trigger was sent to the recording software for RTs measurement of the verbal response to the stimulus. The experiment was run using DMASTR software developed at Monash University and at the University of Arizona by K.I.Forster and J.C.Forster.

Participants were asked to look at the gray circle, behind which a language cue was presented and in the middle of which a digit appeared which should be named as fast and as accurately as possible. For example, if the language cue indicated English and the digit presented was "2", the participant was asked to say out loud "TWO". If the language cue indicated German, and the digit presented was "2", the target response was "ZWEI". Digits were randomized and repetitions of the same digit were allowed only after two trials presenting different digits.

All language cues, trial types (switch trial, repetition trial) and both CSIs were presented equally often. The two different preparation intervals were presented block-wise and used to explore the effect of CSI (i.e., cue-stimulus interval) on digit naming in trilinguals. In the Long CSI condition, we presented the fixation circle for 500ms, then the language cue appeared and after 1000ms the digit was added to the display. In the Short CSI condition, the fixation circle was displayed for 1350ms, then the language cue appeared and after 150ms the digit was added to the display. Overall, participants first practiced the task on four blocks with 24 trials per block before the experiment proper started which included 16 blocks with 50 trials each with long and short CSI fully randomized. Half of the participants started with the short CSI condition, the other half with the long CSI condition.

Data coding and analysis

Two types of languages cues, as described earlier, were used, but collapsed, following Philipp and Koch's (2009) claim that "inhibition is not restricted to a specific cue (…)". The experimental design included accuracy rates and naming latencies as dependent variables. Recordings of participants' verbal responses were compared to the target response list for each participant. A response was deemed incorrect in case of hesitation, errors in language selection, wrong answers, miss-articulation, microphone miss-triggering (e.g., coughing) and represented 12.3% of the data.

Not only is language proficiency a major determinant in speed and accuracy of performance, but for a study, the choice of language with regard to typological distance is crucial when it comes to evaluate the difficulty of trial-by-trial responses in terms of degree of conflict between the target and its two translation equivalents that compete for selection. In particular, the more phonologically similar the target words are in the stimulus materials, the more difficult the inhibition of the two lexical competitors from the languages irrelevant on the current trial, as it has been shown in studies on cognates, homographs and homophones (see e.g., Dijkstra, Grainger & van Heuven, 1999). This assumption is reflected in the error analysis of our digit naming study as well. The digit “6” was named with only 83% accuracy, followed by “5”, “7” and “9” with 85% respectively. The remaining digits were responded to more and more accurately in the following order: “4”, “3”, “2”, “8”, and “1” as the most accurately named digit (see Table 2).

Table 2. Material used (digits) and its translation in English, French and German [IPA in brackets].

DIGIT	ENGLISH	FRENCH	GERMAN
1	one [wʌn]	un [œ̲̃]	eins [aɪ̯ns]
2	two [tuː]	deux [dø̲]	zwei [t͡sva ɪ̯]
3	three [θriː]	trois [tʁ̲wa]	drei [dʁaɪ̯]
4	four [fɔː]	quatre [katr]	vier [fiər]
5	five [fʌɪv]	cinq [sɛ̲̃ŋ̲k]	fünf [fʏnf]
6	six [sɪks]	six [sis]	sechs [zɛks]
7	seven [ˈsɛv(ə)n]	sept [sɛ̲t]	sieben [ˈziːbən]
8	eight [eɪt]	huit [ɥ̲it]	acht [ɑxt]
9	nine [nʌɪn]	neuf [nœ̲f]	neun [nɔɪ̯n]

All participants and all items had an accuracy rate higher than 70% (i.e. participants’ accuracy ranged from 74% to 96% and items’ accuracy ranged from 83% to 92%). Based on this inspection, none of the participants or items were excluded from the analysis.

A second check-up concerned incorrect responses (see above) and outliers. While for the analysis of accuracy both correct and incorrect responses were used, only correct responses were included when analysing naming latencies, which were defined as the interval between the display of the target stimulus and the speech onset. Correct responses were then screened for outliers. Observations more than 2 standard deviations away from the distribution mean were treated as outliers and excluded from further analysis (4.9% of the data). Visual inspection revealed that data were positively skewed, thus violating the normality assumption underlying the general linear model (Baayen & Milin, 2010). To decide on the most appropriate transformation for the data, we used the boxcox function of the MASS package in R (Venables & Ripley, 2002), which suggested to log-transform the data.

All the statistical analyses were implemented in the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) in the R software package (R Development Core Team, 2015). We fit linear mixed-effect models with crossed random effects for participants and items to the naming latencies and generalized linear mixed-effect models with binomial function to the accuracy data (Cnaan, Laird, &

Slasor, 1997; Guo & Zhao, 2000). To select the best-fit models, we compared nested models with increased complexity of fixed and/or random-effect structures, excluding combinations not supported by likelihood ratio tests (e.g., Baayen, Davidson, & Bates, 2008). The best-fit models included Language (L1 vs. L2 vs. L3), Condition (Repetition vs. Switch) and CSI (Short vs. Long) as significant predictors. All fixed effects were coded with sliding contrasts (0.5 vs. –0.5 for two levels factors of Condition and Cue, and 2/3 vs. –1/3 vs. –1/3 and -2/3 vs. 1/3 vs. 1/3 for the three level factor of Language), which allowed to estimate coefficients for both main effects and interactions. Models criticism (Baayen, 2008) was also performed by removing data points with absolute standardized residuals above 2.5 SD (8.2% of the data). In line with the assumption underlying mixed-effect modelling, the trimmed models approximate normality.

Results

To investigate the hypotheses outlined in the introduction, we analysed the effect of speakers' internal (language proficiency) and external (preparation time) factors on a digit naming task involving switching in three languages. Table 1 shows the mean accuracy scores and RTs for all the experimental manipulations.

Table 1. Mean RTs (standard deviations in brackets) and accuracy rates (in percent) for correct trials of L1 vs. L2 vs. L3 in both Repetition vs. Switch condition (upper part) and as a function of Short CSI vs. Long CSI (lower part).

	L1	L2	L3	Mean
Repetition	526ms (165)	585ms (160)	614ms (162)	575ms (167)
Accuracy	94%	88%	86%	89%
Switch	548ms (181)	603ms (168)	623ms (168)	590ms (176)
Accuracy	90%	86%	83%	86%
Switching costs	22ms	18ms	9ms	
Mean	537ms (174)	594ms (165)	618ms (165)	583ms (171)
Accuracy	92%	87%	84%	88%

Short CSI				
Repetition	586ms (167)	614ms (162)	634ms (167)	610ms (166)
Accuracy	93%	87%	85%	88%
Switch	611ms (177)	637ms (170)	654ms (170)	633ms (173)
Accuracy	90%	84%	82%	85%
Mean	598ms (167)	625ms (162)	643ms (167)	621ms (170)
Accuracy	91%	86%	83%	87%
Switching costs	25ms	23ms	20ms	
Long CSI				
Repetition	467ms (142)	558ms (153)	595ms (156)	538ms (159)
Accuracy	96%	90%	86%	91%
Switch	488ms (165)	570ms (160)	593ms (162)	548ms (169)
Accuracy	91%	87%	83%	87%
Mean	478ms (154)	564ms (157)	594ms (159)	543ms (164)
Accuracy	93%	88%	84%	89%
Switching costs	21ms	12ms	-2ms	

Accuracy data

The analysis of accuracy rate revealed that participants responded more accurately on L1 compared to L2 trials (92% vs. 87%, z= 6.64, p< .0001), and that responses in L2 were more accurate than in L3 (87% vs. 84%, z= 3.33, p< .0001). As expected, for every language accuracy rates were lower on switch trials compared to repetition trials (86% vs. 89%, z= 6.39, p< .0001) and on trials with short CSI compared to trials with long CSI (87% vs. 89%, z= 3.73, p< .0001). There was a significant interaction of Language by Condition, indicating that compared to repetition trials accuracy rates in switch trials decreased more for L1 than for L2 (from 94% to 90% for L1 vs. from 88% to 86% for L2, z= 2.67, p< .05). No such difference was found between L2 and L3 (from 88% to 86% for L2 vs. from 86% to 83% for L3, z= 1.35, p= 0.17). Post-hoc analysis showed that relative to the repetition condition, accuracy rates in the switch condition decreased more for L1 compared to L3 (from 94% to 90% for L1 vs. from 86% to 83% for L3, z= 2.35, p< .01).

To investigate the effect of preparation time on error rates, we analyzed accuracy scores by means of a theory-driven analysis, i.e. on Short and Long CSI separately.

In the Short CSI, participants made more errors in L3 compared to L2 (z= 2.07, p< .01) and in L2 than in L1 (z= 2.4, p< .05). As for the collapsed data (i.e., Short and Long CSIs together), also in the Short CSI dataset, Switch trials yielded more errors than Repetition trials (z= 5.45, p< .0001). Furthermore, the marginally significant Language by Condition interaction suggests that error rates tended to increase more in L1 than in both L2 and L3 on Switch compared to Repetition trials (z= 1.83, p< .1 and z= 1.77, p< .1, respectively). However, the increase of error rates on Switch compared to Repetition trials did not differ between L2 and L3 (z= .89, p= .37).

With concern to the Long CSI condition, the analysis showed again a main effect of Language, with L1 being more accurate than L2 (z= 7.00, p< .0001) and L2 having a higher accuracy score than L3 (z= 3.43, p< .05). Repetition trials elicited less errors than Switch trials (z= 5.43, p< .0001) and there was a significant interaction between Language and Condition indicating that the increase of error rates in the Switch compared to the Repetition condition was greater for L1 than for both L2 and L3 (z= 2.16, p< .05 and z= 2.55, p< .05, respectively). Similarly to the Short CSI condition, L2 and L3 error rates did not differ between L2 and L3 on Switch compared to Repetition trials (z= .38, p= .70).

Naming latencies

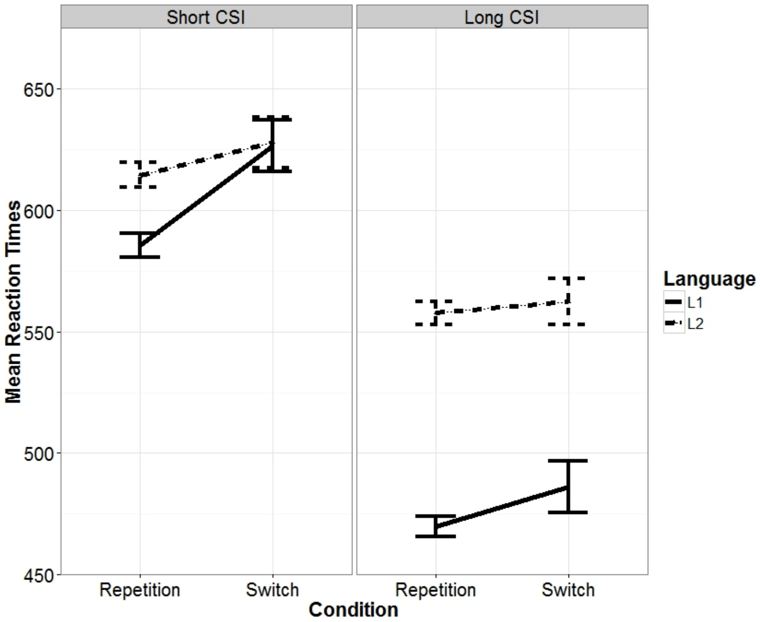
The analysis of correct responses indicated that, as predicted, participants were faster in naming digits in L1 compared to L2 (537ms vs. 594ms, t= 7.44, p< .0001), and in L2 relative to L3 (594ms vs. 618ms, t= 2.49, p< .05). Naming latencies were slower on Switch trials than on Repetition trials (590ms vs. 575ms, t= 4.12, p< .001) and in Short CSI compared to Long CSI (621ms vs. 543ms, t= 13.57, p< .0001).

Language significantly interacted with CSI, suggesting that longer time to prepare for the next trial was more beneficial for L1 than for L2 (119ms vs. 62ms benefit t= 10.40, p< .0001), and for L2 relative to L3 (62ms vs. 49ms benefit, t= 3.23, p< .01). We measured a non-significant interaction between Language and Condition (t= .72, p= .48 for the L1-L2 comparison; t= .83, p= .41 for the L2-L3 comparison and t= 1.56, p= .13 for the L1-L3 comparison) inferring that the overall amount of switching costs did not differ between the three languages. However, Condition interacted significantly with CSI (t= 3.07, p< .01), revealing that the processing system benefited more on Switch trials than on Repetition trials from longer CSI (85ms vs. 72ms preparation benefit, respectively). The beneficial effect of longer CSI on Switch trials was the same for all three languages as suggested by the non-significant three-way interaction of Language, Condition and CSI (t= .27, p= .79, for the L1-L2 comparison; t= 1.24, p= .23, for the L2-L3 comparison and t= .99, p= .33, for the L1-L3 comparison).

In the general analysis, for each language switching costs were calculated for the two non-target languages together, for instance L1 switching costs reflect the mean cost to switch from L2 and L3. However, to measure more detailed language switching costs in relation to the proficiency and language status difference between two languages (and reflecting the native – non-native language distinction in a trilingual setting), we split the data in three language pairs (i.e. L1/L2, L1/L3 and L2/L3). For the language pair L1/L2, mean naming latencies of Repetition and Switch trials in both Short and Long CSI are plotted in Figure 2.

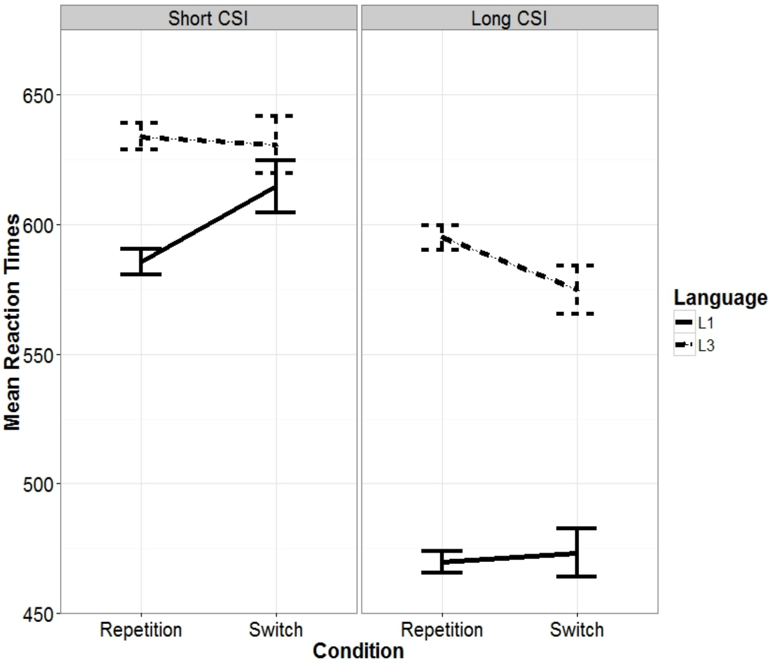
Naming analysis for this language pair revealed that overall L1 was faster than L2 (532ms vs. 587ms, t= 5.78, p< .0001), Switch trials were slower than Repetition trials (577ms vs. 555ms, t= 4.42, p< .001) and compared to Short CSI, Long CSI lead to faster reaction times (respectively 605ms vs. 514ms, t= 15.81, p< .0001). The significant Language by Condition interaction (t= 2.28, p< .05) indicates that L1 (59ms) switching costs were larger than L2 (35ms) ones. Moreover, there was a significant interaction between Condition and CSI (t= 2.74, p< .05), showing that on switch trials the processing system benefited more from longer preparation interval than on Repetition trials (102ms vs. 87ms benefit respectively), which was independent from Language (t= .45, p= .65). Following the significant interaction, we ran post-hoc analyses on Short and Long CSI data separately. In the Short CSI condition, participants were faster to respond in L1 than in L2 (594ms vs. 617ms, t= 4.86, p< .0001) as well as on the Repetition relative to Switch trials (599ms vs. 627ms, t= 3.53, p< .01). The reliable interaction between Language and Condition showed that language switching costs were larger for L1 compared to L2 (41ms vs. 13ms, t= 2.41, p< .05), revealing thus *asymmetrical switching costs* for the L1/L2 language pair in the Short CSI. In the Long CSI condition, however, while the analysis showed that L1 responses were faster than L2 ones (473ms vs. 559ms, t= 8.16, p< .0001), we did not find any difference between Switch and Repetition trials (t= 1.05, p= .30). This null effect of Condition was the same for both L1 and L2 (t= .84, p= .41).

Fig. 2. Naming latencies (in ms) for L1 and L2 as a function of Condition (Repetition vs. Switch) in Short and in Long CSI condition.



With regard to the language pair L1/L3, mean reaction times of L1 and L3 for Switch and Repetition condition in Short as well as Long CSI are illustrated in Figure 3.

Fig. 3. Naming latencies (in ms) for L1 and L3 as a function of Condition (Repetition vs. Switch) in Short and in Long CSI condition.

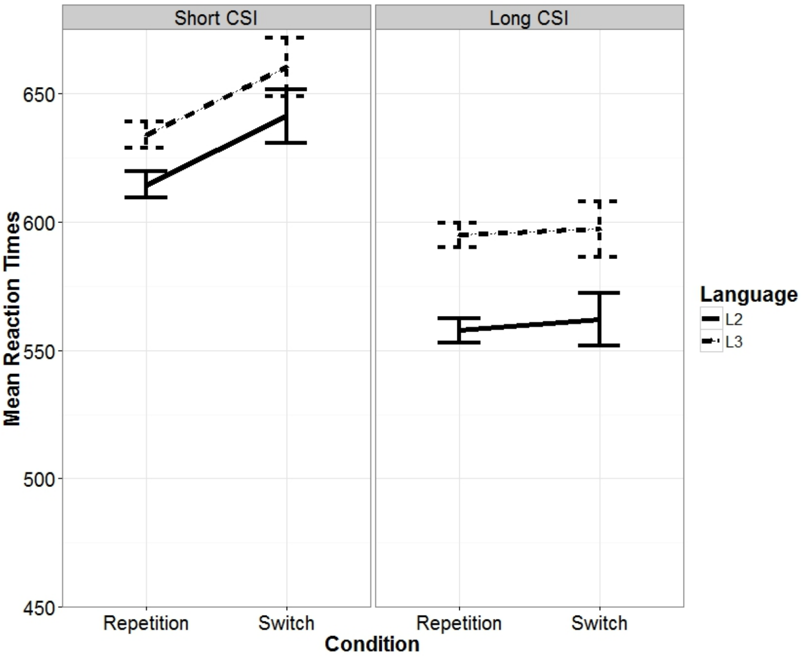


The analysis of naming responses of this language pair showed that L3 trials were slower than L1 trials (612ms vs. 530ms, t= 8.18, p< .0001), Switch condition had higher naming latencies than

Repetition condition (570ms vs. 567ms, t= 2.53, p< .05) and Long CSI trials were faster than Short CSI trials (527ms vs. 611ms, t= 17.38, p< .0001). As suggested by the significant Language by Condition interaction, L1 switching costs were larger than for L3 (17ms vs. –11ms, t= 2.59, p< .05). There was a significant interaction between Language and CSI (t= 10.77, p< .0001) showing that compared to Short CSI, Long CSI speeded L1 naming responses more than L3 ones (121ms vs. 42ms benefit respectively). The three-way interaction of Language, Condition and Cue was not significant (t= 1.27, p= .21). Post hoc analysis on Short CSI trial revealed that responses in L3 were slower than in L1 (633ms vs. 591ms, t= 3.27, p< .01) and that Repetition trials had faster naming latencies than Switch trials (608ms vs. 622ms, t= 2.62, p< .05). The main effects were qualified by a significant Language by Condition interaction (t= 2.37, p< .05), which indicated that switching costs for the stronger L1 (29ms) were larger than those of the weaker L3 (-3ms). This result shows that in the language pair L1/L3, Short CSI yielded *asymmetrical switching costs*. With respect to the Long CSI, we measured faster naming latencies for L1 than for L3 (470ms vs. 591ms, t= 10.29, p< .0001). However, similarly to the results of the L1/L2 language pair in the Long CSI condition, we did not find any effect of Condition (t= 1.04, p= .31) for any of the two languages (t= .74, p= .46) in the Long CSI trials of the L1/L3 language pair.

Finally, consider Figure 4 for the mean reaction times relative to the language pair L2/L3 in Repetition and Switch condition for Short and Long CSI.

Fig. 4. Naming latencies for L2 and L3 as a function of Condition (Repetition vs. Switch) in Short and in Long CSI condition.



The analyses of the naming latencies showed that participants were faster in naming digits in L2 than in L3 (589ms vs. 617ms, t= 2.72, p< .05) and in Repetition trials relative to Switch trials (599ms vs. 615ms, t= 2.33, p< .05). With regard to preparation time, Long CSI yielded faster naming latencies than Short CSI (576ms vs. 629ms, t= 8.11, p< .0001). However, the three main effects were not qualified by any significant interaction (t= 0.14, p= .68 for the Language by Condition interaction; t= 1.88, p= .07 for the Language by Cue interaction; t= 1.01, p= .32 for the Cue by Condition interaction and t= .72, p= .48 for the Language, Condition and Cue three-way interaction). To compare this language pair with the L1/L2 and L1/L3 language pairs, we performed an exploratory analysis on Short and Long CSI separately. Short CSI trials showed no significant difference between L2 and L3 naming responses (620ms vs. 639ms t= 1.27, p= .21), but it revealed that Switch trials were reliably slower than Repetition trials (624ms vs. 650ms, t= 2.75, p< .05). This main effect of Condition was not modulated by Language (t= .42, p= .67), inferring that switching costs for L2 (27ms) and L3 (26ms) were *symmetrical* in Short CSI. In the Long CSI condition we found that L2 naming responses were faster than L3 ones (558ms vs. 595ms, t= 3.56, p< .01). However, similarly to L1/L2 and L1/L3 language pairs in Long CSI trials, we did not find any effect of Condition (Repetition vs. Switch) in the L2/L3 language pair (t= 1.03, p= .31). The null effect was again independent from Language (t= .66, p= .51).

Taken together, the results from this trilingual digit naming task reflect the typical performance pattern related to language proficiency (speakers' internal factor) in terms of accuracy and response latencies (that is declining in accuracy and speed from the language mastered at the highest level of proficiency to the lowest). As regards preparation time (speakers' external factor), these findings suggest that the unbalanced trilinguals tested in the present study responded faster when they had more time to prepare for the upcoming trial. responded faster when they had more time to prepare for the upcoming trial. The processing system benefited especially on switch trials from longer preparation time, leading to the dissipation of language switching costs in the Long CSI time, leading to the dissipation of language switching costs in the Long CSI condition for all languages. In contrast, in the Short CSI condition language switching costs seem to be proficiency-related, i.e., asymmetrical for L1/L2 and L1/L3 (greater switching costs for the stronger language whenever the native language was involved) and symmetrical for L2/L3 (for the two non-native languages).

Discussion

In this study, we aimed at investigating the effect of language proficiency (speakers' internal factor) and the effect of preparation time (speakers' external factor) on a digit naming task involving switching between three languages. Specifically, we wanted to determine the within-subject effect of preparation time on digit naming processing in L1, in L2 and in L3, to highlight the processes involved in switching between the three languages under mixed-language conditions and to define distinct processing differences when manipulating preparation time – if there are any.

Switching between three languages poses large task demands on the processing system. As predicted, accuracy and speed measures are a direct reflection of the respective language proficiency: the higher the language proficiency the higher are accuracy scores and the faster the responses.

The role of language proficiency for language switching costs

Additionally, language switching costs seem to be related to language proficiency, as we found larger switching costs for the stronger of the two languages in the language pairs which involved the native language: L1/L2 and L1/L3 (*asymmetrical switching costs*), but similar amount of switching costs for the L2/L3 language pair (*symmetrical switching costs*). While switching costs involving the native language (L1-L2 and L1-L3 language pairs) replicates asymmetrical switching costs found in naming studies involving languages with different strength (e.g., Costa, Santesteban & Ivanova, 2006), the symmetrical switching costs observed for the two non-native languages (L2-L3 language pair) is unexpected. As reported in the self-rated proficiency assessment and as clearly shown by the naming latencies in the on-line task, all trilinguals tested in this study showed higher levels of proficiency in L2 than in L3. This being said, the symmetrical switching costs for this language pair are not in line with previous studies (see Philipp et al., 2007; Linck et al., 2012). However, as inferred by the accuracy and naming latencies analyses for digit naming in this study, the amount of switching costs for L1-L2 vs. L1-L3 was comparable (28ms vs. 32ms respectively), suggesting that the relative difference between the native and each non-native language was comparable. Accuracy scores confirmed this hypothesis, revealing greater error rates for L1 in switch condition compared to repetition condition but no difference between L2 and L3. To sum up, the language proficiency difference between the two non-native languages in the present study might have been not large enough to elicit asymmetrical switching costs in the L2-L3 language pair on a rather simple task such as digit naming. Even more so, since in the L2/L3 condition, inhibition of the strongest language L1 was necessary on every trial. Therefore, the proficiency difference between L2 and L3 might have been (partially) masked by the continuous inhibition of L1.

The role of CSI

The presence of two CSIs (150ms vs. 1000ms) allowed us to observe how preparation time can modulate language control in trilinguals. With regard to the switch condition, results reveal that the switching cost patterns observed in the Short CSI were not detectable in the Long CSI (for similar results, see Mosca & Clahsen, 2016). However, it should be noted that, in the latter case, in any language pair there was a trend for switch trials to be more costly than repetition trials. This tendency could reflect what has been referred to as "residual switching costs". Residual costs are believed to reflect a type of task-set reconfiguration that takes place only when the target stimulus is presented. Therefore, according to this view, despite the long preparation time between cue and stimulus, speakers are not completely done with the reconfiguration until the stimulus is presented (see also Monsell et al., 2000). With concern to both Switch and Repetition conditions, we observed that in shorter CSI all languages were affected by a general slowdown. Interestingly, the magnitude of this global slowdown was directly related to language proficiency, the stronger the language the greater the reductions of naming speed.

What do these results tell us about language control? The two CSI conditions clearly show that when preparation time was shorter, trilinguals in our study experienced two kinds of costs: Costs related to language switching (which we call "trial-related costs") and costs related to the language as a whole (which we term "block-related costs"), i.e., also repetition trials were affected by the costs. These two types of costs might reflect the two types of control, namely one active on the trial level and another on the entire block level. Consequently, the effects of the trial-related control are reflected in local or switching costs, whereas the effect of block–related control is detectable in global costs. More generally, where switching costs are related to the activation threshold of the language, i.e., the higher the threshold the larger the switching costs, "global costs" are supposed to reflect the difference between a higher demanding (i.e. Short CSI) and a lower demanding task (i.e. Long CSI). Specifically, we suggest that in Short CSI, because of restricted preparation time (150ms), advanced reconfiguration of the upcoming trial could not be fully accomplished, whereas in Long CSI the interval was sufficient (1000ms) to prepare for the next trial, leading to performance facilitation. Interestingly both kinds of control (or costs) are strongly related to language proficiency. We discuss this in the next sections.

CSI × Language

The present study clearly shows that preparation time affected L1, L2 and L3 naming latencies differently. Specifically, we found that preparation time benefit was greater for the native language L1 compared to the two non-native languages, L2 and L3. Additionally, we found that longer CSI had a greater benefit on L2 than on L3. Evidence of greater preparation time facilitation for the native compared to the non-native language is not a new result (e.g., Fink & Goldrick, 2015). However, the presence of an additional non-native language in our study unambiguously led to the evidence that preparation time benefit increases along with language proficiency. This finding allows disentangling which factor – language status (native vs. non-native) and language proficiency (more proficient vs. less proficient) – is actually driving preparation time benefit.

Finally, it should be considered that, to our knowledge, there are two other studies in which the CSI has been manipulated in a trilingual digit naming task. In one of these studies, CSI affected all three languages equally (Guo et al., 2013) and in the other one, longer preparation time yielded a traditional L1 advantage over L2 and L3 but no difference between the two non-native languages (Philipp et al., 2007). However, it is worth noting that in the first study, authors measured n-2 repetition costs and not switching costs and that the CSI manipulation was a between-subject factor and not a within-subject factor like in the present study. In the second study, trilinguals' switching costs were measured in three separate experimental contexts, where only two languages at a time were used, i.e. L1-L2, L1-L3 and L2-L3. Therefore, it is conceivable to suppose that the trilinguals' performances might have been influenced by the fact that, as mentioned earlier, they were tested in bilingual contexts (i.e. two languages at a time) and not in a single trilingual setting as in our study. Therefore, because of profound methodological differences it is hard to directly compare the results of the present study with those obtained in the previous ones.

To conclude, we suggest that in conditions where more time is available to prepare for the next trial, the speed of the reconfiguration mechanism is related to language proficiency levels, i.e. the stronger the language the faster and more effective (in terms of facilitation) the reconfiguration process.

CSI × Condition

In the present study, we observed a beneficial effect of longer CSI over shorter CSI that was reflected in both faster reaction times and lower error rates. This finding replicates the traditional preparation time effect consistently reported in task as well as language switching studies (e.g., Fink & Goldrick, 2015; Meiran, 2000). We also found that longer preparation time was more advantageous for switch trials than for repetition trials, as revealed by the absence of switching costs in the Long CSI. This effect was the same for all three languages.

Why is preparation time particularly beneficial for switch trials? As already discussed, switching costs might reflect the extra time needed to update a new task-set in the switch trial (Roger & Monsell, 1995) or an inertial interference coming from the previous task-set (Allport et al. 1994). In our study, we had two CSI types (i.e. 150ms vs. 1000ms) and since the Response to Stimulus interval (RSI) was kept constant (at 1500ms), we also had two types of Response to Cue interval (RCI – i.e. 1350ms for Short CSI vs. 500ms for long CSI). We observed switching costs only in the Short CSI condition where RCI was longer (1350ms) and we did not find evidence of language switching costs in long

CSI condition, where RCI was shorter (500ms). This suggests that the switching costs measured in the Short CSI of the present study might reflect costs of preparation for the upcoming trial, rather than interference from what was relevant in the trial before. This does not rule out interference from the previous trial as possible source of switching costs, as observed in several studies investigating backwards effects of inhibition (e.g., Mayr & Keele, 2000; Philipp et al., 2007). Indeed according to a more hybrid approach, both reconfiguration of the new task and passive decay of the previous task might influence switching cost magnitude (Monsell, 2003).

Conclusion

The goal of the present study was to investigate which are the mechanisms that underlie trilingual language control. The manipulation of preparation time (Short CSI vs. Long CSI) in a language switching task involving three languages of unequal proficiency (L1 vs. L2 vs. L3) allowed us to detect two different types of costs, i.e., trial-related and block-related costs. With reference to trial-related costs, they were found only in the Short CSI condition and were influenced by language proficiency. Specifically, when splitting the data into language pairings for a closer look at language status, we found asymmetrical switching costs for the language pairings L1-L2 and L1-L3, with larger switching costs for the native compared to a non-native language. Concerning block-related costs, we found that preparation time benefit increased along with language proficiency, with greater preparation time for L1 than for L2, and for L2 relative to L3. Overall, results indicate that both types of costs are modulated by language proficiency, suggesting that the multilinguals' language control system is a highly adaptive system.

Finally, we believe that only the presence of a second non-native language in the study allowed us to verify that the observed effects were due to the factors under study and not to the fact that we were comparing a native vs. a non-native language (as in most of the studies involving bilinguals). Indeed, as indicated by the increasing benefit of preparation time on the three languages, with the weakest benefit for L3 and the strongest for L1, we were able to reveal that the system adapted step-wise to each language according to the proficiency. This observation would not have been possible when comparing only two languages. Therefore, we believe that trilingualism provides data and conclusions that cannot be gathered from bilingual studies.

Appendix A: *Language history and self-rating scores of the participants*

Participants were all native speakers of English (L1) learners of French and German. For 11 participants French was the stronger foreign language (L2) and German the weaker foreign language (L3) and for nine participants German was the stronger non-native language (L2) and French the weaker one (L3). In the self-rating task participants had to self-assess their spoken level of French and German based on a Likert scale ranging from 1 ("no knowledge of the language") to 20 ("like a native speaker"). Mean score of spoken language self-ratings and of amount of time spent in a French and/or German speaking milieu are illustrated in Table A1. Participants were all university students and were specializing in the following languages: 5 (French and German), 3 (French), 3 (German), 3 (French and Spanish DP), 2 (Italian and French DP), 1 (English and French), 1 (German and Spanish), 1 (German, Italian and French) and 1 in Physics. Beyond the three languages under investigation, 13 participants reported knowing one additional language: for 7 of them was Spanish and for 6 participants was Italian. Only in one case, a participant reported knowing three more languages (Punjabi, Hindi and Urdu).

In the questionnaire, participants were also requested to report the level of language proficiency attained for both French and German. The proficiency levels ranged from low (corresponding to the A2 Level of the Common European Framework of Reference) to high level (corresponding to the level achieved in a successful 4th university year in Modern Language Programme). With regard to French, from the L2 French/ L3 German subgroup 2 participants reported having reached a high proficiency level (4th year university level), 4 considered themselves being at an intermediate stage (2nd year university level), 3 at an intermediate/low level (1st year university level) and only 2 reported having a low level of language attainment (A2 level). From L2 German/L3 French subgroup, 2 participants reported having attained a high acquisition stage (4th university year level), 1 an intermediate level (2nd university year level), 1 an intermediate/low level (1st university year level), and 5 a basic level (A2 level). With concern to German, in the L2 French/ L3 German subgroup almost everyone reported having achieved a low proficiency level (A2) except for 2 participants with a low/intermediate (1st university year) and intermediate (2nd university year) level each. Conversely, in the L2 German/L3 French subgroup most of the subjects stated having reached a high level of proficiency (4th university year level), except for 2 participants with low/intermediate (1st university year) and 1 with low (A2) proficiency level.

Table A1. Mean score (standard deviations in brackets) of spoken language self-ratings (1–20) and of amount of time (in months) spent in a French and/or German setting for native speakers of English (L1) learners of French and German (i.e. L2 French/L3 German L3 on the left; L2 German /L3 French on the right).

		L2 French/L3 German			*L2 German/L3 French*	
	L1	**L2**	**L3**	**L1**	**L2**	**L3**
Self-rating	20	13.1 (1.2)	10.6 (0.8)	20	13.2 (2.6)	9 (3.9)
Time of immersion (months)	-	16.6 (47.7)	0.5 (1.8)	-	0.2 (0.4)	3.2 (3.9)

Acknowledgements

This study was conducted at and supported by the University of Exeter (Department of Cognitive Science), when Julia Festman had the great opportunity to stay shortly in Exeter and to work with Stephen Monsell and David Green at University College London. Michela Mosca is part of the International Doctorate for Experimental Approaches to Languages And Brain (IDEALAB), University of Potsdam, Germany; University of Groningen, The Netherlands; University of Trento, Italy; Newcastle University, United Kingdom and Macquarie University, Australia.

5 Publication III

Under review with: the *Quarterly Journal of Experimental Psychology*

Trilinguals' language switching: A strategic and flexible account

Michela Mosca[a,b]

[a] International Doctorate in Experimental Approaches to Language And Brain (IDEALAB), University of Potsdam, Germany; University of Groningen, the Netherlands; University of Trento, Italy; Newcastle University, UK and Macquarie University, Australia.

[b] Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam, Germany

Abstract

The goal of the present study was to determine how trilinguals select the language they intend to use in a language switching context. Two accounts are examined: 1) a *language specific* account, according to which language selection considers the activation level of the words of the intended language only (i.e. language coactivation without language competition) and 2) a *language non-specific* account, where activated words from both the intended and non-intended languages compete for selection (i.e. language coactivation with language competition).

Results showed that, in both groups, all three languages competed for selection, and that selection was achieved by inhibiting the currently non-relevant languages. Moreover, extending findings from previous research, the study reveals that, in both Exp. 1 and Exp. 2, the amount of inhibition was not only influenced by language proficiency, but also by the typological similarity between languages.

Overall, the study shows that language switching performance can be accounted for by a strategic and flexible inhibitory account. In particular, the controlling system is "strategic" in the sense that it aims at preventing potential conflicting situations, such as typological closeness between languages, and it is "flexible" in that it adjusts languages' activation levels, depending on the conflict to be solved.

Keywords:

Language switching; language selection; trilingualism; inhibition; language typology

Introduction

While walking through the city a person speaking language A asks you for street directions. Although language B is your native-language, you know you have to reply in language A and not in language B in order to be understood by your interlocutor. This is a common situation for multilinguals, who according to context, have to adapt their response language in order to successfully communicate. The ability to use the intended language while avoiding interference from the non-intended language(s) is known as “language control”. In the past, researchers have supported the idea of a language-switching device (Macnamara, 1967; Penfield & Roberts, 1959), to activate the intended language and deactivate the unintended one (Grainger, 1993; Grainger & Beauvillain, 1987). However, it is now acknowledged that speakers are not equipped with such a mental device specific to language and that language switching might rely on cognitive control mechanisms also involved in non-linguistic tasks (e.g., Abutalebi & Green, 2007; Festman, 2012; Green & Abutalebi, 2013). Additionally, it is now generally agreed that while multilinguals use one specific language, also non-selected languages might be coactivated (e.g., Colomé, 2001; de Bot, 2004; Dijkstra & van Hell, 2003; Grosjean, 2001).

Despite the common belief that relevant and non-relevant languages are coactivated, researchers seem to disagree on how language selection is ultimately achieved. On the one hand, some authors support the idea that language selection is *language specific*, where despite languages’ joint activation only words of the target language are considered for selection (e.g., Costa & Caramazza, 1999; Costa, Miozzo & Caramazza). This hypothesis assumes that after languages’ coactivation a lexicon-external device determines which of the two lexicons has to be consulted. This monitoring device is sensitive to specific properties of the lexical nodes (e.g., grammatical class, language membership), so that in the multilingual case it restricts lexical search to lexical nodes of the relevant language only, while ignoring activated words in the non-intended languages, i.e. language coactivation without language competition (Costa & Caramazza, 1999). On the other hand, it has been suggested that selection is *language non-specific*, where after languages’ joint activation words from both target and non-target language start competing for selection, i.e. language coactivation with language competition (e.g., de Bot, 1992; Poulisse & Bongaert, 1994). This last view has found major support in studies investigating multilinguals’ language control, in that both languages are believed to compete for selection, which can be achieved by a) activating the words of the relevant language more (e.g., Finkbeiner, Gollan & Caramazza, 2006; La Heij, 2005) or b) activating the words of the intended language while suppressing the words of the non-intended language (e.g., Green, 1998). While the first language non-specific approach (a) is more parsimonious than the second one (b), since no extra mechanisms are required to achieve language selection, it fails to make clear predictions on language selection behaviour and, therefore, to account for the available data.

For example, La Heij's (2005) suggests that selection of a word is achieved when its activation level exceeds the activation levels of the competing words from both the relevant and the non-relevant language. According to this, words with higher baseline activation level are retrieved faster relative to words with lower baseline activation level and the baseline activation level of a word depends on its frequency of use, i.e. more frequently used words (e.g., words in the L1, the native language) have higher activation level compared to less frequently used words (e.g., words in the L2, the first non-native language). The author further suggests that bilingual language selection behaviour may reflect processes involved in the incorporation of the language cue (a piece of information that specifies the language to be used) in the preverbal message (a conceptual structure containing word's non-verbal information, such as pragmatic and affective characteristics). However, this explanation leaves unclear how the incorporation is specifically fulfilled, whether the mechanism is the same for L1 and L2 words and, importantly, if and how the incorporation mechanism is affected by the different baseline activation levels of L1 and L2 words.

Consequently, as an alternative to the language specific view, the present paper will consider only the language suppression hypothesis (b) as a comprehensive language non-specific account. According to the language suppression hypothesis elaborated in the Inhibitory Control (IC) Model proposed by Green (1986, 1993, 1998), the amount of suppression applied on non-target words depends on their level of activation, namely words from a stronger language are more strongly inhibited than words from a weaker language. The amount of inhibition applied will then affect the speed with which words will be reactivated, i.e. more strongly inhibited words will need more time to overcome suppression compared to less strongly inhibited words.

The Language Switching Paradigm

In order to investigate language control in multilinguals, a growing body of studies has relied on language switching paradigms, where languages need to be constantly selected and deselected. The language switching paradigm includes two types of trials, *repetition* (same language as in the preceding trial) and *switch* trials (different language compared to the trial before). Responses in switch trials are usually slower and more error prone than in repetition trials; the difference between switch and repetition trial is known as "switching costs" (e.g., Roger & Monsell, 1995). Larger switching costs for a stronger compared to a weaker language, i.e. *asymmetrical switching costs*, have been taken as evidence in favour of the IC-Model (e.g., Costa & Santesteban, 2004; Kroll, Bobb; Misra & Guo, 2008; Linck, Schwieter & Sunderman 2012; Meuter & Allport, 1999; Schwieter, 2013; Wang, Xue, Chen, Xue & Dong, 2007). However, many studies testing highly proficient bilinguals with a stronger L1 and a weaker L2 have shown that, despite the proficiency difference, switching costs for the two languages might be the same, i.e. *symmetrical switching costs*. Based on these results, it has been suggested that when the difference in language proficiency is relatively small, the strength of suppression applied on the two languages is comparable, leading to a similar amount of reactivation costs (e.g., Meuter & Allport, 1999; Fink & Goldrick, 2015).

Importantly, some authors have suggested that both the IC-Model and a language specific mechanism are possible in multilinguals, with a lexical selection process requiring inhibitory control in speakers with low L2 proficiency and a lexical access process relying on a language specific mechanism in speakers with high L2 proficiency (Costa & Santesteban, 2004; Costa, Santesteban & Ivanova, 2006; Schwieter & Sunderman, 2008). Within this "hybrid" approach, researchers have considered asymmetrical switching costs as an evidence for a dominance-related inhibition (the stronger the language the greater the inhibition applied on it) and symmetrical switching costs, as an index of a language specific mechanism. Precisely, supporters of this account specified that once the lexical selection process has shifted from an inhibitory mechanism to a mechanism where no inhibition is needed since only words from the relevant language are considered (see Schwieter & Sunderman, 2008, for an estimate of the critical point in L2 lexical robustness in which the shift supposedly takes place), the latter will be used in other cases of language switching, for example when switching between the stronger L1 and the weaker second non-native language, i.e. the L3 (Costa & Santesteban, 2004).

Crucially, in a cued language-switching paradigm Verhoef, Roelofs, & Chwilla (2009) found that preparation time affected switching costs pattern in a group of low-proficient Dutch-English bilinguals. In particular, the authors found asymmetrical switching costs (larger for the L1 Dutch than for the L2 English) when the interval between the cue and the stimulus (CSI) was shorter (500ms) and symmetrical switching costs when the CSI was longer (1250ms). This result indicates that both asymmetrical and symmetrical switching costs can be found within the same group of participants, suggesting that the amount of competition between words' belonging to different languages can be modulated by several more factors, such as preparation time, than only by language proficiency (for a review on these factors see Bobb & Wodniecka, 2013).

Languages' Strength of Activation

As proposed by Grosjean (1998, 2001) speakers are sensitive to context and simply knowing that a certain language might become relevant will enhance the activation level of that language. Similarly, Grainger & Dijkstra (1992) suggested that the task goal can change the languages' activation levels, with contextual more appropriate languages being more activated than contextual less relevant languages. This phenomenon known as “proactive task adaptation” is supposed to take place before the onset of the task (de Groot, 1998). Therefore, the fact that trilinguals’ language control has been often investigated in language switching tasks involving only two languages at a time (i.e. L1-L2, L1-L3 and L2-L3 separately), instead of all three languages simultaneously, might have affected languages’ activation and competition level (e.g., Abutalebi et al. 2013; Costa & Santesteban, 2004; Costa et al., 2006; Philipp, Gade & Koch, 2007). For example, it might be the case that while in a trilingual setting (e.g., L1, L2 and L3) all three languages are activated to a certain extent because they are all relevant to the task, in a dual-language setting (e.g., L1-L3) activation is predominantly sent to the two languages relevant to the task but not so to the non-relevant language (e.g., L2). This might be connected to the fact that the trilingual and the dual-language setting differ in complexity. While in the former, multiple responses are possible (e.g., L1, L2 or L3), i.e. multiple responses task, the latter includes only two types of responses (e.g., either L1 or L3), i.e. two choice task. Depending on the properties of the task, the same subject can perform differently, that is rely on different problem-solving strategies (Paquette & Kida, 1988). Therefore, the activation level of the two languages in the dual-language setting and that of the same two languages in the trilingual setting might vary as a function of task complexity.

A way to measure to what extent languages’ strength of activation is affected by context, is to compare, for each language, performance between trials in a “single-language” block (where only one language is required for the task) with trials of the same kind, i.e. repetition trials, but in a “mixed-language” block (where more than one language are involved in the task). Trials in a single-language task are usually responded faster and more accurately than repetition trials in a mixed-language block, the difference between single- and mixed-language trials is known as “mixing costs”. Mixing costs are believed to reflect the sustained effort of maintaining two or more tasks/languages active in the mixed- relative to the single-language block (Braver, Reynolds & Donaldson, 2003). However, there is growing evidence that mixing costs are not so sustained or “fixed”, in that they can be affected by preparation time (Mosca & Clahsen, 2016) and tend to vary according to language dominance, i.e. greater mixing costs for the stronger than for the weaker language (e.g., Prior & Gollan, 2011). Additionally, some studies have shown that mixing costs can be absent in the weaker but not in the stronger language (e.g., Christoffels, Firk & Schiller, 2007), or that they can turn into mixing benefits for the weaker but not so for the stronger language (e.g., Hernandez, Dapretto, Mazziotta & Bookheimer, 2001). Overall, results suggest that mixing costs might not only reflect the sustained cost of maintaining multiple tasks/languages active, because in that case the amount of mixing costs should not vary across languages or experimental manipulations (see Ma, Li & Guo, 2016).

With this regard, it has been proposed that the different amount of mixing costs between the stronger and the weaker language might reflect an unconscious strategy to facilitate performance in the weaker of the two languages, by inhibiting the stronger language more, i.e. lowering its baseline activation level, and/or by activating the weaker language more, i.e. enhancing its baseline activation level (Mosca & Clahsen, 2016; for a review on mixing costs see Festman & Schwieter, 2015). Therefore, within a language non-specific account, the symmetrical switching costs usually found between the stronger L1 and the weaker L2 of highly proficient bilinguals could reflect the fact that languages' relative strength of activation is modulated by the task (L2 activation level is lowered and/or L1 activation level is enhanced) rather than by the relatively small proficiency difference between the two languages.

Consequently, it is difficult to determine how language context has influenced languages' strength of activation in previous trilinguals' language control studies testing two languages at a time and how this was then related to languages' switching costs. Moreover, studies investigating trilinguals' language control in a trilingual context have focused on the overall amount of switching costs for each language, but not on the switching costs' asymmetry between two languages (e.g., Linck et al. 2012; Schwieter & Sunderman, 2011, Schwieter, 2013 but see Festman & Mosca, 2016), leaving unclear which mechanism underpins language control in this population.

Hence, by using a trilingual context, the first goal of the present study is to determine whether language selection in trilinguals with high L2 proficiency is language specific or whether it relies on an inhibitory language non-specific mechanism. To control for languages' strength of activation, for each language performance are compared between trials of the single-language block and repetition trials of the mixed-language block.

Languages' Typological Distance

As Costa et al. (2006) noted, the way languages are controlled could be affected by languages' similarity, with more similar languages being more likely to interfere with each other compared to less similar languages. To investigate this aspect, authors compared performance of highly proficient bilinguals switching between two typologically distant languages (L1 Spanish - L2 Basque) with that of highly proficient bilinguals switching between two closely related languages (L1 Spanish - L2 Catalan). Results revealed symmetrical switching costs for both groups of bilinguals, leading authors to conclude that language similarity does not significantly affect the way bilinguals control languages.

While this might be the case for bilingual language processing, there is growing evidence suggesting that language similarity might play a role in L3 processing. As indicated by a wealth of research on trilinguals, L3 acquisition can be subject to cross-language influences from both the native L1 and the non-native L2 (e.g., Singleton, 1987; Festman, 2008). Which of the two languages will mostly influence L3 processing will depend on several factors (Williams & Hammarberg, 1998), with the most influential one being “language typology” (e.g., Kellerman, 1983; Ringbom, 1987). According to the “language typological distance” hypothesis, the language typologically closer to the weaker L3 will become the main source of cross-language interference during L3 production. If L1 and L2 are similarly close to the weaker L3, then the L2 is likely to dominate cross-language transfer into the L3 (e.g., Cenoz, 2001, 2003; Hammarberg, 2001). Specifically, as supported by the “foreign language effect” theory, during L3 production it is easier to inhibit the native L1 rather than the non-native L2, because the L1 is perceived as qualitatively different from the non-native L2 and L3. Moreover, when acquiring the second non-native language (L3), speakers might unconsciously rely on the same strategies previously used to acquire the first non-native language (L2). This might lead the L2 to become the main source of transfer during L3 processing (Hammarberg, 2001).

The effects predicted by the language typological distance hypothesis could be accounted for by de Bot’s (2004) multilingual production model, according to which the multilingual lexicon is composed of language-specific subsets and depending on linguistic closeness, the language subsets might overlap with each other to different degrees. When a language is activated, elements shared with other languages might also be activated. By extension, the greater the degree of overlap between two subsets, the stronger the co-activation (and thus interference) of the two languages.

Furthermore, de Bot’s (2004) multilingual production model proposed that since the L1 is used more, its activation/deactivation networks might be stronger relative to those of the non-native languages. Thus, suppressing the native language during L3 production should be easier than suppressing the non-native L2. It follows that the amount of inhibition applied on a language where suppression is more easily exerted is greater compared to a language where inhibition is more difficult to be applied on. Therefore, if as predicted by the language typological distance hypothesis, L3 processing is hampered by a typological closer L2, this should influence the way these two languages are controlled in a language switching context.

With this respect, the second aim of the present study is to investigate the effects of language typological closeness between the two non-native languages on trilingual language control. According to the Typological Primacy Model (TPM; Rothman, 2015) typological proximity between languages can refer to four different levels, i.e. lexicon, phonological/phonotactic cues, functional morphology and syntactic structure (ordered according to their relative impact). In the present study languages’ typological closeness refers to the distance at the lexicon level between two languages. Specifically, language switching costs were measured in a trilingual picture naming task involving two groups of participants, speakers of German, English and Italian. Measurements of typological distance at the lexicon level among the three languages were taken from the study conducted by

Schepens, Dijkstra and Grootjen (2012). Based on their calculations of relative cognate frequency between six European language pairs (English, German, French, Spanish, Italian and Dutch), the authors were able to measure the degree of lexical closeness between several language pairs. As expected, on the lexicon level, German and English were shown to be more closely related than Italian and English or Italian and German (Schepens et al., 2012).

Overview of the Experiments

The present study is structured as follows. In Exp. 1 native speakers of Italian (L1), highly proficient speakers of German (L2), learners of English (L3) were tested. For this first group of trilinguals, the L3 English is typologically closer to the L2 German, but typologically more distant to the L1 Italian. Exp. 2 included native speakers of German (L1), highly proficient speakers of English (L2), learners of Italian (L3). For this second group of trilinguals, the L3 Italian is typologically more distant to both L2 English and L1 German. Overall, the aim of the study is twofold: 1) to investigate whether the lexical selection process in trilinguals with high L2 proficiency relies on a language specific or an inhibitory language non-specific mechanism. Specifically, if trilinguals rely on a language specific mechanism, for both groups symmetrical switching costs and similar error rates in each language pairing (L1-L2, L1-L3 and L2-L3) are expected. In contrast, if they rely on an inhibitory mechanism, asymmetrical switching patterns and different error rates are predicted in each language pairing (L1-L2, L1-L3, and L2-L3). In particular, switching costs and error rates should be greater for the stronger L1 and smaller for the weaker L3, while language switching asymmetry between two languages is expected to increase along with the relative proficiency difference of two languages, i.e. larger asymmetry for L1-L3 than for L1-L2 or L2-L3. 2) The second aim of the study is to assess whether language control in trilinguals is influenced by language lexical similarity. More precisely, if typological closeness to the L3 does not affect trilinguals' language control, then irrespective of whether language control in trilinguals is supported by a language specific or non-specific mechanism, the same patterns of results are expected for Exp. 1 and Exp. 2. In contrast, if trilinguals' language control is sensitive to language typological closeness to the L3, then different language switching patterns are expected in Exp. 1 compared to Exp. 2.

Experiment 1

Participants

Thirty-two native speakers of Italian (eight men, mean age: 28.3 years) with good proficiency of German and English participated in this experiment. They were recruited by means of flyers, social networks or via the university blackboard and were paid for their participation. All the participants gave informed consent and filled out a background questionnaire to gain information on their linguistic and demographic history. They were all born in Italy and were living in the Berlin area (Germany) at the time of testing (mean length of immersion in the German environment: 2.8 years). Some of them were enrolled in a Master study program in the host country, while others had finished their university studies in Italy and had moved to Germany to work. All the participants had grown up in an Italian monolingual environment and started learning a non-native language at school or later. They reported acquiring German from the mean age 16.6 years (8-36 years, sd: 6.2) and English from the mean age of 10.4 years (5-34 years, sd: 4.9). On a daily basis, participants reported speaking Italian most of the time (63.8% - with their partner, children, friends and family), then German (30% -with their partner, extended family, friends, university peers and working colleagues) and little English (6.2% – with friends, university peers and working colleagues). The oral and written exposure (watching TV, listening to the radio and reading of books, newspapers, etc.) did also differ between the three languages: German was the language they were exposed to most (50.5%) followed by Italian (31.4%) and then by English (18.1%).

To screen for languages' proficiency, participants were asked to self-rate their abilities to speak, comprehend, write and read in the three languages (Italian, German and English). The assessment was based on a ten-point scale, in which 0 = "no knowledge of the language" and 10 = "proficient like a native speaker". Results of the self-rating task are shown in Table 1.

Table 1. Experiment 1: Mean scores (standard deviations in brackets) of the self-rating task for speaking, comprehension, writing and reading skills in Italian, German and English.

	Speaking	*Comprehension*	*Writing*	*Reading*	*Mean*
Italian	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)
German	7.8 (0.9)	8.3 (1.2)	7.7 (1.3)	8.4 (1.3)	8.1 (1.1)
English	6.1 (1.5)	6.8 (1.9)	6.3 (1.5)	7.7 (1.3)	6.7 (1.7)

In general, participants considered themselves as proficient as native speakers in Italian, highly proficient in German and less proficient in English. Sliding contrasts on the mean scores of the three languages revealed that the difference between the three languages was significant (i.e. Italian vs. German, β= 1.89, SE= .18, t= 10.28, p< .0001; German vs. English, β= 1.34, SE= .27, t= 4.92, p< .0001). Therefore, despite the fact that English was the foreign language they started learning first, they considered German their stronger non-native language.

To further investigate language proficiency, participants were administered a verbal fluency task, which is considered a reliable indicator of lexical robustness, i.e. the size and the strength of the lexicon (Gollan, Montoya & Werner, 2002; Schwieter & Sunderman, 2011). In the verbal fluency task, participants are asked to name as many words as possible starting with a given letter (i.e. phonemic subtest) or belonging to a specific semantic category (i.e. semantic subtest) within 60s. Verbal fluency performance is believed to reflect the ease with which lexical representations are retrieved from the mental lexicon and is considered, therefore, a good indicator of language proficiency (Schwieter & Sunderman, 2008). For each language, two letters and two semantic categories were selected; for details see Appendix A. The total score for each language was calculated by adding the correct responses from the phonemic and the semantic subtests. Proper names, mispronounced terms and words repeated more than once were deemed as incorrect responses. As expected, participants performed better in the native language Italian compared to the non-native languages. In this latter case, they named more words in German than in English. Sliding contrasts revealed that the performance difference among the three languages was significant (i.e. Italian vs. German, $\beta= 20.9$, $SE= 2.22$, $t= 9.41$, $p< .0001$; German vs. English, $\beta= 6.09$, $SE= 2.22$, $t= 2.74$, $p< .05$). Hence, despite English having been acquired before German, German seemed to be the stronger non-native language for the participants tested in the present experiment. This proficiency pattern is in line with the self-rating scores, according to which participants considered themselves being more proficient in German than in English. Based on these results, Italian was labelled as the L1, German as the L2 and English as the L3 of the participants.

Materials and Method

Materials

Eighteen pictures were selected from the "Colorized Snodgrass and Vanderwart pictures" set (Rossion & Pourtois, 2004) representing simple concrete objects. Items were presented at the centre of a 15-inch computer screen set to 1280x800 pixel resolution. They had a size of 197x281 pixel and were seen from a distance of approximately 80cm. DMDX (Forster and Forster, 2003) was used for items presentation and CheckVocal (Protopapas, 2007) for recording & measuring speech-onset latencies. Stimuli were matched for lemma frequency, words length (in letters), cognateness and semantic category based on the values of the International picture naming project (IPNP) database (Bates et al., 2003). One-way ANOVA among the three languages revealed that items did not differ for lemma frequency (Italian: 2.317, German: 2.688 and English: 3.193; $p> .05$) nor for letters length (Italian: 6.1, German= 5.8 and English= 5.3; $p> .05$). No cognate words were used in the experiment and it was made sure that all the test pictures belonged to different semantic categories to avoid uncontrolled cumulative semantic interference (Howard, Nickels, Coltheart, & Cole-Virtue, 2006). Moreover, all the selected pictures were evaluated as conceptual simple (conceptual complexity= 1, see also Bates et al. 2003). The materials were identical to those used by Mosca and Clahsen (2016); the complete list of the items is reported in Appendix B.

Procedure

Participants sat in front of a computer screen in a quiet laboratory room. They were first given oral instructions by the researcher about the task, followed by written instructions on the computer screen. Participants were asked to name presented pictures in either L1 (Italian), L2 (German) or L3 (English) as fast and accurate as possible. To avoid effects of preparation time on words' competition level, participants were given no time to prepare for the upcoming trial (CSI= 0ms). The language to be used was indicated by the colour of the computer screen background (i.e. yellow= L1; blue= L2 and red= L3). A trial consisted of (i) a fixation cross (for 500ms), (ii) a blank screen (for 300ms), (iii) a picture together with the language cue (1.500ms), (iv) a blank screen (for 2.400ms). Each trial had a fixed duration of 4.700ms, independently of subjects' response speed.

The experimental session consisted of a *single-language* followed by a *mixed-language* block. In a single-language block, pictures had to be named in either L1, L2 or L3 separately. A single-language block was composed by a total of 108 trials. For each language, the 18 pictures were randomly presented in two consecutive sessions (total of 36 trials per language). The language of instructions corresponded to the language in which the upcoming task had to be performed (e.g., instructions in English for the upcoming L3 part). There were two lists of the single-language block, one starting with the L2 (blue screen background) and one beginning with the L3 (red screen background). The L1 (yellow screen background) was always seen the last. Lists' order was counterbalanced across participants. The mixed-language block entailed two types of trials, "repetition" and "switch" trials. In a repetition trial, the presented picture had to be named in the same language as in the trial before (e.g., L1-L1), whereas in a switch trial a given picture had to be named in a different language compared to the trial before (e.g., L2-L1). The mixed-language block consisted of 432 trials (324 repetition and 108 switch trials). Moreover, each 1/3 of the trials (144 trials) had to be named in either L1, L2 or L3. The trials were grouped in pseudo-randomized language chunks, where 75% were repetition and 25% were switch trials. For example, a language chunk such as L1-L1-L1-L2 implied that three consecutive pictures had to be named in the L1 and the fourth picture in the L2. For the analysis, only the second half of the chunk was used (x-x-L1-L2), so that it was possible to compare 108 repetition trials with 108 switch trials. This chunk system was used to make sure that every repetition trial was coming from a "clean" repetition trial (note the difference between *L2*-*L1*-L1-L2 and *L1*-*L1*-L1-L2) and to avoid effects of backward inhibition in switch trials (note the difference between L1-*L2*-L1-*L2* and L1-*L1*-L1-*L2*), for a review on backward inhibition see Koch, Gade, Schuch, and Philipp (2010). A specific type of chunk pattern (e.g., L1-L1-L1-L2) did not appear more than two times in a row. The native language (Italian) was used as language of instructions throughout the entire block. The same 18 pictures of the single-language block were used in the mixed-language block. Each picture was seen 24 times throughout the mixed-language experiment, which is once in each position of the language chunk (i.e. first, second, third and fourth position), and in each of the possible language chunks (i.e. L1-L1-L1-L2; L2-L2-L2-L1; L1-L1-L1-L3; L3-L3-L3-L1; etc.). The same picture was not seen within five trials. The mixed-language block was divided into six parts; each part (72 trials) contained the same number of trials for each language (24 trials) and was followed by a short break for the participant. Six different lists of mixed-language block were created, from which each participant was administered only one. Two of the lists started with the L1, two with the L2 and two with the L3. The lists order was counterbalanced across participants. Because of these precautions, trials' order was unpredictable.

Before each experimental part, participants were familiarized with the task by means of a practice session. There were 3 practice sessions for each language part of the single-language block (6 trials each) and one practice session for the mixed-language block in which all the possible chunk structures were trained (total of 24 trials). The pictures used for the practice sessions differed from the ones used in the main experiment. The approximate duration of one experimental session (including instructions and pauses) was 1 hour.

Data cleaning

The dependent variables of the present study were accuracy of responses and naming latencies. Before running the statistical analyses, the data set was cleaned up based on subjects' and items' error rates. Participants and/or items with accuracy scores lower than 70% were not submitted to further analysis. Based on this threshold, two participants (participants' accuracy rate ranged from 50% to 99%) but no item (items' accuracy rate ranged from 75% to 94%) were discarded from subsequent analysis. Both correct and incorrect responses were included in the analysis of the accuracy score, whereas only correct responses were used to analyze naming latencies (defined as the interval between the presentation of the target and the speech onset). A response was classified as incorrect in case of missing response, selection of the wrong language or of the wrong word, and in case of microphone miss-triggering, such as coughing or hesitation (accounting for 9.7% of the data points).

After excluding incorrect responses, naming latencies were screened data for outliers. Implausible fast responses (<100ms) and too slow values (>3500ms) were treated as extreme outliers (see also Baayen & Milin, 2010; Luce, 1986) and trimmed from the dataset (0.12% of the data). Despite outliers trimming, visual inspection of the naming latencies revealed a highly skewed distribution of the data. The most appropriate transformation for the data was chosen by means of the boxcox function of the MASS package in R (Venables & Ripley, 2002), which suggested performing a logarithmic transformation. Moreover, a visual comparison between log and inverse transformed data, confirmed that the log transformation was the most appropriate to approximate normality. For detailed information on data analysis see Appendix C.

Results

To investigate the assumptions outlined in the introduction, mixing and switching costs were measured for the three languages under scrutiny: L1 Italian, L2 German and L3 English. For each language, *mixing costs* were calculated as the difference between trials in the single-language block and repetition trials in the mixed-language block; whereas *switching costs* were measured by comparing performance in repetition vs. switch trials of the mixed-language block. Mean naming latencies and accuracy scores for both single- and mixed-language block are illustrated in Table 2.

Table 2. Experiment 1: Mean naming latencies (standard deviations in brackets) and accuracy rates (in percentages) for correct responses of L1 vs. L2 vs. L3 in single-language block (upper part) vs. mixed-language block (lower part). In the mixed-language block, mixing costs (calculated as the difference between trials in the single-language block and repetition trials in the mixed-language block) and switching costs (calculated as the difference between repetition and switch trials) for L1, L2 and L3 are reported.

		L1	**L2**	**L3**
Single-language block				
Mean	*Naming latencies*	773ms (258)	1000ms (313)	1052ms (380)
	Accuracy rates	97%	87%	83%
Mixed-language block				
Repetition	*Naming latencies*	861ms (278)	1021ms (347)	1097ms (398)
	Accuracy rates	98%	91%	88%
MIXING COSTS		*88ms*	*21ms*	*45ms*
Switch	*Naming latencies*	1093ms (320)	1295ms (418)	1225ms (388)
	Accuracy rates	93%	86%	86%
SWITCHING COSTS		*232ms*	*274ms*	*128ms*
Mean	*Naming latencies*	975ms (321)	1154ms (407)	1160ms (398)
	Accuracy rates	96%	88%	87%

Accuracy rates. Consider first the single-language block. The analysis of the error rates revealed that participants responded more accurately in the native language than in the two non-native languages (β= 1.66, SE= .25, z= 6.47, p< .0001 for the L1 vs. L2 comparison and β= 1.95, SE= .25, z= 7.68, p< .0001 for the L1 vs. L3 comparison), while no difference was found between L2 and L3 (p> .05). Similarly, in the mixed-language block, participants made fewer errors in L1 responses compared to L2 and L3 ones (β= 1.81, SE= .33, z= 5.34, p< .0001 for the L1 vs. L2 comparison and β= 1.99, SE= .35, z= 5.67, p< .0001 for the L1 vs. L3 comparison), while accuracy rates for L2 and L3 were not significantly different (p> .05). With regard to Block Type, trials in the single-language block were responded less accurately than repetition trials in the mixed-language block (β= .58, SE= .17, z= 3.44, p< .01). This effect was independent of Language, as suggested by the non-significant interaction between Language and Block Type (all p> .05).

As expected, in the mixed-language block, switch trials elicited more errors than repetition trials (β= .69, SE= .14, z= 4.81, p< .0001). This effect was modulated by Language, indicating more errors for the L1 than for the L2 and the L3 in switch compared to repetition trials (trend towards significance for L1 vs. L2 β= .64, SE= .37, z= 1.75, p= .08 and significant effect for L1 vs. L3 β= .97, SE= .36, z= 2.62, p< .01). The interaction of Language by Condition was not significant for the L2 vs. L3 comparison (p> .05).[2]

Naming latencies. Analysis of the naming latencies in the single-language block indicated that participants were faster in the L1 than in the L2 or L3 (β= .26, SE= .04, t= 6.28, p< .0001 for L1 vs. L2 and β= .31, SE= .03, t=8.06, p< .0001 for L1 vs. L3). Naming latencies in L2 and L3 did not differ significantly (p> .05). In the mixed-language block, results revealed similar pattern of naming latencies for the three languages as in the single-language block, which is faster naming latencies for L1 than either L2 or L3 (β= .16, SE= .02, t= 7.09, p< .0001 for L1 vs. L2 and β=.16, SE= .002, t= 7.90, p< .0001 for L1 vs. L3) and no difference for L2 vs. L3 responses (p> .05). With regard to Block Type, analysis of the naming latencies showed that participants responded significantly slower in repetition trials of mixed-language block than in the single-language block (β= .04, SE= .01, t= 4.00,

---

[2] Please note that the analysis of the accuracy rates revealed a different pattern of results compared to what reported in Tables 2, where percentages of the accuracy scores seem to suggest that the difference between switch and repetition trials was smaller for the L3 (i.e. 88% vs. 86%) compared to both L1 and L2 (98% vs. 93% and 91% vs. 86% respectively) and that there was no difference between L1 and L2. The difference between the percentages reported in Table 2 and the estimates of the model are due to the fact that in binomial data, such as accuracy rate (i.e. correct vs. incorrect responses), a change in percentages around 50% corresponds to a smaller change in predictors than the same change in percentages close to 0% or 100%. This difference is captured by the logistic regression model that operates on log-odds rather than on raw proportions (Jaeger, 2008). Therefore, with reference to Table 2, the analysis of the accuracy rates of switch and repetition trials showed that the difference between 98% and 93% for the L1 is not the same as the difference between 91% and 86% for the L2, but rather that the latter resembles the difference between 88% and 86% found in the L3 (for an in-depth discussion on categorical data analysis see Agresti, 2002).

p< .001), indicating that naming repetition trials in a mixed-language condition was more costly than in a single-language context, i.e. *mixing costs*. Mixing costs were larger for the L1 than for both L2 (β= .10, SE= .02, t= 4.36, p< .0001) and L3 (β= .08, SE= .01, t= 4.67, p< .0001). Mixing costs for the two non-native languages, L2 and L3, did not differ significantly (p> .05).

In the mixed-language block, switch trials were responded slower than repetition trials (β= .20, SE= .001, t= 12.10, p< .0001), i.e. *switching costs*. This effect was significantly smaller for the weaker L3 compared to the L1 or the L2 (β= .12, SE= .001, t= 3.81, p< .001 for L3 vs. L1 and β= 0.12, SE= .003, t= 3.88, p< .001 for L3 vs. L2), while no significant difference was detected between the L1 and the L2 (p> .05). Moreover, the analysis revealed that, compared to L1 and L3, L2 switching costs were more strongly modulated by the language of the preceding trial, as indicated by the significant three-way interactions of Language, Trial Type and Preceding Language (β= .14, SE= .006, t= 2.15, p< .05 for L1 vs. L2 and β= .15, SE= .003, t= 2.34, p< .05 for L2 vs. L3). The difference between L1 and L3 was not significant (p> .05).

Briefly, in both single and mixed-language block participants responded faster in their native than in a foreign language and within the mixing language block they suffered from smaller switching costs in the weaker L3 compared to the stronger L1 and L2. Moreover, participants were slower in repetition trials of the mixed-language block compared to trials of the single-language block and this effect was larger for the L1. Finally, the language of the preceding trial affected switching costs, particularly in the L2.

Preceding Language and Switching Costs

To disentangle the effect of previous trial language on switching costs, post-hoc tests were run on the data split in three language pairing subsets, namely L1-L2, L1-L3 and L2-L3. Each language pairing included repetitions and switch trials from two languages only (e.g., the subset L1-L2 included L1 and L2 repetition trials as well as L1-L2 and L2-L1 switch trials).

If language selection relies on a dominance-related inhibitory mechanism, switching costs are expected to be larger for the stronger compared to the weaker language in each language pairing, i.e. L1-L2, L1-L3 and L2-L3 asymmetrical switching costs. In contrast, if lexical selection relies on a mechanism where only words from the intended language are considered for selection, then the same amount of switching costs are predicted for the stronger and the weaker language of any language pairing, i.e. L1-L2, L1-L3 and L2-L3 symmetrical switching costs. Mean naming latencies and accuracy rates for each language pairing are illustrated in Table 3.

Table 3. Experiment 1: Mean naming latencies (standard deviations in brackets) and accuracy rates (in percentages) for correct responses of the language pairings L1-L2, L1-L3 and L2-L3 in repetition and switch trials. For each language pairing, switching costs (calculated as the difference between repetition and switch trials) are reported in italics.

L1-L2 language pairing

		L1	**L2**
Repetition	*Naming latencies*	861ms (278)	1021ms (347)
	Accuracy rates	98%	91%
Switch	*Naming latencies*	1070ms (291)	1354 (431)
	Accuracy rates	93%	84%
Mean	*Naming latencies*	929ms (299)	1126ms (406)
	Accuracy rates	96%	89%
SWITCHING COSTS		*209ms*	*333ms*

L1-L3 language pairing

		L1	**L3**
Repetition	*Naming latencies*	861ms (278)	1097ms (398)
	Accuracy rates	98%	88%
Switch	*Naming latencies*	1117ms (346)	1237 (401)
	Accuracy rates	94%	84%
Mean	*Naming latencies*	944ms (325)	1142ms (404)
	Accuracy rates	96%	87%
SWITCHING COSTS		*256ms*	*140ms*

L2-L3 language pairing

		L2	L3
Repetition	*Naming latencies*	1021ms (347)	1097ms (398)
	Accuracy rates	91%	88%
Switch	*Naming latencies*	1238ms (398)	1213ms (374)
	Accuracy rates	87%	87%

Mean	*Naming latencies*	1092ms (378)	1136ms (394)
	Accuracy rates	90%	88%
SWITCHING COSTS		*217ms*	*116ms*

Accuracy rates. The analysis of the accuracy rates showed that in the L1-L2 subset participants made more errors when responding in the L2 than in the L1 ($\beta= 1.33$, $SE= .18$, $z= 7.17$, $p< .0001$) and in switch compared to repetition trials ($\beta= .97$, $SE= .18$, $z= 5.23$, $p< .0001$). The effects were not modulated by language as suggested by the non-significant interaction between Language and Trial Type ($p> .05$). In the L1-L3 subset, responses in the L1 were significantly more accurate than in the L3 ($\beta= 1.79$, $SE= .37$, $z= 4.85$, $p< .0001$) and in repetition compared to switch trials ($\beta= .82$, $SE= .25$, $z= 3.26$, $p< .01$). The interaction between Language and Trial Type failed to reach significance ($p= .09$). Finally, for the L2-L3 subset results indicate that participants responded equally accurate in the L2 and in the L3 ($p> .05$) as well as in switch compared to repetition trials ($p> .05$). The interaction of Language and Trial Type was also not significant ($p> .05$).

Naming latencies. For the L1-L2 language pairing, analyses of the naming latencies revealed that trials in the L1 were named faster than in the L2 ($\beta= .19$, SE .02, $t= 8.12$, $p< .0001$) and that switch trials were responded slower than repetition trials ($\beta= .26$, $SE= .02$, $t= 11.95$, $p<. 0001$). Despite the difference between L1 and L2 raw mean switching costs (respectively 209ms and 333ms) reported in Table 3, the interaction between Language and Trial Type based on log-transformed data did not reach significance ($p= .11$). This indicates that on a log scale the degree of switching costs for L1 and L2 was similar, i.e. *symmetrical switching costs*. Concerning the language pair L1-L3, results showed that responses were faster in the L1 than in the L3 ($\beta= .16$, $SE= .02$, $t= 6.75$, $p< .0001$) and in repetition trials compared to switch trials ($t= 9.28$, $p< .0001$). The significant interaction of Language by Trial Type ($\beta= .19$, $SE= .02$, $t= 3.37$, $p< .01$) indicated that switching costs for the L1 were larger than for the L3, i.e. *asymmetrical switching costs*. Finally, the analysis of the L2-L3 language pairing revealed that mean responses latencies of two languages did not differ ($p> .05$) and that switch trials were responded slower than repetition trials ($\beta= .15$, $SE= .02$, $t= 7.10$, $p< .0001$). This effect was stronger for the L2 compared to the L3 ($\beta= .08$, $SE= .04$, $t= 2.06$, $p< .05$), indicating that switching costs for the L2 were larger than for the L3, i.e. *asymmetrical switching costs*.

Generally, analysis of naming latencies revealed that, in any given language pairing participants responded faster in the stronger compared to the weaker language. Switching costs were symmetrical for the L1-L2 language pairing and asymmetrical for both the L1-L3 and L2-L3 language pairings, i.e. with larger switching costs for the stronger compared to the weaker language.

Cross-language Interference

The mechanisms underlying trilingual language control were further explored by analyzing the speech errors participants made during the mixed-block of the picture naming task. Specifically, the analysis focused on non-target language intrusions while the target language was intended, i.e. "cross-language interference" (CLI). Differently from the bilingual case, where CLI can only come from one non-target language, in the trilingual case any of the non-target languages can act as source of interference while the target language is planned. Thus, the study of CLI can reveal to what extent the three languages are connected to each other (e.g., Ringbom, 1987; Williams & Hammarberg, 1998) and how trilinguals control languages (e.g., Festman, 2008). The total number of CLI for the three languages (L1, L2 and L3) in the two experimental trial types (Repetition and Switch) are illustrated in Table 4.

Table 4. Experiment 1: Mean percentages (raw values in parentheses) of Cross-Language Interference (CLI) for L1 vs. L2 vs. L3 in repetition and switch trials (upper part) and as a function of CLI - Direction, namely from L1 vs. from L2 vs. from L3 (lower part).

CLI - GENERAL			
	L1	**L2**	**L3**
Repetition	3.7% (6)	13.3% (22)	14.5% (24)
Switch	11.5% (19)	39.4% (65)	17.6% (29)
CLI - DIRECTION			
	L1	**L2**	**L3**
From L1			
Repetition	-	0.6% (1)	2.4% (4)
Switch	-	14% (23)	4.2% (7)
Total	-	14.6% (24)	6.6% (11)

From L2			
Repetition	2.4% (4)	-	12.1% (23)
Switch	5.5% (9)	-	13.2% (22)
Total	7.9% (13)	-	25.4% (42)
From L3			
Repetition	1.2% (2)	12.6% (21)	-
Switch	6.1% (10)	25.5% (42)	-
Total	7.3% (12)	38.1% (63)	-
TOTAL	15.2% (25)	52.7% (87)	32.1% (53)

In most of the cases, participants of the present study successfully selected words from the target language while ignoring words from the non-target languages. However, in some occasions (19.2% of the incorrect responses) CLI from any of the two non-target languages occurred. For example, suppose the L3 target language response "pumpkin", two types of errors were included in the CLI analysis: 1) Complete shift to a non-target language, e.g., the L2 non-target language response "kürbis"; 2) Integration of the non-target and target language, e.g., "kürbis.. pumpkin" or "kürb.. pumpkin". Both types of errors 1 and 2 were factored in the data analysis as “interference from the L2”.

The analysis of the CLI rates revealed that cross-language interference was more frequent in switch than in repetition trials (68.5% vs. 31.5%, β= 1.08, SE= .25, z= 4.34, $p< .0001$) and in the non-native languages compared to the native language (52.7% vs. 15.1%, for L2 vs. L1, β= 2.43, SE= .47, z= 5.18, $p< .0001$ and 32.2% vs. 15.1%, for L3 vs. L1, z= 2.52, $p< .05$). The difference between the two non-native languages was also significant (52.7% vs. 32.2% for L2 vs. L3, β= 1.07, SE= .36, z= 2.99, $p< .01$), indicating that overall the L2 was the language that mostly suffered from cross-language disturbances. Interestingly, the analysis revealed that most of the CLI came from the weaker L3 rather than from the stronger L2 or L1 (45.6% vs. 32%, for L3 vs. L2, β= 1.76, SE= .59, z= 2.99, $p< .01$ and 45.6% vs. 22.4%, for L3 vs. L1, β= 1.74, SE= .54, z= 3.23, $p< .01$). The difference between the amount of CLI coming from the L2 and from the L1 was not significant (32% vs. 22.4%, p> .05). Moreover, for each language CLI direction was explored, i.e. whether CLI came from the stronger or from the weaker of the two non-target languages. For example, if the L3 were the target language, L1 was considered the stronger non-target language and L2 the weaker non-target language. The analysis of the CLI direction showed that when the target language was a non-native language most of the CLI came from the other non-native language, i.e. from the weaker of the two non-target languages. Specifically, Pearson's Chi square test of independence for each of the target language showed that when L2 was the target language, most of CLI came from the L3 than from the L1 (72% vs. 28% respectively, x2= 17.48, p< .0001); when L3 was the intended language, interference mostly came from the L2 rather than from the L1 (79% vs. 21% respectively, x2= 18.13, p< .0001). In contrast, no difference was found in the amount of CLI coming from L2 and L3 when the L1 was the target language (52% vs. 48%, for L2 vs. L3, x2= .04, p= .84). This indicates that when the native language was intended, interference was equally coming from the stronger and from the weaker non-native language.

Overall, the analysis of CLI rates showed that the amount of cross-language interference was higher in switch than in repetition trials. The language mostly subject to unwanted interference was the L2, while the language that mainly served as source of interference was the weaker L3. With regard to CLI direction, when the two non-target languages were a native vs. a non-native language, CLI was more likely to come from the non-native language. However, when the native language was the target language, then the amount of CLI from the two non-native languages did not differ.

Discussion

As indicated by the language switching costs, Experiment 1 demonstrated that trilingual language switching is a costly process and that this was true for all the three languages under scrutiny, namely L1, L2 and L3. Moreover, the presence of mixing costs indicated that it was more difficult to name repetition trials' in a language switching context compared to a single language setting.

In Exp. 1, mixing costs were larger for the L1 compared to both the L2 and the L3, suggesting that compared to the single-language block, in the mixed-language block the native language was inhibited to a greater extent than the two non-native languages. One reason for that could be that participants unconsciously aimed at facilitating naming in the weaker non-native languages, by overall inhibiting the stronger native language more. This view would correspond to the idea of "proactive task adaptation", according to which the activation level of each language depends on the task goal (de Groot, 1998). Within this framework, proactive control is set before the onset of the task and can be distinguished from the type of control operating at the trial level, e.g., when a language switch is required. Specifically, the Dual Mechanisms of Control (DMC) framework proposes that cognitive control relies upon two different control mechanisms, i.e. *proactive control* for anticipating and preventing interference and *reactive control* for resolving interference after its onset (Braver, 2012). Recent research has confirmed that proactive and reactive language control are indeed reflected by language mixing and switching costs respectively (Ma et al., 2016).

With regard to the latter, in the present experiment the overall amount of switching costs was smaller for the L3 than for the L1 or the L2, while no difference was found between L1 and L2. This finding challenges the general assumption of the inhibitory dominance-related account (cf. Meuter & Allport, 1999; Philipp et al., 2007), according to which switching costs are determined by language dominance, i.e. the stronger the language the larger the switching costs. Switching costs within language pairings (i.e. L1-L2, L1-L3 and L2-L3) did also only partially confirm the dominance-related hypothesis, in that results revealed asymmetrical switching costs for the L1-L3 and L2-L3 language pairings, with larger switching costs for the stronger than for the weaker language as predicted by the dominance-related account, but not for the L1-L2 language pairing, where switching costs were symmetrical. This result could be explained by the fact that the participants of the present study were so proficient in their L2 to minimize the dominance difference between the L1 and the L2, leading to symmetrical switching costs. However, this hypothesis is hard to support if we consider that the dominance difference between the L2 and the L3 was even smaller (as revealed by the results of the verbal fluency, the self-rating and the single-language naming tasks) and switching costs for the L2-L3 language pairing were asymmetrical. According to the language specific mechanism (Costa & Santesteban, 2004; Costa et al. 2006), a relatively high L2 proficiency level can boost the shift from an inhibitory mechanism to a mechanism where inhibition of the non-intended languages is not necessary. Within this framework, the symmetrical switching costs found in L1-L2 language pairing could be considered evidence for the shift to a language specific mechanism. However, the findings that switching costs were asymmetrical in the L1-L3 and L2-L3 language pairing do not support the language specific hypothesis, according to which once the language specific mechanism has kicked in, it will operate also in other instances of language switching involving fairly strong languages. Crucially, both symmetrical and asymmetrical switching costs can be accounted within the IC-Model, with asymmetrical switching costs indicating different amount of inhibition for the two languages and therefore, different reactivation costs, and symmetrical switching costs indexing similar amount of suppression and hence, similar reactivation costs. Consequently, it can be assumed that, for reasons to be explored, in the L1-L2 condition the two languages suffered the same amount of inhibition. Importantly, some other studies have offered a complementary explanation to the inhibitory view to account for asymmetrical switching costs, which is *persisting activation* from the previous trial into the current relevant trial (Philipp et al. 2007). In this scenario, performing in the weaker language requires more activation than performing in the stronger language and in both cases, the residual activation will spread into the upcoming trial. It follows that in case of language switching, persisting activation from the weaker language will be stronger than persisting activation from the stronger language, so that performance in the stronger language is more hampered than in the weaker language (asymmetrical switching costs). Although this explanation suits well with language switching performance, it cannot easily account for the effects of language *backward inhibition* found in both task and language switching studies (e.g., Mayr & Keele, 2000; Philipp & Koch, 2009; Philipp et al. 2007), where performance in n-2 repetition trials (e.g., A-B-*A*) is found to be slower than in no repetition trials (e.g., C-B-*A*). This effect is explained by the fact that persisting inhibition of a previously relevant task hampers performance when the same task is reactivated in an n-2 trial (in this example task “A”).

As outlined in the introduction, languages’ activation level can be modulated by several more factors than just by language proficiency. The analysis of the speech errors represented a useful tool to deeper investigate language switching behaviour. In this respect, the analysis showed an unexpected pattern of interference, revealing that the L2 was the language that mostly suffered from cross-language interference and that most of the unwanted influences were coming from the weaker L3. Why is the L2 the most vulnerable language in terms of cross-language interference and not the weaker L3? Moreover, why does the majority of the cross-language interference come from the weaker L3 rather than from the stronger L1 or L2?

In the present experiment, Italian was the native language (L1) and German and English the two non-native languages (L2 and L3 respectively). Thus, the language pair German-English not only shared the foreign language factor, but they were also typologically closer compared to the language pairs Italian-German and Italian-English. In this scenario, according to the language typological distance hypothesis, the L2 German should act as the strongest source of interference for the L3 English. However, in Exp. 1 we observed the opposite pattern of cross-language influences, which is more interference from the weaker L3 into the stronger L2.

As suggested by Los (1996), in case of enhanced task uncertainty such as in a mixed-language block, speakers tend to prepare for the “worst possible case”. It is reasonable to suppose, therefore, that participants might have relied on involuntary naming strategies to be prepared for the most difficult situation, which in this case is naming in the L3. Thus, to avoid strong interference from the L2 into the weaker L3, trilinguals might have inhibited the L2 more, leading to greater switching costs for the L2 than those predicted by a dominance-related account. This would explain why overall L2 switching costs are as large as for the L1 (rather than smaller than the L1 as predicted by the dominance-related hypothesis), and also why L2 switching costs in the L1-L2 language pairings are symmetrical rather than asymmetrical (which would imply smaller switching costs for the L2 compared to the L1). Finally, the attempt to suppress the stronger L2 in favor of the weaker L3 would explain why the L2 is the language, which mostly suffers from cross-language transfer and not the weaker L3, and why cross-language interference mostly come from the weaker L3 rather than from the stronger non-native language L2.

This result suggests that the degree of language inhibition can be modulated by other factors, such as typological closeness of the L2 to the L3. However, it would be also reasonable to argue that the observed language switching behavior could have been boosted by the shared foreign language status of L2 and L3 and not by their typological closeness. Therefore, Exp. 2 aims at disentangling the effect of language typological proximity from that of language status in a similar trilingual language switching task.

Experiment 2

In Exp. 1, it was proposed that trilinguals whose non-native languages were typologically closer inhibited the L2 more in order to facilitate performance in the weaker L3. Exp. 2 aims at distinguishing the role of language lexical distance (typologically closer vs. typologically less close language) from that of language status (native vs. non-native language) in trilingual language control. To this end, unbalanced trilinguals with German as their L1, English as their L2 and Italian as their L3 were tested in a picture naming task involving language switching. The fact that the two non-native languages English and Italian were typologically more distant than the native German and the non-native English, allowed us to investigate whether the observed effect in Exp. 1 were due to the typological closeness of L2 and L3 or by their shared foreign language status.

Participants

Thirty-two native speakers of German (14 men, mean age: 29.5) with good proficiency of English and Italian took part in the experiment. They were recruited by means of flyers, social networks or via the university blackboard and were paid for their participation. All the participants gave informed consent and filled out the same background questionnaire used in Experiment 1 to gain information on their linguistic and demographic history. Twenty-three participants reported having spent some time in a non-German speaking country for study or work reasons (i.e. Australia, Canada, Czech-Republic, France, Korea, Italy, Luxemburg, Norway, Spain, The Netherlands and USA), whereas the rest reported having spent their entire life in Germany and having visited a foreign country only for short vacation. Most of them were university students, from which some of them were specializing in English and/or Italian language and culture. Some of the participants had already accomplished their studies, whereas only few of them did not attend university and were employees. All the participants were born and raised in a German monolingual environment and started learning a foreign language at school (except for one participant that started learning a foreign language by the age of 4). Specifically, they reported learning English from the mean age of 10 years (4-14 years, sd: 1.9) and Italian from the mean age of 22.7 years (13-48 years, sd: 9.5). On a daily basis, they reported speaking German most of the time (86.8% - with their partner, children, family, extended family, friends, university peers and working colleagues) rather than English (8% - with friends, university peers and working colleagues) or Italian (5.2% - with their partner, extended family, friends, university peers and working colleagues). The oral and written exposure (watching TV, listening to the radio and reading books, newspapers, etc.) did also differ between the three languages, with German being the language they were exposed the most (73.4%) followed by English (18.9%) and by Italian (7.7%).

To gain some insight into their languages' proficiency, participants were asked to self-rate their ability to speak, comprehend, write and read in the three languages (German, English and Italian). The self-rating was based on a 10-point scale, in which 0= no knowledge of the language and 10= proficient like a native speaker. Mean scores of the self-rating task are reported in Table 5.

Table 5. Experiment 2: Means scores (standard deviations in brackets) of the self-rating task for speaking, comprehension, writing and reading skills in German, English and Italian.

	Speaking	*Comprehension*	*Writing*	*Reading*	*Mean*
German	10 (0)	10 (0)	10 (0)	10 (0)	10 (0)
English	7.2 (1.1)	7.8 (1)	7.1 (1.1)	8.1 (1.1)	7.5 (1.1)
Italian	5.1 (1.4)	5.6 (1.7)	5.1 (1.3)	6.1 (1.3)	5.5 (1.4)

Overall, participants considered themselves as proficient as native speakers of German, advanced speakers of English with a lower proficiency in Italian. Sliding contrasts on the self-rating scores revealed that the difference between the three languages was significant (β= 2.42, SE= .17, t= 13.77, p< .0001 for L1 vs. L2 and β= 2.07, SE= .24, t= 8.63, p<. 0001 for L2 vs. L3). To further investigate languages' proficiency, participants were administered the same verbal fluency task used in Experiment 1. Mean scores of the verbal fluency task are reported in Appendix A. Overall, participants named more words in German, followed by English and then by Italian. The difference between the three languages was significant as revealed by the analysis of the mean scores (β= 19.8, SE= 3.19, t= 6.22, p< .0001 for L1 vs. L2 and β= 18.8, SE= 3.1, t= 5.90, p< .0001 for L2 vs. L3).

Based on the results of both the language self-rating and verbal fluency tasks, German was considered the L1, English the L2 and Italian the L3 of the trilinguals tested in the present experiment.

Materials and Method

The same materials and procedure of Experiment 1 were used here.

Data Cleaning

The parameters used for data cleaning were the same as in Experiment 1. Because of low accuracy score (participants' and items' accuracy rate ranged respectively from 57% to 99% and from 64% to 98%), data from two participants and from one item were excluded from further analysis. For the remaining data, accuracy scores and naming latencies were measured. The analysis of the accuracy rates included both correct and incorrect responses (13.4% of the data points), while only correct responses were used in the naming latencies analysis. Before the analysis of naming latencies, data were screened for outliers where implausible fast responses (< 100ms) and extremely slow responses (>3500ms) were excluded from subsequent analysis (0.12% of the data). Despite data cleaning, visual inspection revealed that the data were positively skewed. Therefore, to approximate normality the data were log transformed as suggested by the boxcox function of the MASS package in R (Venables & Ripley, 2002). The visual comparison between logarithmic and inverse transformed data, confirmed the former one being the more appropriate to approach normality. Finally, the procedure for selecting and comparing the best-fit models for both naming latencies and accuracy data was the same as for Experiment 1.

Results

Table 6 shows mean naming latencies and accuracy scores for both single-language and mixed-language block.

Table 6. Experiment 2: Mean naming latencies (standard deviations in brackets) and accuracy rates (in percentages) for correct responses of L1 vs. L2 vs. L3 in single-language block (upper part) vs. mixed-language block (lower part). In the mixed-language block, mixing costs (calculated as the difference between trials in the single-language block and repetition trials in the mixed-language block) and switching costs (calculated as the difference between repetition and switch trials) for L1, L2 and L3 are reported.

		L1	**L2**	**L3**
Single-language block				
Mean	*Naming latencies*	826ms (273)	934ms (356)	1025ms (366)
	Accuracy *rates*	97%	92%	69%
Mixed-language block				
Repetition	*Naming latencies*	893ms (268)	1007ms (335)	1034ms (312)
	Accuracy rates	97%	91%	77%
MIXING COSTS		*67ms*	*73ms*	*9ms*
Switch	*Naming latencies*	1161ms (318)	1154ms (349)	1214ms (312)
	Accuracy rates	94%	89%	75%
SWITCHING COSTS		*268ms*	*147ms*	*180ms*
Mean	*Naming latencies*	1024ms (322)	1080ms (350)	1123ms (347)
	Accuracy rates	95%	90%	76%

Accuracy rates. In the single-language block, the accuracy rate analysis revealed that participants made more errors in the weaker L3 compared to the stronger L1 and L2 (β= 3.27, SE= .54, z= 6.05, p< .05 for L3 vs. L1 and β= 2.23. SE= .40, z= 5.47, p< .0001 for L3 vs. L2); the analysis also showed a trend for the L1 to be more accurate than the L2, which, however, did not reach significance (p=.07 for L1 vs. L2). In the mixed-language block, participants were significantly more accurate in the native language compared to the non-native languages (β= 0.79, SE= .19, z= 4.13, p< .0001 for L1 vs. L2 and β= 1.96, SE= .27, z= 7.02, p< .0001 for L1 vs. L3). The difference between the two non-native languages was also significant, with L2 being responded more accurately than L3 (β= 1.16, SE= .28, z= 4.11, p< .0001). Concerning the Block Type, there was no difference between the overall accuracy rates in single-language block compared to the repetition trials in the mixed-language block (p> .05). However, compared to L1 and L2, responses in L3 were significantly more accurate in the mixed-language compared to the single-language block (β= .70, SE=.30, z= 2.35, p< .05 for L3 vs. L1 and β= .96, SE= .23, z= 4.13, p< .0001 for L3 vs. L2). No significant difference was found between L1 and L2 (p> .05).

With regard to Trial Type in the mixed-language block, switch trials were more error prone than repetition trials (β= .39, SE= .13, z= 3.01, p< .01) and this effect was independent from Language (all p> .05).

Overall, the accuracy rates analysis showed that participants made more errors in the weaker L3 than in the stronger L1 or L2 and in switch trials compared to repetition trials. There was no accuracy score difference between trials in the single-language block and repetition trials in the mixed-language block. However, L3 responses showed significantly better performance in the mixed- compared to the single-language block.

Naming latencies. In single-language block, the analysis of the naming latencies showed that the native language was responded faster than the non-native languages (β= .13, SE= .04, t= 2.85, p< .01 for L1 vs. L2 and β= .23, SE= .04, t= 4.66, p< .0001 for L1 vs. L3) and that naming latencies in the stronger non-native language were faster than in the weaker one (β= .09, SE0 .04, t= 2.26, p< .05 for L2 vs. L3). Similarly, in the mixed-language block participants were faster in their native language compared to the non-native languages (β= .05, SE= .01, t= 2.93, p< .01 for L1 vs. L2 and β= .09, SE= .02, t= 4.44. p< .0001 for L1 vs. L3) and in the stronger non-native language than in the weaker one (β= .04, SE= .02, t= 2.05, p< .05 for L2 vs. L3). Concerning Block Type, trials in the single-language block were responded faster than repetition trials in the mixed-language block (β= .07, SE= .02, t= 3.33, p< .001), indicating mixing costs in the mixed-language compared to the single-language block. Differently from what raw means in Table 6 suggest, the amount of mixing costs on a log scale did not differ among the three languages (all p> .05).

With regard to Trial Type in the mixed-language block, switch trials were responded slower than repetition trials (β= .19, SE= .01, t= 11.15, p< .0001). The effect was larger for the L1 compared to either L2 or L3 (β= .11, SE= .03, t= 3.45, p< .05 for L1 vs. L2 and β= .09, SE= .03, t= 2.87, p< .01 for L1 vs. L3), while no difference was detected between the L2 and L3 (p> .05). This indicates that overall switching costs were larger for the native language than for the non-native languages. The significant interaction of Trial Type by Preceding Language (β= .06, SE= .02, t= 2.52, p< .05) indicated that switching costs were modulated by the language of the preceding trial. The three-way interaction of Language, Trial Type and Preceding Language did not reach significance (p= .21 for L1 vs. L2; p= .09, for L1 vs. L3 and p= .64, for L2 vs. L3).

Generally, the analysis showed that in both the single- and in the mixed-language block participants' naming latencies depended on language dominance, i.e. responses were faster in the L1 and slower in the L3. In the mixed-language block, switch trials were responded slower than repetition trials, while the latter were responded slower compared to trials in the single-language block. Finally, in each language the amount of switching cost was modulated by the language of the preceding trial.

Preceding Language and Switching costs

To disentangle the effects of preceding trial language on switching costs and compare them to Exp.1, post-hoc analyses were run on the data split in three into three language pairings subsets (L1-L2, L1-L3 and L2-L3). By doing this, it was possible to test the dominance-related hypothesis, according to which language proficiency differences between two languages affect switching costs pattern. Mean naming latencies and accuracy rates for each language pairing are illustrated in Table 7.

Table 7. Experiment 2: Mean naming latencies (standard deviations in brackets) and accuracy rates (in percentages) for correct responses of the language pairing L1-L2, L1-L3 and L2-L3 in repetition and switch trials. For each language pairing, switching costs (calculated as the difference between repetition and switch trials) are reported in italics.

L1-L2 language pairing

		L1	**L2**
Repetition	*Naming latencies*	893ms (268)	1007ms (335)
	Accuracy rates	97%	91%
Switch	*Naming latencies*	1120ms (315)	1166 (361)
	Accuracy rates	95%	89%
Mean	*Naming latencies*	967ms (303)	1059ms (352)
	Accuracy rates	96%	90%
SWITCHING COSTS		*227ms*	*159ms*

L1-L3 language pairing

		L1	**L3**
Repetition	*Naming latencies*	893ms (268)	1034ms (312)
	Accuracy rates	97%	77%
Switch	*Naming latencies*	1202ms (316)	1207 (381)
	Accuracy rates	93%	76%

Mean	*Naming latencies*	933ms (319)	1091ms (346)
	Accuracy rates	96%	77%
SWITCHING COSTS		*309ms*	*173ms*

L2-L3 language pairing

		L2	**L3**
Repetition	*Naming latencies*	1007ms (335)	1034ms (312)
	Accuracy rates	91%	77%
Switch	*Naming latencies*	1142ms (337)	1221ms (331)
	Accuracy rates	90%	74%
Mean	*Naming latencies*	1052ms (342)	1095ms (330)
	Accuracy rates	91%	76%
SWITCHING COSTS		*137ms*	*187ms*

Accuracy rates. Consider first the L1-L2 language pairing. The analysis of the accurate rates showed that responses were more accurate in the L1 than in the L2 (β= .99, SE= .18, z= 5.27, p< .0001) and in repetition compared to switch trials (β= .48, SE0 .18, z= 2.57, p< .01). The interaction between Language and Trial Type was not significant (p> .05). With regard to the L1-L3 language pairing, performance was more accurate in the L1 than in the L3 (β= 2.00, SE= .33, z= 6.02, p< .0001). More errors were detected in switch than repetition trials (β= .48, SE= .18, z= 2.58, p<. 01) and this effect was larger for the L1 compared to the L3 (β= .76, SE= .37, z= 2.03, p= .05). Finally, in the language pairing L2-L3 error rates were lower for the L2 compared to the L3 (β= 1.24, SE= .31, t= 3.94, p< .0001, while neither the main effect of Trial Type nor the Language by Trial Type interaction were significant (all p< .05).

Naming latencies. In the language pairing L1-L2, the analysis of the naming latencies indicated that responses in the L1 were faster than in the L2 (β= .11, SE= .02, t= 4.12, p< .0001) as well as in repetition compared to switch trials (β= .23, SE= .03, t= 7.67, p< .0001). The interaction between Language and Trial Type was also reliable (β= .08, SE= .04, t= 2.07, p< .05), indicating larger switching costs for the L1 compared to the L2, i.e. *asymmetrical switching costs.* In the language pairing L1-L3, naming latencies were faster in L1 than in L3 responses (β= .06, SE= .02, t= 2.69, p< .05). Furthermore, switch trials were responded slower than repetition trials (β= .23, SE= .02, t= 9.41,

$p < .0001$) and this effect was larger for the L1 compared to the L3 (β= .15, SE= .04, t= 3.68, $p < .01$), i.e. *asymmetrical switching costs*. Finally, in the language paring L2-L3 results showed faster responses for the L2 than for the L3 (β= .04, SE= t= 2.25, $p < .05$) and in repetition than in switch trials (β= .16, SE= .19, t= 8.17, $p < .0001$). Differently from what suggested by the raw means in Table 7, on a log scale the effect of Trial Type was not modulated by Language ($p > .05$), indicating same degree of switching costs for the two languages, i.e. *symmetrical switching costs*.

To summarize, the analyses of the naming latencies on the language pairings L1-L2, L1-L3 and L2-L3 showed that within any language pairing participants responded faster in the stronger than in the weaker language and in repetition relative to switch trials. Switching costs were asymmetrical in the language pairings L1-L2 and L1-L3 (with larger switching costs for the native language compared to the non-native) and symmetrical for the language pairing L2-L3.

Cross-language Interference

The total number of CLI for the three languages (L1, L2 and L3) in the two experimental trial types (Repetition and Switch) are illustrated in Table 8.

Table 8. Experiment 2: Mean percentages (raw values in parentheses) of Cross-Language Interference (CLI) for L1 vs. L2 vs. L3 in repetition and switch trials (upper part) and as a function of CLI - Direction, namely from L1 vs. from L2 vs. from L3 (lower part).

CLI - GENERAL			
	L1	**L2**	**L3**
Repetition	4.4 % (4)	13.3% (12)	10% (9)
Switch	15.6% (14)	31.1% (28)	25.6% (23)
CLI - DIRECTION			
	L1	**L2**	**L3**
From L1			
Repetition	-	6.7% (6)	1.1% (1)
Switch	-	15.6% (14)	14.4% (13)
Total	-	22.3% (20)	15.4% (14)

From L2			
Repetition	2.2% (2)	-	8.9% (8)
Switch	10% (9)	-	11.1% (10)
Total	12.2% (11)	-	20% (18)
From L3			
Repetition	2.2% (2)	6.6% (6)	-
Switch	5.6% (5)	15.6% (14)	-
Total	7.8% (7)	22.2% (20)	-
TOTAL	20% (18)	33.4% (40)	

In most cases, participants of Experiment 2 responded in the target language and only in some occasions (12.3% of the incorrect responses), they experienced CLI from a non-target language while the target one was planned. As expected, the analysis showed that CLI were more frequent in switch relative to repetition trials (72.2% vs. 27.8%, β= 1.0, SE= .28, z= 3.54, p< .001). Overall, the amount of CLI was higher in the non-native languages than in the native language; however, while the difference of L2 vs. L1 was significant (44.4% vs. 20%, β= .81, SE= 34, z= 2.38, p< .05), the L3 vs. L1 comparison did not reach significance (35.6% vs. 20%, β= .59, SE= .35, z= 1.69, p= .09). Finally, the frequency of CLI did not differ between L2 and L3 (44.4% vs. 35.6%, p> .05). With regard to the source of interference, raw percentages suggest that CLI were mostly coming from the native language and less from the non-native languages (37.8% for L1, 32.2% for L2 and 30% for L3); however, the difference was not significant (all p> .05). As in Experiment 1, it was measured whether given a target language, CLI were mostly coming from the stronger or the weaker non-target language. The analysis of CLI direction revealed CLI were coming to a similar extent from the stronger and weaker non-target language in any given target language (i.e. for L1, x2= .88, p= .34; for L2, x2= 0, p= 1 and for L3, x2= .43, p= .56).

Overall, the analysis showed that switch trials were more subject to cross-language interference than repetition trials. Moreover, results revealed that L2 was the language that mostly suffered from

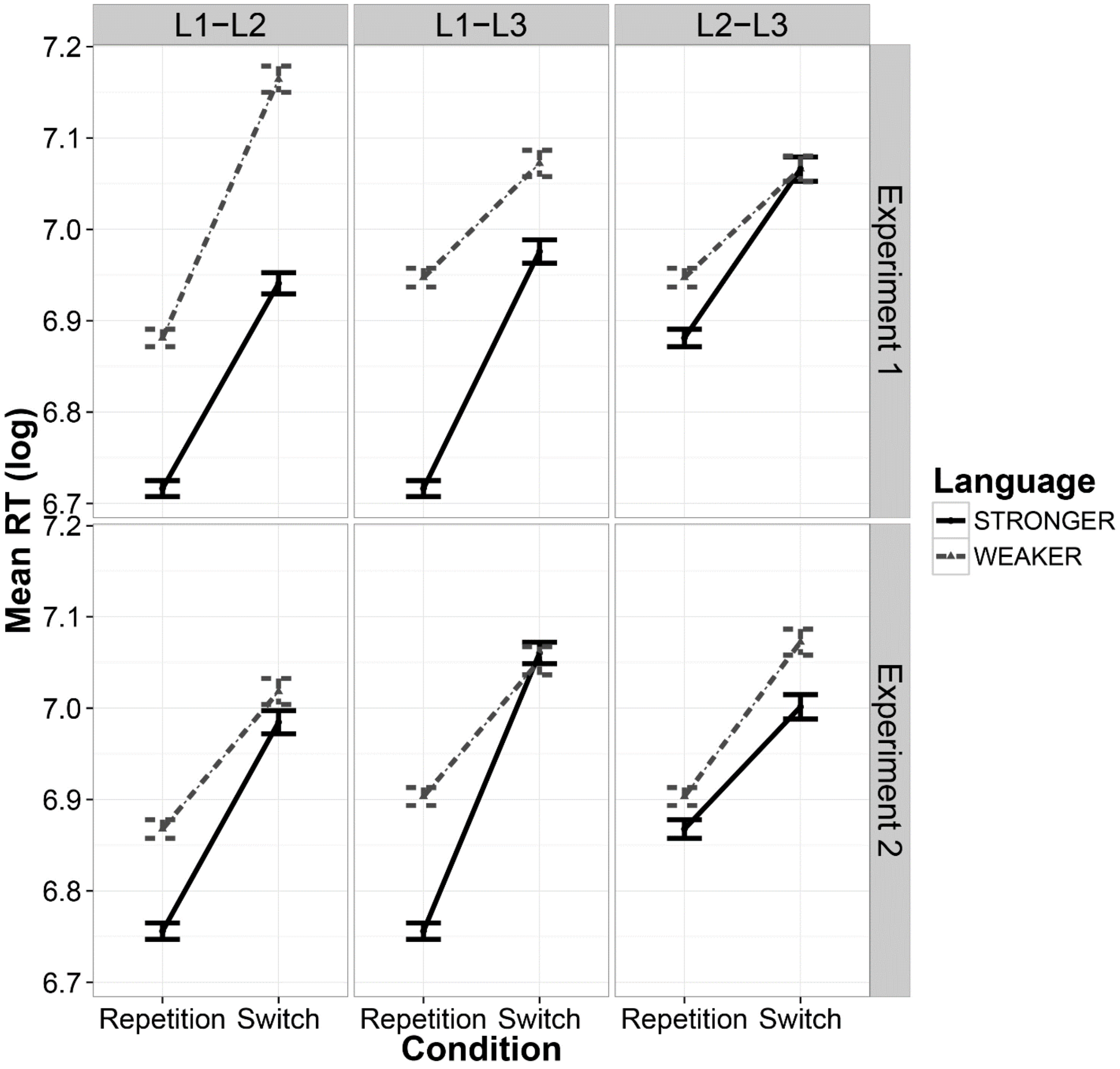
CLI and that in each of the three languages, CLI were equally coming from the stronger and the weaker non-target language.

Discussion

Similarly, to Exp. 1, Exp. 2 showed that context-driven language switch is a costly process for trilingual speakers and that performance on repetition trials is worse in the language switching context compared to the single language setting. Dissimilarly from Exp. 1, where mixing costs were larger for the native language compared to the non-native languages, in Exp. 2 the amount of mixing costs did not differ among the three languages. This suggests that unlike the findings in Exp. 1, in Exp. 2 participants did not adapt languages' activation level to facilitate naming in the weaker languages, or at least not to the same extent trilinguals in Exp. 1 did. Thus, it might be speculated that while in Exp. 1 trilinguals needed to unconsciously facilitate naming in the two non-native languages because of their conflicting typological closeness, as shown by the larger mixing costs found for the native compared to the non-native languages, in Exp. 2 the strong conflict between the two non-native languages was less present, leading to a similar amount of mixing costs for the three languages. Additionally, in both groups of trilinguals, L3 proportions of accuracy rates were higher in the mixed- than in the single-language block and the effect was stronger in Exp. 2 than in Exp. 1 (i.e. significant interaction of Language by Block Type for Exp. 2, but only a main effect of Block Type likely driven by the L3 in Exp. 1). These patterns support the idea that the system strategically tries to compensate for languages' differences during language switching, by fostering weaker languages and/or penalizing stronger ones. However, further studies are necessary to determine which factors influence bilinguals' and trilinguals' proactive adaptation to the task.

With regard to switching costs, in Exp. 2 the overall amount of switching costs was greater for the L1 compared to the L2 and the L3, while there was no difference between the two non-native languages. Moreover, the switching costs pattern within language pairings (i.e. L1-L2, L1-L3, L2-L3) showed asymmetrical switching costs for L1-L2 and L1-L3, with larger costs for the native compared to the non-native language, and symmetrical switching costs for L2-L3. Therefore, despite the fact that participants were more proficient in the L2 than in the L3 (as revealed by the verbal fluency, the self-rating and the single-language naming tasks), the amount of switching costs for the two non-native languages was comparable. This finding fails to support the dominance-related account (cf. Meuter & Allport, 1999; Philipp et al., 2007), according to which language switching costs increase along with language proficiency. In contrast, symmetrical switching costs for language pairs with different proficiency levels are accounted for by the language specific hypothesis (e.g., Costa & Santesteban, 2004), and are believed to indicate that lexical selection has moved from an inhibitory process to a language selection mechanism that does not require inhibition. However, the asymmetrical switching pattern found in both the L1-L2 and the L1-L3 language pairing cannot be explained within the language specific framework. The overall switching costs pattern for the trilinguals tested in Exp. 1 and Exp. 2 are shown in Figure 1.

Fig. 1. Mean naming latencies (log transformed) for correct responses in repetition and switch trials for the language pairing L1-L2, L1-L3 and L2-L3 in Experiment 1 (upper part) and Experiment 2 (lower part). For each language pairing, the stronger language is indexed by darker solid line and the weaker language by the lighter dashed line.



As for Exp. 1, the analyses of cross-linguistic transfers in Exp. 2 revealed interesting patterns of interference. In particular, results showed that in this group of trilinguals most of the interference was coming from the L1 German and that L2 English was the language that mostly suffered from unintended transfers. Similarly, Festman (2008) found that speakers of German (L1), English (L2) and French (L3) performing a trilingual picture naming task showed higher cross-language interference in the L2 English, and that the L1 German was the strongest source of interference. These findings seem to suggest that the typological closeness of L1 and L2 might affect language control performance in trilinguals. More precisely, it might be speculated that in this group of trilinguals, the most difficult situation would be naming in the L2 English because of the linguistic closeness to the L1 German. If this was the case, it should be possible to measure either a stronger inhibition of the L1 or a facilitating naming strategy for the L2, i.e. less inhibition on the L2 as the one predicted by the dominance-related account. In this regard, it is interesting to note that on average L2 switching costs were much smaller compared to both L1 and L3. Hence, a facilitating strategy in favor of the L2 English might explain why the amount of switching costs does not differ between L2 and L3 and why switching costs in the L2-L3 language pairing are symmetrical. Moreover, results of Exp. 2 suggest that the switching pattern observed in Exp. 1, are mostly driven by the typological closeness of the two non-native languages (i.e. German and English) than by their shared foreign language status. Indeed, if it was the case that as default trilinguals suppress the stronger non-native language in order to facilitate performance in the weaker non-native language, it should have been possible to see this switching behavior also in Exp. 2. On the contrary, Exp. 2 did not show a stronger suppression of the L2, but rather a weaker one than the one predicted by a dominance-related account.

Finally, are there other differences between the two groups of trilinguals that might have led to the occurrence of the effects found in the study?

The main difference between the two groups is the language of the environment, since trilinguals of Exp. 1 were living in a L2 environment, while trilinguals in Exp. 2 were living in a L1 environment. Albeit, in both trilinguals' groups the L1 was the language mostly used on a daily basis, followed by the L2 and this by the L3, in absolute percentages, there was a difference of ca. 10-12% in the amount of L1 and L2 language use between the two groups (i.e. Exp. 1, 63.8% - 30% for L1-L2 and Exp. 2, 73.4% - 18.4% for L1-L2). The amount of languages' passive exposure did also differ between the two groups, in that trilinguals of Exp. 1 were mostly exposed to their L2 and trilinguals of Exp. 2 to their L1. Although the difference in the amount of languages' usage is relatively small and language exposure is more likely to play a role in language reception rather than in language switching in production, we can still suppose that in Exp. 1 the higher exposure and frequency of use of the L2 had strengthened its lexical representation, making the L2 in Exp. 1 more automized compared to the

L2 in Exp. 2. However, language strength of representation alone cannot explain why, for example, in Exp. 1, the L2 is the language that mostly suffered from cross-language interference and why interference mostly came from the weaker L3.

Conclusion

The goal of the present study was to investigate whether language control in trilinguals with a high L2 proficiency level is language specific (e.g., Costa & Santesteban, 2004; Costa et al. 2006) or whether it relies on an inhibitory language non-specific mechanism (e.g., Green, 1998; Meuter & Allport, 1999). Moreover, the study assessed whether trilinguals' language control is affected by language typological similarity. With this aim, cross-language interference and language switching costs were measured in two groups of unbalanced trilinguals, with a relatively high L2 proficiency level, while performing a trilingual picture naming task. Overall, results on cross-linguistic transfers suggest that during trilingual language switching all three languages competed for selection. This was clearly indicated by the findings that each of the three languages suffered from unwanted interference from the other two languages. With regard to the language switching patterns, they showed that language switching costs tended to be larger for the stronger compared to the weaker languages. This result can be accounted by a dominance-related inhibitory mechanism (e.g., Green, 1998; Meuter & Allport, 1999), according to which during language selection non-relevant languages are inhibited and that the amount of inhibition increases along with language dominance. Results of the present study suggest, therefore, that language selection in the group of trilinguals tested is language non-specific.

Additionally, while the study shows that language dominance is a strong influential factor in language switching behavior, it also suggests it is not the only one. Specifically, Exp. 1 revealed that language switching patterns were asymmetrical in the L1-L3 and L2-L3 language pairings, with larger switching costs for the stronger compared to the weaker language. Similarly, Exp. 2 showed asymmetrical switching costs for L1-L2 and L1-L3, with larger switching costs for the stronger than the weaker language. These findings are in line with an inhibitory dominance-related hypothesis, which predicts larger switching costs for the stronger than for the weaker language. However, the symmetrical switching costs found in the L1-L2 and the L2-L3 language pairings, in Exp. 1 and Exp. 2 respectively, suggest that factors other than only language dominance might play a role in trilinguals' language control. Specifically, converging evidence shows that the typological closeness between two languages might have influenced the degree to which those languages were activated, and thus suppressed. In Exp. 1, L2 German and L3 English were the typologically closer languages for the group of trilinguals tested. Results showed that the overall amount of switching costs for the

L2 was larger than predicted by dominance-related account, and switching costs in the L1-L2 language pairing were symmetrical and not asymmetrical, with weaker costs for the L2 compared to the L1, as expected in an inhibitory framework. For this experiment, it has been proposed that in order to prevent unwanted interference from the L2 into the typologically closer L3, trilinguals might have relied on unconscious naming strategies, such as suppressing the "disturbing" L2 more strongly. This greater suppression would explain the increased L2 switching costs overall in the mixed-language block as well as in the L1-L2 pairings. In Exp. 2, L1 German and L2 English were the typologically closer languages for the group of trilinguals tested. In this experiment, the L2 showed a weaker degree of switching costs than the one predicted by a dominance-related account. Moreover, switching costs in the L2-L3 language pairing were symmetrical and not asymmetrical, with larger costs for the L2 than the L3, as postulated by a dominance-related account. Altogether, these results indicate that the L2 of this group of trilinguals experienced a weaker degree of inhibition as the one predicted by a dominance-related inhibition. In particular, in the present paper it has been suggested that because of the typological closeness of the L1 German and the L2 English, trilinguals of Exp. 2 might have unconsciously facilitated their L2 by suppressing it less strongly. A weaker suppression of the L2 would explain why switching costs were overall smaller for this language than expected based on a dominance-related account and why L2-L3 switching costs were symmetrical, and not larger for the L2 than for the L3 as the inhibitory account would predict. Considered this, the symmetrical switching costs found between a stronger and a weaker language in Costa and colleagues' (2004, 2006) studies might have indicated that the two languages were experiencing a similar amount of inhibition and not that speakers' lexical selection process had undergone a shift from an inhibitory to a language specific mechanism as proposed by the authors. As mentioned, language control in these trilinguals was investigated in a bilingual setting (i.e. two languages at one time) rather than in a trilingual one (all three languages simultaneously). Therefore, it is reasonable to suppose that the experimental setting might have influenced the activation level (and thus the amount of inhibition required) of the languages under scrutiny. Indeed, both Costa and Santesteban (2004, Exp. 4) and Costa et al. (2006, Exp. 2) showed that overall (i.e. repetition and switch trials) the weaker language was responded faster than the stronger one, suggesting that speakers were indeed relying on unconscious strategies to facilitate naming in the weaker language.

To conclude, the results of the present study suggest that language selection in unbalanced trilinguals with a high L2 proficiency is language non-specific and that this mechanism predominantly relies on an inhibitory process that is dominance-related. However, the study also suggests that dominance alone is not enough to explain trilinguals' language control mechanisms. Specifically, it showed that typological closeness to the L3 modulates trilinguals' language control (Exp. 1) and that the effect of language typological proximity is not restricted to the L3, but extend to L1 and L2 as well (Exp. 2). Altogether, the results of the present research indicate that language control is a flexible mechanism that strategically adapts to the context. In particular, it seems that the controlling system aims at preventing potential conflicting situations, such as typological closeness between languages, by adjusting languages' activation levels. However, in order to be able to determine the strategic mechanisms underlying language control, the effects of language typological similarity need to be implemented in future models of trilinguals' language control. Consequently, further research is necessary to systematically investigate the role of language typological proximity on trilinguals' language control, for example, by comparing the effect of typological closeness to the L3 of either the L1 or the L2 or to test trilinguals, whose three languages are similarly close or similarly distant to each other.

Appendix A. Verbal fluency task

In the verbal fluency task participants are asked to name as many words as possible starting with a given letter (phonemic fluency) or belonging to a determined semantic category (semantic fluency) in 1 minute time. For the phonemic subtest, two letters were selected for each language, according to their frequency as word initial letters. Frequency values for Italian have been taken from Lusetti (2001), those for German from *Vogelsang (*2003) and frequency values for English from Borkowski, Benton and Spreen (1966). For the semantic subtest, two semantic categories were selected for each language, with one supposed to elicit more responses than the other (see also Schwieter & Sunderman, 2008; Stokholm, *Jorgensen & Boge,* 2013). Table A1 and Table B1 illustrate the details of the verbal fluency task in Experiment 1 and Experiment 2 respectively.

Table A1. Experiment 1: Mean scores (standard deviation in brackets) of the verbal fluency task in the three languages (Italian, German and English). For the phonemic subtest (on the top), the Letter 1 represents a more frequent and the Letter 2 a less frequent word initial letter in a given language (frequency percentage in brackets). For the semantic subtest (at the bottom), Category 1 represents the category, which elicits a higher number of responses and Category 2 indicates the category that elicits a smaller number of responses.

Phonemic subtest			
Language	Italian	German	English
Letter 1	D (4.9%)	M (5.5%)	W (5.7%)
Letter 2	F (4.5%)	T (4%)	L (3.9%)
Total frequency	9.4%	9.5%	9.6%
Score	29.7 (6.6)	22 (5.8)	20 (5.4)

Semantic subtest			
Category 1	Cities	Proper names (persons)	Countries
Category 2	Clothes	Supermarket	Animals
Score	51.2 (9.9)	37.5 (8.7)	32.5 (12.6)
Total score	80.5 (16.5)	59.5 (14.8)	44.5 (18.1)

Table A2. Experiment 2: Mean scores (standard deviation in brackets) of the verbal fluency task in the three languages (German, English and Italian). For the phonemic subtest (on the top), the Letter 1 represents a more frequent and the Letter 2 a less frequent word initial letter in a given language (frequency percentage in brackets). For the semantic subtest (at the bottom), Category 1 represents the category, which elicits a higher number of responses and Category 2 indicates the category that elicits a smaller number of responses.

Phonemic subtest

Language	German	English	Italian
Letter 1	M (5.5%)	W (5.7%)	D (4.9%)
Letter 2	T (4%)	L (3.9%)	F (4.5%)
Total frequency	9.5%	9.6%	9.4%
Score	30 (8.5)	22.2 (5.8)	17.2 (5.9)

Semantic subtest

Category 1	Proper names (persons)	Countries	Cities
Category 2	Supermarket	Animals	Clothes
Score	51.3 (13.2)	38.9 (8.8)	25.1 (7.8)
Total score	81.3 (20.7)	61.1 (12.1)	42.3 (11.1)

Appendix B: Materials (Italian, German and English):

List of the items used:

	Italian	*German*	*English*
1	Albero	Baum	Tree
2	Campana	Glocke	Bell
3	Cintura	Gürtel	Belt
4	Cipolla	Zwiebel	Onion
5	Collana	Kette	Necklace
6	Cucchiaio	Löffel	Spoon
7	Farfalla	Schmetterling	Butterfly
8	Foglia	Blatt	Leaf
9	Freccia	Pfeil	Arrow
10	Fungo	Pilz	Mushroom
11	Orologio	Uhr	Watch
12	Porta	Tür	Door
13	Ruota	Rad	Wheel
14	Scopa	Besen	Broom
15	Sedia	Stuhl	Chair
16	Uva	Weintraube	Grapes
17	Vestito	Kleid	Dress
18	Zucca	Kürbis	Pumpkin

Appendix C: Data Analysis

All the analyses were carried out in GNU-R (R Development Core Team, 2015) using the lme4 package (Bates, Maechler, Bolker & Walker, 2014). Reaction times data were fitted in linear mixed effect models, whereas accuracy binary data (correct= 1; incorrect= 0) were fitted in generalized linear mixed effect model with a logistic link function (Bates, 2010). All models included crossed random effects for participants and items. The best-fit models were selected by using a forward stepwise procedure in which non-significant predictors were removed from the models. Only significant predictors were used to compare nested models of increasing complexity, the parameters of which were estimated using the Maximum Likelihood for fixed effects and Restricted Maximum Likelihood for random effects comparisons (Pinheiro and Bates, 2000). Models that failed to converge were not included in the models' comparison. The fitted models were compared using likelihood-ratio chi-square tests (Baayen, 2008) and the quality of the fit was determined by means of the Akaike's Information Criterion (AIC). Following this procedure, the best-fit models included Language (L1 vs. L2 vs. L3), Condition (Switch vs. Repetition), Block Type (Single vs. Mixed) and Preceding Language (L1 vs. L2 vs. L3) as significant predictors. Fixed effects and, when justified, their interactions were coded using successive differences coding (which compares each level to the level before). More specifically, the three levels factors of Language and Preceding Language were coded as 2/3, -1/3, -1/3 and 1/3, 1/3, -2/3 contrasts, while the two levels factors of Condition and Block were coded as .5 vs. - .5. Finally, to approximate residuals normality and thus minimize distorting effects of outliers on the models' coefficients, models' criticism on the fitted data was performed by removing standardized residuals larger than 2.5 (see also Baayen, 2008).

Acknowledgments

I am indebted to Julia Festman and Harald Clahsen for their expert suggestions on the entire research project. I also would like to thank the members of the Potsdam Research Institute of Multilingualism (PRIM) for their valuable comments throughout the project progress. Finally, I am very grateful to all the participants that supported this research by taking part at the study.
The research presented in this article was supported by an Erasmus-Mundus Joint Doctoral scheme for 'International Doctorate for Experimental Approaches to Language and Brain' (IDEALAB) awarded to the author (MM) by the European Commission under grant no. <2012-0025/2013-1458-EMII EMJD>.

6 Publication IV

Under review with: *Frontiers in Psychology*

Bilingual Language Switching: Production vs. Recognition

Michela Mosca[a, b] and Kees de Bot[c]

[a] International Doctorate in Experimental Approaches to Language And the Brain (IDEALAB), University of Potsdam, Germany; University of Groningen, the Netherlands; University of Trento, Italy; Newcastle University, UK; and Macquarie University, Australia.

[b] Potsdam Research Institute for Multilingualism (PRIM), University of Potsdam, Germany

[c] University of Pannonia, Hungary

Abstract

This study aims at assessing how bilinguals select words in the appropriate language in production and recognition, while minimizing interference from the non-appropriate language. Two prominent models are considered which assume that when one language is in use, the other is suppressed.

The Inhibitory Control (IC) model suggests that, in both production and recognition, the amount of inhibition on the non-target language is greater for the stronger compared to the weaker language. In contrast, the Bilingual Interactive Activation (BIA) model proposes that, in language recognition, the amount of inhibition on the weaker language is greater than otherwise.

To investigate whether bilingual language production and recognition can be accounted for by a single model of bilingual processing, we tested a group of native speakers of Dutch (L1), advanced speakers of English (L2) in a bilingual recognition and production task. Specifically, language switching costs were measured while participants performed a lexical decision (recognition) and a picture naming (production) task involving language switching.

Results suggest that while in language recognition the amount of inhibition applied on the non-appropriate language increases along with its dominance as predicted by the IC model, in production the amount of inhibition applied on the non-relevant language is not related to language dominance, but rather it may be modulated by speakers' unconscious strategies to foster the weaker language. This difference indicates that bilingual language recognition and production might rely on different processing mechanisms and cannot be accounted within one of the existing models of bilingual language processing.

Keywords: Language switching; IC model; BIA model; bilingual production and recognition; language inhibition.

Introduction

When a speaker of more than one language (hereafter "bilingual") processes a language, words from the non-relevant language might be activated and interfere. This can happen while speaking, but also during writing, listening and reading. The ability to confine processing to the relevant language is called "language control" and is essential for successful communication. Despite the importance of this phenomenon, research on language control has predominantly concentrated on language production (e.g., Abutalebi & Green, 2007, 2008; Calabria et al., 2012; Costa & Santesteban, 2004; Filippi, Karaminis & Thomas, 2014; Finkbeiner, Almeida, Janssen & Caramazza, 2006; Goldrick, Runnqvist & Costa, 2014; Gollan & Ferreira, 2009; Jackson et al., 2001; La Heij, 2005; Linck, Schwieter & Sunderman, 2012; Meuter & Allport, 1999), while much less attention has been devoted to language recognition (e.g., Grainger & Beauvillain, 1987; Orfanidou & Sumner, 2005; Thomas & Allport, 2000; von Studnitz & Green, 1997; von Studnitz & Green, 2002; Wang, 2015). Moreover, language production and recognition have been often investigated separately, leaving unclear whether the two processes rely on the same or different mechanisms. The present paper focusses on bilingual language control in production and recognition.

Language control in production and recognition

To investigate language control, most of the studies have focussed their attention on spoken (but not written) word production and visual (but not spoken) word recognition. In spoken word production, language control refers to the capacity of a bilingual person to speak in the intended language, while avoiding interferences from the non-intended language. In visual word recognition, language control indicates the ability to understand the meaning of written words belonging to a certain language, while reducing interference from the non-target language. These processes are far from being effortless. Indeed, it is generally agreed that when a word from the intended language is processed, also words from the irrelevant language are coactivated and might interfere during processing (e.g., for production: Colomé, 2001; Hermans et al., 1998; Kroll, Bobb, & Wodniecka, 2006, Poulisse & Bongaert, 1994; for recognition: Dijkstra, Timmermans & Schriefers, 2000; Dijkstra, van Jaarsveld & ten Brinke, 1998; van Heuven, Dijkstra & Grainger, 1998). Much evidence has suggested that interference from the irrelevant language might be resolved by suppressing words of the non-target language (e.g., for production Green, 1998; Kroll, Bobb, Misra & Guo, 1998; but see La Heij, 2005; Finkbeiner et al., 2006; Costa, Santesteban & Ivanova, 2006; for recognition: Dijkstra & van Heuven, 1998, 2002; van Heuven et al., 1998). According to the Inhibitory Control (IC) model proposed by Green (1986, 1993, 1998), the amount of inhibition applied on the non-relevant language depends on its dominance. This means that a greater magnitude of inhibition is needed to suppress the stronger language (e.g., the native language or “L1”) compared to the weaker language (e.g., a later acquired language or “L2”). It follows that the cost to reactivate a language depends on the strength of inhibition previously applied, with more strongly inhibited languages having larger reactivation costs than less strongly inhibited languages. Hence, within the IC model reactivation costs (also “switching costs”) are predicted to be larger for the stronger than for the weaker language. In this framework, switching costs are *dominance-related*.

Even though the predictions of the IC model have been mainly used to investigate language control in production, the model was conceived to be applicable on both bilingual production and recognition (e.g., Green, 1986, 1998). Research on bilingual recognition, however, has mainly relied on a computational model called Bilingual Interactive Activation (BIA; Dijkstra & van Heuven, 1998; Grainger & Dijkstra, 1992; van Heuven, Grainger & Dijkstra, 1998). According to the BIA, when a word is presented, similar words from both the relevant and the irrelevant language are activated. The activated words send activation to the respective language node (representational layer containing language tags). Language competition is solved via inhibition from the language node to the words of the other language. The amount of inhibition applied to the words of the other language depends on the strength of activation of the language node. Specifically, the stronger the activation of the language node the greater the inhibition of the words of the other language, i.e. “asymmetric inhibition”. Because of this, words from the non-relevant language are more strongly inhibited than words from the relevant language. In this way, the words of the relevant language that best matches the input becomes most active and crosses the “recognition threshold” (Dijkstra & van Heuven, 1998). More generally, the BIA suggests that the activation of the language node reflects the amount of activity in the lexicon. Since L1 words have a higher baseline activation level than L2 words, the L1 language node is more activated than the L2 language node, and inhibition is greater on L2 than on L1 words (Dijkstra & van Heuven, 1998; van Heuven, Grainger & Dijkstra, 1998). Therefore, if inhibition is asymmetrical, namely greater for L2 words than for L1 words, then also switching costs for L2 and L1 are expected to be asymmetrical, that is larger for the weaker L2 than for the stronger L1. In this scenario, switching costs are *dominance-reversed*.[3] A similar interpretation of the BIA model was provided by Grainger and colleagues (2010), suggesting that since L1 words have a higher resting level of activation than L2 words, on a switch trial interference from the L1 into the L2 is greater than otherwise, leading to larger costs for the weaker than for the stronger language.

---

[3] In the extended version of the BIA model, the BIA+ model (Dijkstra & van Heuven, 2002), the asymmetrical inhibition from language node to words is removed. As replacement to the language nodes, the BIA+ model introduces a task/decision system that should perform the same functions. However, how the mechanism exactly works is not specified, leaving the BIA+ model an incomplete model of bilingual processing (see also Bultena & Dijkstra, 2012).

Language Switching

To measure switching costs and thus cast light on how bilinguals control their languages, a language switching paradigm is required. The language switching paradigm includes two types of trials, *repetition trials* (stimuli in the same language as in the preceding trial, e.g.,L1-L1) and *switch trial*s (stimuli in a different language compared to the preceding trial, e.g.,L2-L1). Responses on switch trials are usually less accurate and slower compared to repetition trials, and this difference is known as language "switching costs" (e.g., Meuter & Allport, 1999).

Several studies on bilingual language control in production have interpreted larger switching costs for the stronger than for the weaker language, i.e. *asymmetrical switching costs*, as evidence for dominance-related inhibition as predicted by the IC model (e.g., Meuter & Allport, 1999; Fink & Goldrick, 2015; Macizo, Bajo & Paolieri, 2012; Peeters, Runnqvist, Bertrand & Grainger, 2014; Reynolds, Schlöffel & Peressotti, 2016; but see Bobb & Wodniecka, 2013). However, it has been also shown that when the dominance difference between the L1 and the L2 is relatively small, the amount of switching costs for the two languages becomes comparable, i.e. *symmetrical switching costs* (e.g., Christoffels, Firk & Schiller, 2007; Declerck, Koch & Philipp, 2012; Fink & Goldrick, 2015; Schwieter & Sunderman, 2008). Based on this, the predictions of the IC model have been expanded by suggesting that when the dominance difference between two languages is relatively small, the amount of inhibition applied on the two languages is similar, yielding comparable switching costs (e.g., Meuter & Allport, 1999; but see Costa & Santesteban, 2004; Costa, Santesteban & Ivanova, 2006; Verhoef, Roelofs & Chwilla, 2009; for alternative explanations).

As far as recognition is concerned, few studies have investigated how bilinguals control languages (e.g., Thomas & Allport, 2000; von Studnitz & Green, 1997; for a review on bilingual word recognition see van Assche, Duyck & Hartsuiker, 2012). Interestingly enough, most of the studies reported a similar magnitude of switching costs for the stronger L1 and the weaker L2, that is symmetrical switching costs (e.g., for lexical decision task: Orfanidou & Sumner, 2005; Thomas & Allport, 2000; von Studnitz & Green, 1997; for categorization task: Macizo et al., 2012; von Studnitz & Green, 2002; but see Jackson et al., 2004). Unfortunately, symmetrical switching costs do not clearly indicate whether bilingual language control in recognition can be better accounted for by the IC or the BIA model. Indeed, it is the direction of switching costs asymmetry (whether switching costs are larger for the stronger or for the weaker language) that more clearly indicates whether language control in recognition can be better explained by the IC model (predicting larger switching costs for the stronger than for the weaker language) or the BIA model (predicting larger costs for the weaker than for the stronger language). Furthermore, these results have led to the question whether the preponderant presence of asymmetrical switching costs in production relative to recognition tasks is due to methodological inconsistencies across studies or to the fact that bilingual language control in production and recognition rely on two different mechanisms (Reynolds et al. 2016). The goal of the present study is to shed light on these issues. More precisely, the study aims at investigating whether 1) language control in recognition is dominance-related (as the IC model predicts) or dominance-reversed (as the BIA model predicts); 2) language control in bilingual recognition and production rely on the same mechanisms. To do it, we tested one group of unbalanced bilinguals (advanced L2 speakers, with a stronger L1 and a weaker L2) in both a recognition and a production task involving language switching. As to recognition, participants were administered a bilingual lexical decision task; regarding production, participants performed a bilingual picture naming task. In both tasks, switching costs in L1 and L2 were measured. Moreover, to assess languages' baseline strength of activation (believed to be a reflection of language dominance), a "single-language" block was included in the picture naming task, in which pictures had to be named in either L1 or L2 separately.

Concerning the predictions, we expect one of two possible outcomes: 1) If the amount of inhibition of the non-relevant language strongly depends on language dominance, we expect to find asymmetrical switching costs in both the lexical decision and the picture naming task. In this framework, if the magnitude of inhibition is dominance-related (as the IC model predicts), then switching costs are expected to be larger for switching to the stronger than for switching to the weaker language in both tasks. If, however, language inhibition in recognition is dominance-reversed (as predicted by the BIA model), switching costs in the lexical decision tasks are expected to be larger for the weaker than for the stronger language. 2) Alternatively, if asymmetric inhibition is applied only in cases of extreme dominance difference between two languages, then switching costs are expected to be symmetrical in both the recognition and the production task.

Participants

For the present study, we recruited thirty-two native speakers of Dutch (6 men, mean age: 21.8 years, sd: 4.49) from the student population of the University of Groningen (mean years of formal education: 16.7, sd: 1.93). Participants were tested in Dutch and English and were paid for their participation. All participants were right handed, had normal or corrected to normal vision and had never been diagnosed with reading, learning or language disability. Before the experiment, participants gave their written consent and filled out the Language Experience and Proficiency Questionnaire (LEAP-Q, Marian, Blumenfeld & Kaushanskaya, 2007) to assess their language profiles.

All participants had acquired Dutch from birth as the native language (L1) and English at school as a second language (L2) for a minimum of 6 years (mean L2 AoA: 9.35 years, sd: 2.34). On a daily basis, speakers were mostly exposed more to Dutch (57%) than to English (34%) or to other languages (9%). To screen for language skills, participants were asked to self-rate their ability to speak, understand and read in Dutch and English based on a ten-point scale, where 0= no knowledge and 10= perfect knowledge. Overall, results revealed that participants considered themselves excellent users of Dutch (mean score: 9.16, sd: 0.85) and good users of English (mean score: 7.65, sd: 1.04). Their L2 proficiency level was tested using the grammar part of the paper-based Oxford Placement Test (Allan, 2004). The test yielded a mean score of 84.4% (sd: 4.72) correct answers, indicating that participants were highly proficient L2 users (C1-C2 level), according to the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001). For detailed information on participants, see Appendix A.

Materials

Stimuli of the lexical decision task were 28 words representing simple concrete objects and 28 non-existing words, i.e. pseudowords. With regard to the words, half of the items were in Dutch and the other half were their English translations[4]. Based on the information of the webCELEX database (http://celex.mpi.nl/), words were matched for word form and lemma frequency (t= 1.49, p> 0.5 and t= 1.25, p> 0.5) as well as for letter orthographic length (t= 1.22, p> 0.05). Words were also matched according to their orthographic neighborhood density within each language (t= 1.54, p> 0.05), between the two languages (t= 0.75, p> 0.05) and across other languages (German, t= 0.47, p> 0.05; French, t= 1.18, p> 0.05 and Spanish, t= 1.32, p> 0.05). The values for the orthographic neighborhood density were taken from the Cross Linguistic Easy Access Resource for Phonological and Orthographic Neighborhood Densities database (CLEARPOND; Marian et al., 2012). Furthermore, we made sure that words were legal string of letters in both Dutch and English. To do it, we checked that all bigram transitions and single letter positions were probable and that probability was comparable in the two languages. Finally, we did not include cognates, homophones or words belonging to the same semantic category.

With concern to pseudowords, they were created by modifying the first or the last subsyllabic elements of the words selected for the lexical decision task. Apart from this change, pseudowords maintained the subsyllabic structure of the words they were generated from (e.g., from the words "glass-es" and "wo-lk" the pseudowords "<u>smart</u>-es" and "wo-<u>lm</u>" were derived). Half of the

---

[4] In the lexical decision task, we used words representing the same objects in L1 (Dutch) and L2 (English) in order to be consistent with the picture naming, where the same pictures were used to test performance in the two languages.

pseudowords were created starting from the selected Dutch words and the other half from the selected English words by using the multilingual pseudoword generator Wuggy (Keuleers & Brysbaert, 2010). To refrain participants from relying on subtle cues to differentiate between words and pseudowords, we matched words and pseudowords on several levels. Specifically, all pseudowords obeyed to the phonotactic constraints of both Dutch and English (i.e. they were legal string of letters in both languages) and they had an overlap ratio of 2/3 to existing words (i.e. they looked and sounded like real words). To make sure that the generated pseudowords equally represented words in Dutch and English, they were matched as closely as possible according to their bigram transitions and single letter positional probability in the two languages. Moreover, pseudowords were matched according to their orthographic Levenshtein distance to real words in the lexicon ($t= 0.31$, $p> 0.05$). The neighborhood size difference between pseudowords and words was kept as minimal as possible (neighborhood size difference< 0.45, see also Keuleers & Brysbaert, 2010) and this was matched across pseudowords ($t= 0.51$, $p< 0.05$). Finally, words and pseudowords had the same orthographic length.

For the picture naming task, we used 14 pictures corresponding to the words used in the lexical decision task. Pictures were taken from the "Colorized Snodgrass and Vanderwart pictures" set (Rossion & Pourtois, 2004) and had to be named either in Dutch or in English.

All items were presented in the centre of a 15-inch computer screen set to 1280x800 pixel resolution and they were seen from a distance of approximately 80cm. Words and pseudowords were presented in white lowercase letters (font: Courier New, point size: 36) against a black background. Pictures had a size of 197x281 pixel and were presented against a coloured (green or blue) background. Stimuli were presented using the software E-Prime Professional version 2.0 (Psychology Software Tools, Pittsburgh, PA). See Appendix B for detailed information on the stimuli used.

Procedure

Participants were tested individually in a quiet laboratory room. They were first given verbal instructions about the task, followed by written instructions displayed on the computer screen. The experimental session consisted of the lexical decision task followed by the picture naming task. In both tasks, participants were asked to respond as quickly and accurately as possible.

In the lexical decision task, E-Prime was used for data collection. Reaction times were measured as the interval between the display of the string of letters and the onset of the manual response. In the picture naming task, naming latencies were manually checked with the software Praat version 5.4.08 (Boersma & Weenink, 2015) and were measured as the interval from the presentation of the picture until the speech onset. In both tasks, a trial consisted of (i) a fixation cross for 250ms, (ii) a blank screen for 250ms, (iii) the target item (together with the language cue in the picture naming task) for 1.500ms and (iv) blank screen for 1.000ms. Independently from subjects' response speed, each trials had a fixed duration of 3.000ms.

In the lexical decision task, participants were instructed to decide whether a presented string of letters was a real word or not by pressing either a YES or a NO button. Participants responded by using the index finger of their right or left hand to press a button on the right or left side of the keyboard, respectively. The assignment of the button was counterbalanced across participants. The lexical decision task consisted of 336 trials: 1/3 of pseudowords (112 trials) and 2/3 of words (224 trials). Half of the words belonged to the L1 and the other half to the L2. Participants were told that words' language membership was irrelevant for the task.

Items were displayed singularly, but they were organized in pseudo-randomized chunks. There were two types of chunks, namely full and partial word type chunk. The full word type chunk included only words and were composed of 75% repetition and 25% switch trials (total of 252 and 84 trials respectively). For example, in the chunk L1-L1-L1-L2, the first three members belonged to the same language (L1) and the last one to the other language (L2). Each of the four elements had to be classified as existing words, irrespective of language membership. Only the last two elements of the chunk were included in the analysis (e.g., X-X-L1-L2). We used this system to make sure that every repetition trial was coming from a "pure" repetition trial (note the difference between L2-L1-L1-L2 and L1-L1-L1-L2) and to exclude effects of backwards inhibition on switch trials (note the difference between L1-L2-L1-L2 and L1-L1-L1-L2), for a review on backward inhibition see Koch et al., (2010). Each word was seen only once in each chunk position (i.e. first, second, third and fourth position) of the two language chunks (L1-L1-L1-L2 and L2-L2-L2-L1).

A partial word type chunk entailed both words and pseudowords: the first three components belonged to the same category (word or pseudoword), while the fourth element was in a different category. For example, in a word-word-word-pseudoword chunk, the first three elements had to be classified as existing words and the fourth one as a non-existing word. If the first three components were words, they always belonged to the same language (e.g., L1), while the following pseudoword was generated either from the same language (e.g., L1) or from a different one (e.g., L2). If the first three components of the chunk were pseudowords, they could have been generated from both the L1 and the L2 and the following word was either in the L1 or in the L2. Only the last two elements of the chunk were included in the analysis (e.g., X-X-word-pseudoword).

If two items (e.g., bottle-glasses-X-X) had occurred together in a specific chunk, this items' combination was not repeated twice; additionally, their translations (e.g., fles-bril-X-X) and their derived pseudowords (e.g., boddle-smartes-X-X) were never presented together within a chunk. To avoid orthographic priming, items sharing more similar orthographic patterns (such as the pseudowords derived from the item words) never occurred in the same chunk. Moreover, a given item was never seen within the next 5 trials and the same type of chunk never occurred more than twice in a row. Because of these constraints, order of the trials was unpredictable. Four lists of the task were created, two starting with a word-word-word-pseudoword and the other two with a pseudoword-pseudoword-pseudoword-word chunk type. Each participant was administered with one list only. The language of instructions for the lexical decision task was Dutch. Before the main experiment, participants were given a practice session of 24 trials. Practice items were not included in the main experiment.

In the picture naming task, participants had to name a presented picture either in their L1 (Dutch) or in their L2 (English). The language to be used was signalled by the background colour of the screen (e.g., blue= L1 and green= L2). The assignment of the colour cue to the response language was counterbalanced across participants. The picture naming task consisted of a single-language followed by a mixed- language block. In the single-language block, pictures had to be named in either L1 or L2 separately. It included a total of 56 trials: Half of the items had to named in the L1 and the other half in the L2. For each language, the 14 pictures were randomly presented for two times consecutively. The order of languages' presentation was counterbalanced across participants. The language of instruction corresponded to the language in which the upcoming task had to be performed (English for the upcoming L2 part and Dutch for the following L1 part).

The mixed-language block involved two kinds of trials, repetition and switch trials. In a repetition trial, a given picture had to be named in the same language as in the trial before (e.g., L1-L1). In a switch trial, a presented picture had to be named in a different language compared to the trial before (e.g., L2-L1). The mixed-language block was composed by 112 trials: Half of them had to be named in the L1 and the other half in the L2. Trials were organized in pseudo-randomized language chunks composed of 75% repetition and 25% switch trials (84 and 28 trials respectively). For example, the language chunk L1-L1-L1-L2 implied that the three consecutive pictures had to be named in the L1 and the fourth one in the L2. For the analysis, only the second part of the language chunk was used (X-X-L1-L2). The same chunk type was never displayed more than two times in a row. Each picture was seen eight times within the mixed-language block, which is once in each position of the language chunk (first, second, third of fourth position) of the two language chunks (L1-L1-L1-L2 and L2-L2-L2-L1). The same picture was not seen within 5 trials. Again the order of the items was unpredictable. Four lists of mixed-language block were created: two starting with a L1-L1-L1-L2 and the other two with a L2-L2-L2-L1 chunk type. Each participant saw only one list. In the mixed language block, the language of instructions was Dutch.

Participants were given one practice sessions before each language part of the single-language block (total of 12 trials) and one practice session before the mixed-language block, in which the two chunk types were trained (total of 16 trials). Practice trials did not appear in the experimental sessions. On average, the experiment (including instructions and breaks) lasted 30 minutes.

Data cleaning

The dependent variables of the present study were accuracy rates and reaction times. Before the statistical analyses, we cleaned the dataset based on participants' and items' error rates. Because of low accuracy scores (< 66%), one participant and two items were removed from subsequent analysis (one participant from the single-language block of the picture naming task and the pseudowords "hers" and "croud" of the lexical decision task). Accuracy rates for all other participants and items ranged respectively from 73% to 100% and from 79% to 100%.

Both correct and incorrect responses were included in the accuracy rates' analyses, whereas only correct responses were used to analyse reaction times. We deemed a response as incorrect in case of wrong button pressing for the lexical decision task (4.5% of the data) and in case of microphone miss-triggering (e.g., coughing, hesitation, utterance repairs, etc.) and selection of the wrong word and/or language for the picture naming task (9.8% of the data). In both tasks, missing responses were classified as incorrect.

After exclusion of the incorrect responses, we screened reaction times for extreme values. Extremely fast (< 150ms) or particularly slow (> 2.500ms) reaction times were removed from the dataset (0.03% of the data). Visual inspection revealed that the removal of the extreme values did not result in normality, violating the normality assumption underlying the general linear model (Baayen & Milin, 2010). Tests for skewness indicated that the data were positively skewed (i.e. data skewness > 0.80). To decide on a suitable transformation, we estimated the lambda parameter of the Box-Cox transformation, which ensures that residuals RTs are approximately normal (Box & Cox, 1964). Lambda values of -0.50, -0.14 and -0.30 (for lexical decision, and single- and mixed-language picture naming) indicated that a reciprocal cube root and a log transformation was appropriate for the lexical decision and the picture naming task, respectively. Visual inspections of the data together with the skewness tests confirmed the reciprocal square root and the log transformation as appropriate to approximate normality in the two tasks (data skewness < 0.15).

All the analyses were carried out in GNU-R version 3.2.2 (R Core Team, 2015) using the lme4 package version 1.1-9 (Bates et al., 2015). Reaction times data were fitted in linear mixed effects models, while accuracy binary data (1= correct, 0= incorrect) were fitted in generalized linear mixed effects models with a logistic link function. All models included crossed random effects for participants and items. The best-fit models were selected using a forward stepwise procedure, where only significant predictors were used to compare nested models of increasing complexity. During models comparisons, parameters were estimated using the Maximum Likelihood for fixed effects and Restricted Maximum Likelihood for random effects comparisons (Pinheiro & Bates, 2000). Models that failed to converge were excluded from comparisons of further models. The quality of the fit was determined by the Akaike's information criterion (AIC; Akaike, 1998) and its significance by the likelihood ratio test (see also Matuschek et al., 2015).

Based on this criteria, the best-fit model for the lexical decision data included the experimental factors of Language (L1 vs. L2), Condition (Repetition vs. Switch), Word Type (Word vs. Pseudoword) and Word Type Change (Yes vs. No) as significant predictors. The model had fully random slopes and intercepts by participants and items structure (the so-called maximal model, Barr et al., 2013; Bates et al., 2015). Main effects and interactions were coded using sum contrasts (i.e. -0.5 vs. +0.5). For the picture naming task, the best-fit model included the experimental factor of Language (L1 vs. L2) and Block (Single vs. Mixed). In addition to this, the factor Condition (Repetition vs. Switch) was included in the mixed-language block. The models had random intercepts by participants (for the factor Language) in the single-language block and by participants and items (for the factor Condition) in the mixed-language block. Before accepting the models, we checked whether the they provided a satisfactory fit to the data, i.e. “models’ criticism”. Visual inspection revealed that the distribution of the residuals tended to have a thicker right tail than expected for a normal distribution and that variance was not uniformly distributed (see also Baayen, 2008). To approximate residuals normality and to stabilize variance (“homoscedasticity”), we removed standardized residuals larger than 2.5 (see also Baayen, 2008). Mild a-priori screening for outliers together with model criticism are considered two essential procedures within mixed-modelling approach to avoid model distortion (Baayen & Milin, 2010).

Results

To investigate the assumptions outlined in the introduction, we measured L1 and L2 switching costs (calculated as the RT difference between repetition and switch trials) in the lexical decision and in the picture naming tasks. Consider first the lexical decision task. Mean accuracy rates and reaction times are illustrated in Table 1. Results of the statistical analysis for the accuracy rate and the reaction time data are reported in Table 2 and Table 3, respectively.

Table 1. Correct mean reaction times in milliseconds (standard deviation in brackets) and accuracy rates in percent for Words and Pseudowords as a function of Category Change (no vs. yes), Language (L1 vs. L2) and Condition (Repetition vs. Switch). Language membership in Pseudowords indicates whether they were generated from a L1 or a L2 word. Switching costs (calculated as the difference between repetition and switch trials) are reported in italics.

		Category Change (no)			Category Change (yes)		
		L1	L2	Mean	L1	L2	Mean
a) Word							
Repetition	*Reaction times*	468ms (127)	511ms (170)	490ms (149)	585ms (174)	583ms (150)	584ms (162)
	Accuracy rates	99% (8)	98% (13)	99%	97% (18)	93% (25)	95%
Switch	*Reaction times*	503ms (144)	508ms (142)	505ms (143)	562ms (129)	584ms (120)	572ms (125)
	Accuracy rates	98% (12)	96% (19)	97%	97% (17)	93% (26)	95%
SWITCHING COSTS		*35ms*	*-3ms*		*-23ms*	*1ms*	
Language Mean	*Reaction times*	480ms (134)	510ms (161)	495ms (149)	573ms (152)	583ms (135)	578ms (144)
	Accuracy rates	99% (10)	97% (15)	98% (13)	97% (17)	93% (26)	95% (22)
b) Pseudoword							
Repetition	*Reaction times*	625ms (178)	630ms (168)	627ms (173)	617ms (122)	619ms (153)	618ms (138)
	Accuracy rates	96% (18)	99% (10)	98%	82% (38)	88% (33)	85%
Switch	*Reaction times*	637ms (189)	630ms (212)	634ms (201)	589ms (92)	604ms (93)	596ms (92)
	Accuracy rates	94% (25)	94% (25)	94%	93% (25)	93% (25)	93%
SWITCHING COSTS		*12ms*	*0ms*		*-28ms*	*-15ms*	
Language Mean	*Reaction times*	631ms (184)	630ms (191)	630ms (188)	612ms (117)	618ms (148)	615ms (133)
	Accuracy rates	95% (22)	96% (19)	96% (21)	84% (36)	88% (32)	86% (34)

Lexical decision task

The analysis of the accuracy rates revealed that responses were significantly more accurate in words than in pseudowords (β= 0.44, SE= 0.09, z= 4.64, $p < 0.0001$). Category change (e.g., from pseudoword to word) yielded significantly less accurate responses compared to when the same category was repeated (β= 0.60, SE= 0.09, z= 6.26, $p < 0.0001$). The significant interaction of Language by Category (β= 0.33, SE= 0.09, z= 3.43, $p < 0.01$), indicated that compared to pseudowords, words were responded more accurately in the L1 than in the L2. Specifically, while accuracy rates did not differ significantly for pseudowords generated from the L1 and those generated from the L2 ($p > 0.05$), L1 words were responded more accurately than L2 words (β= 0.44, SE= 0.11, z= 3.93, $p < 0.001$). Finally, we found a significant interaction between Condition and Category Change (β= 0.37, SE= 0.09, z= 3.80, $p < 0.001$), indexing that a change in category affected repetition and switch trials differently. In particular, compared to a situation in which the same category is repeated (e.g., word-word), a change in category (e.g., pseudoword-word) yielded significantly lower accuracy rates for repetition trials (β= 2.40, SE= 0.21, t= 11.19, $p < 0.0001$), but not so for switch trials ($p > 0.05$).

As expected, the analysis of the reaction times showed that pseudowords were responded significantly slower than words (β= 1.57, SE= 0.17, t= 9.24, $p < 0.0001$) and that category change yielded slower responses compared to category repetition (β= 0.88, SE= 0.15, t= 5.56, $p < 0.0001$). The main effects of Language and Condition were not significant ($p > 0.05$). However, the significant interactions of Language by Category (β= 0.17, SE= 0.07, t= 2.31, $p < 0.05$) and of Condition by Category (β= 0.18, SE= 0.08, t= 2.21, $p < 0.05$) revealed that compared to pseudowords, words were responded faster in the L1 than in the L2 and in repetition compared to switch trials. Both the interactions of Category by Category Change and of Condition by Category Change were also significant (β= 0.21, Se= 0.91, t0 2.32, $p < 0.05$ and β= 0.93, SE= 0.07, t= 12.32, $p < 0.0001$), indicating that a category change was more costly for repetition compared to switch trials and for words than for pseudowords. Both the three-way interactions between Language, Condition and Category and between Condition, Category and Category Change were significant (β= 0.23, SE= 0.08, t= 2.65, $p < 0.05$ and β= 0.20, SE= 0.08, t= 2.32, $p < 0.05$, respectively). All the other interactions were not significant ($p > 0.05$).

To investigate these interactions, we split the data into words and pseudowords. With regard to words, we found that responses were faster in the L1 than in the L2 (β= 0.84, SE= 0.16, t= 5.17. $p < 0.0001$) and in repetition compared to switch trials (β= 0.38, SE= 0.17, t= 2.23, $p < 0.05$). The effect of Category Change was also significant (β= 3.62, SE= 0.28, t= 12.60, $p < 0.0001$), indicating that responses were significantly slower if the preceding trial was a pseudoword compared to a word. This effect was smaller for switch compared to repetition trials (β= 1.21, SE= 0.34, t= 3.55, p< 0.01). The three-way interaction between Language, Condition and Category Change was also significant (β= 1.44, SE= 0.67, t= 2.14, p< 0.05). No other interaction was significant (p> 0.05). Words were further divided based on the category of the preceding trial (Category Change: no vs. yes).

When the same category was repeated (i.e. Category Change: no), words were responded faster in the L1 relative to L2 (β= 0.99, SE= 22, t= 4.49, p< 0.0001) and in repetition compared to switch trials (β= 0.91, SE= 0.17, t= 5.18, p< 0.0001). The difference between repetition and switch trials was larger for L1 compared to L2 words (β= 1.40, SE= 0.35, t= 3.97, p< 0.001), that is *asymmetrical switching costs*. In particular, the effect of language switching was significant for the L1 (β= 1.68, SE= 0.42, t= 4.00, p< 0.001), but not for the L2 (p> 0.05). With reference to words preceded by pseudowords (i.e. Category Change: yes), responses were faster in the L1 compared to the L2 (β= 0.70, SE= 022, t= 3.14, p< 0.01). No other main effect or interaction was significant (p> 0.05) in this condition. Concerning pseudowords, items generated from the L1 were responded equally fast as those generated from the L2 (p< 0.05). All the other main effects and interactions were also not significant (p> 0.05).

To sum up, performances were more accurate and faster for words than for pseudowords and in case of category repetition (e.g., word-word) compared to category change (e.g., pseudoword-word). Changing category was particularly costly for words and repetition trials than for pseudowords and switch trials respectively. The effect of Language was influential for words, in that L1 words were responded faster and more accurately than L2 words, but not for pseudowords. The effect of Condition also affected words only, with responses on switch trials being slower and less accurate than on repetition trials. This difference was significant for the L1 and not significant for the L2, i.e. asymmetrical switching costs.

Table 2. Estimated coefficients standard errors (SE) and z values from the best-fit generalized linear mixed-effects models for the accuracy data. Asterisks (∗) indicate: $p< 0.05$ (∗), $p< 0.01$ (∗∗), $p< 0.001$(∗∗∗) and $p< 0.0001$ (∗∗∗∗).

	Accuracy rates		
	Estimate	SE	z-value
a) Overall model			
Intercept	3.56	0.21	16.79****
Language (L1 vs. L2)	0.12	0.09	1.30
Condition (Repetition vs. Switch)	0.15	0.09	1.56
Category (Word vs. Pseudoword)	0.44	0.09	4.64****
Category Change (no vs. yes)	0.60	0.09	6.26****
Language*Condition	- 0.13	0.09	- 1.35
Language*Category	0.33	0.09	3.43**
Condition*Category	0.05	0.09	0.56
Language*Category Change	- 0.06	0.09	- 0.66
Condition*Category Change	0.37	0.09	3.80***
Category*Category Change	- 0.01	0.09	- 0.18
Language*Condition*Category	0.10	0.09	1.12
Language*Condition*Category Change	- 0.02	0.09	- 0.27
Language*Category*Category Change	0.05	0.09	0.60
Condition*Category*Category Change	- 0.16	0.09	- 1.73
Language*Condition*Category*Category Change	0.05	0.09	0.58
*Formula: Accuracy ~ Language*Condition*Category*Category Change + (1+ | subject) + (1+ | item)*			
b) Word			
Intercept	4.133	0.247	16.71****
Language (L1 vs. L2)	0.440	0.112	3.93***
Formula: Accuracy ~ Language + (1+ | subject) + (1+ | item)			
c) Pseudoword			
Intercept	2.763	0.226	12.17****
Language (L1 vs. L2)	- 0.172	0.091	- 1.89
Formula: Accuracy ~ Language + (1+ | subject) + (1+ | item)			
d) Repetition trials			
Intercept	4.814	0.282	17.05****
Category Change (no vs. yes)	2.401	0.214	- 11.19****
Formula: Accuracy ~ Category Change + (1+ | subject) + (1+ | item)			

e) Switch trials			
Intercept	3.629	0.262	13.81****
Category Change (no vs. yes)	- 0.436	0.243	- 1.79
Formula: Accuracy ~ Category Change + (1+ | subject) + (1+ | item)			

Table 3. Estimated coefficients, standard errors (SE) and t values from the best-fit linear mixed effects models run on reciprocal square root-transformed RTs. Asterisks (∗) indicate: $p< 0.05$ (∗), $p< 0.01$ (∗∗), $p< 0.001$(∗∗∗) and $p< 0.0001$ (∗∗∗∗).

	Reaction times		
	Estimate	SE	t-value
a) Overall model			
Intercept	- 42.76	0.44	- 96.61****
Language (L1 vs. L2)	- 0.21	0.17	- 1.22
Condition (Repetition vs. Switch)	0.08	0.08	0.95
Category (Word vs. Pseudoword)	- 1.57	0.17	- 9.24****
Category Change (no vs. yes)	- 0.88	0.15	- 5.56****
Language*Condition	- 0.14	0.08	- 1.80
Language*Category	- 0.17	0.07	- 2.31*
Condition*Category	- 0.18	0.08	- 2.21*
Language*Category Change	- 0.05	0.07	- 0.67
Condition*Category Change	- 0.21	0.09	- 2.32*
Category*Category Change	- 0.93	0.07	- 12.32****
Language*Condition*Category	0.07	0.08	- 0.92
Language*Condition*Category Change	- 0.23	0.08	- 2.65*
Language*Category*Category Change	- 0.06	0.07	0.92
Condition*Category*Category Change	- 0.20	0.08	- 2.32*
Language*Condition*Category*Category Change	- 0.02	0.08	- 0.24
*Formula: RT ~ Language*Condition*Category*Category Change + (1+ Language+Category+Category Change| subject) + (1+Language+Category+Category Change| item)*			
b) Word – Overall model			
Intercept	- 44.30	0.45	- 96.71****
Language (L1 vs. L2)	0.84	0.16	5.17****
Condition (Repetition vs. Switch)	0.38	0.17	2.23*
Category Change (no vs. yes)	3.62	0.28	12.60****
Language*Condition	- 0.49	0.33	- 1.48
Language*Category Change	- 0.44	0.33	- 1.36
Condition*Category Change	- 1.21	0.34	- 3.55**
Language*Condition*Category Change	1.44	0.67	- 2.14*

*Formula: RT ~ Language*Condition* Category Change + (1+ Category Change| subject) + (1 | item)*

c) Word – *Category Change (no)*

Intercept	- 46.13	0.50	- 91.42****
Language (L1 vs. L2)	0.99	0.22	4.49****
Condition (Repetition vs. Switch)	0.91	0.17	5.18****
Language*Condition	- 1.40	0.35	- 3.97***

*Formula: RT ~ Language*Condition + (1+ Language| subject) + (1+ Language*Condition | item)*

d) Word – *Category Change (no) L1*

Intercept	- 47.43	0.53	- 88.27****
Condition (Repetition vs. Switch)	1.65	0.41	3.95***

Formula: RT ~ Condition + (1 | subject) + (1+ Condition | item)

e) Word – *Category Change (no) L2*

Intercept	- 45.70	0.63	- 72.10****
Condition (Repetition vs. Switch)	0.10	0.49	0.21

Formula: RT ~ Condition + (1 | subject) + (1+ Condition | item)

f) Word – *Category Change (yes)*

Intercept	- 42.61	0.47	- 89.80****
Language (L1 vs. L2)	0.70	0.22	3.14**
Condition (Repetition vs. Switch)	- 0.31	0.33	- 0.95
Language*Condition	- 0.08	0.55	- 0.15

*Formula: RT ~ Language*Condition + (1+ Condition| subject) + (1 | item)*

g) Pseudoword – Overall model

Intercept	- 41.09	0.49	- 83.21****
Language (L1 vs. L2)	0.15	0.29	- 0.53
Condition (Repetition vs. Switch)	0.20	0.15	1.26
Category Change (no vs. yes)	0.05	0.15	0.36
Language*Condition	- 0.16	0.14	- 1.10
Language*Category Change	0.02	0.20	0.13
Condition*Category Change	0.17	0.17	0.98
Language*Condition*Category Change	- 0.30	0.17	- 1.76

*Formula: RT ~ Language*Condition*Category Change + (1+ Language*Category Change| subject) + (1+ Language*Category Change | item)*

Picture naming task

Consider now the picture naming task. Mean accuracy rates and reaction times are presented in Table 4. Results of the statistical analysis for the accuracy rate and the reaction time data are reported in Table 5 and Table 6, respectively.

Table 4. Mean reaction times in milliseconds (standard deviations in brackets) and accuracy rates in percent for correct responses of L1 vs. L2 in single-language block (upper part) vs. mixed-language block (lower part). For the mixed-language block repetition vs. switch trials are reported. Mixing costs (calculated as the difference between trials in the single-language block and repetition trials in the mixed-language block) and switching costs (calculated as the difference between repetition and switch trials) and for the L1 and the L2 are reported in italics.

		L1	**L2**	**Mean**
Single-language block				
Mean	*Reaction times*	704ms (175)	729ms (173)	716ms (174)
	Accuracy rates	92% (27)	90% (30)	91% (28)
Mixed-language block				
Repetition	*Reaction times*	765ms (181)	740ms (161)	752ms (171)
	Accuracy rates	90% (30)	94% (23)	92% (27)
MIXING COSTS		*61ms*	*11ms*	
Switch	*Reaction times*	809ms (177)	794ms (176)	801ms (177)
	Accuracy rates	88% (33)	88% (32)	88% (32)
SWITCHING COSTS		*44ms*	*54ms*	
Mean	*Reaction times*	787ms (179)	766ms (171)	776ms (175)
	Accuracy rates	89% (31)	91% (28)	90% (30)

The analysis of the accuracy rates showed that in the single-language block responses in L1 and L2 were equally accurate ($p > 0.05$). In the mixed-language block, responses were significantly more accurate in the L2 than in the L1 ($\beta = 0.38$, $SE = 0.17$, $t = 2.20$, $p < 0.05$) and in repetition compared to switch trials ($\beta = 0.57$, $SE = 0.16$, $t = 3.42$, $p < 0.01$). The accuracy rates' difference between the single- and the mixed-language blocks was not significant ($p > 0.05$). However, there was a significant interaction between Language and Block (β= 0.24, SE= 0.08, t= 3.08, p< 0.01), indicating that compared to the single-language block, mixed-language block responses were more accurate for the L2 than for the L1. Specifically, while the L2 was responded to significantly better in the mixed- compared to the single-language block (β= 0.39, SE= 0.12, t= 3.16, p< 0.01), L1 accuracy rates was not influenced by the type of block (p> 0.05).

The analysis of the reaction times revealed that there was no significant difference between L1 and L2 responses in the single-language block (p> 0.05) and that in the mixed-language block, responses were faster in the L2 than in the L1 (β= 0.02, SE= 0.008, t= 3.54, p< 0.01). Overall, participants responded slower on repetition trials of the mixed-language block compared to trials in the single-language block (β= 0.05, SE= 0.02, t= 2.70, p< 0.05), i.e. *mixing costs*. This difference was greater for the L1 than for the L2, as indicated by the significant interaction of Language and Block (β= 0.07, SE= 0.02, t= 2.85, p< 0.05). In particular, while L1 responses were significantly slower in the mixed- compared to the single-language block (β= 0.04, SE= 0.01, t= 4.34, p< 0.0001), L2 responses were not affected by the type of block (p> 0.05). In the mixed language block, switch trials were responded to slower than repetition trials (β= 0.06, SE= 0.01, t= 3.49, p< 0.01), and this effect was comparable for the L1 and the L2 (p> 0.05), which means that we found symmetrical *switching costs*.

Briefly, both accuracy rates and reaction time analyses showed that in the single-language block there was no difference between L1 and L2. In the mixed-language block, responses were more accurate and faster in the L2 compared to the L1 and in repetition than in switch trials. Responses in the L1 were equally accurate in repetition trials of the mixed-language block compared to trials of the single-language block and were responded to slower in the former than in the latter, which means that there were mixing costs. Responses in the L2 were more accurate in repetition trials of the mixed-language block than in trials of the single-language block and did not suffer from mixing costs. Finally, within the mixed-language block, the difference between repetition and switch trials was similar for the L1 and the L2, indexing symmetrical switching costs.

Table 5. Estimated coefficients standard errors (SE) and z values from the best-fit generalized linear mixed-effects models for the accuracy data. Asterisks (*) indicate: p< .05 (*), p< .01 (**), p< .001(***) and p< .0001 (****).

	Reaction times		
	Estimate	SE	z-value
a) Single-language block			
Intercept	2.66	0.19	13.34****
Language (L1 vs. L2)	0.07	0.15	0.46
Formula: Accuracy ~ Language + (1+ Language | subject) + (1+ Language | item)			
b) Mixed-language block			
Intercept	3.00	0.25	11.82****
Language (L1 vs. L2)	- 0.38	0.17	- 2.20*
Condition (Repetition vs. Switch)	- 0.57	0.16	- 3.42**
Language*Condition	0.29	0.16	1.72
*Formula: Accuracy ~ Language*Condition + (1+ Language | subject) + (1+ Language | item)*			
c) Mixing costs – Overall model			
Intercept	2.755	0.191	14.36****
Language (L1 vs. L2)	- 0.186	0.130	- 1.42
Block (Single vs. Mixed)	- 0.138	0.081	- 1.71
Language*Block	0.249	0.080	3.08**
*Formula: Accuracy ~ Language*Block + (1+ Language | subject) + (1 | item)*			
d) Mixing costs – L1			
Intercept	2.574	0.202	12.71****
Block (Single vs. Mixed)	6.805	0.102	1.13
Formula: Accuracy ~ Block + (1 | subject) + (1 | item)			
e) Mixing costs – L2			
Intercept	2.970	0.272	10.89****
Block (Single vs. Mixed)	-11.91	45.542	- 3.16
Formula: RT ~ Block + (1 | subject) + (1 | item)			

Table 6. Estimated coefficients, standard errors (SE) and t values from the best-fit linear mixed effects models run on log-transformed RTs. Asterisks (*) indicate: p< .05 (*), p< .01 (**), p< .001(***) and p< .0001 (****).

	Reaction times		
	Estimate	SE	t-value
a) Single-language block			
Intercept	6.54	0.02	271. 77****
Language (L1 vs. L2)	- 0.01	0.01	- 0.96
Formula: RT ~ Language + (1+ Language | subject) + (1 | item)			
b) Mixed-language block			
Intercept	6.631	0.024	274.24****
Language (L1 vs. L2)	- 0.029	0.008	- 3.54**
Condition (Repetition vs. Switch)	- 0.061	0.017	- 3.49**
Language*Condition	0.007	0.167	0.43
*Formula: RT ~ Language*Condition + (1+ Condition | subject) + (1+ Condition | item)*			
c) Mixing costs – Overall model			
Intercept	6.531	0.031	209.62****
Language (L1 vs. L2)	- 0.033	0.033	- 0.98
Block (Single vs. Mixed)	- 0.092	0.026	- 3.53**
Language*Block	- 0.071	0.024	- 2.85**
*Formula: RT ~ Language*Block + (1+ Language*Block | subject) + (1+ Language*Block | item)*			
d) Mixing costs – L1			
Intercept	6.576	0.028	231.53****
Block (Single vs. Mixed)	- 0.048	0.011	- 4.34***
Formula: RT ~ Block + (1 | subject) + (1+ Block | item)			
e) Mixing costs – L2			
Intercept	6.576	0.025	254.95****
Block (Single vs. Mixed)	- 0.010	0.009	- 1.09
Formula: RT ~ Block + (1+ Block | subject) + (1 | item)			

Discussion

The goal of the study was to investigate language control in bilingual recognition and production. More precisely, we aimed at assessing whether 1) language inhibition in recognition is dominance-related (as the IC predicts) or dominance-reversed (as the BIA predicts); 2) language control in bilingual recognition and production rely on the same or different mechanisms. To address these issues, we measured language switching costs in a group of native speakers of Dutch (L1) - proficient learners of English (L2) performing a bilingual lexical decision and a bilingual picture naming task.

Lexical decision task

In the lexical decision task, participants responded faster and more accurately to words than to pseudowords replicating the well established facilitatory effect of words over pseudowords (e.g., Meyer & Schvaneveldt, 1971; Weekes, 1997; MacGregor et al. 2015). Even though pseudowords were generated starting from existing L1 and L2 words, the effect language did not influence participants' performances. This indicates that overall pseudowords were properly matched, in that it was no longer possible to quickly associate a given pseudoword to a specific language. However, a closer inspection to the data revealed that there was a tendency for pseudowords generated from the L1 to be more error prone than the ones generated from the L2. As noted in previous studies, responses on pseudowords with many neighbors are hampered to a greater extent compared to pseudowords with fewer neighbors (Balota et al. 2004, Coltheart et al., 1977). To avoid this effect, pseudowords of the present study were matched according to their orthographic Levenshtein distance to real words in the lexicon. However, it might be possible that if a speaker is more dominant in one language (L1) than in another (L2), pseudowords generated from a more dominant language will coactivate a larger number of neighbors, than pseudowords generated from a less dominant language. This imbalance in the speakers' languages dominance could explain why pseudowords created from the L1 tended to be less accurate than pseudowords generated from the L2.

A change in the category type was costly only for words, but not for pseudoword. This means that words were responded slower and less accurately if preceded by pseudowords and that pseudowords performances were not affected by the category type of the preceding trial. This effect might be explained by the fact that in a lexical decision task, participants are faced with two tasks of unequal difficulty, as recognizing pseudowords is more costly than recognizing real words. Specifically, it might be the case that the effort of recognizing pseudowords has a carryover-effect on the next trial. This carry-over effect on upcoming trials is detectable in case of relatively fast trials, such as word recognizing, while it takes more time to dissipate in case of relatively slower trials, such as pseudowords recognition. It should be noted that in previous monolingual lexical decision tasks, both words and pseudowords were found to be responded to slower if the preceding trial was a pseudoword compared to a word (Lima & Huntsman, 1997; Perea & Carreiras, 2003). However, the difference between previous monolingual studies and the present study might be explained by the fact that pseudowords created from a single language are faster to recognize than pseudowords sharing orthographic rules with two languages. Therefore, potential carry-over effect from the preceding trial is still detectable when pseudowords are relatively easier to recognize (monolingual task), but less so when pseudowords are more difficult to identify (bilingual task). In this respect, it is interesting to note that, in the present study, there is a numerical trend for pseudowords preceded by pseudowords to be slower than pseudowords preceded by words.

Language switching costs. As expected, the difference between repetition and switch trials was significant for words, but not for pseudowords. As to words, we found that language switching was costly only when words were preceded by words, but not when words were preceded by pseudowords. The absence of language switching costs in case of category change (pseudowords-words) replicates previous studies on bilingual recognition (Thomas & Allport, 2000; von Studnitz & Green, 2002). This effect has been explained by the fact that a language change might unconsciously boost a change in response and, therefore, lead to faster responses when it overlaps with category change compared to when the same category is repeated (von Studnitz & Green, 2002). However, if this explanation is correct, then it implies that pseudowords' language membership is still detectable and that it affects speakers' responses on upcoming trials. Alternatively, it might be supposed that in case of category change, language switching is not costly because pseudowords do not belong to any language and, therefore, cannot lead to language switching costs on upcoming words.

With regard to words preceded by words, language switching costs were larger for the L1 compared to the L2, i.e. asymmetrical switching costs. More precisely, responses were significantly slower on switch than on repetition trials for the L1, but not for the L2 where reaction times in switch and repetition trials were comparable. This result replicates Jackson et al.'s (2004) study, where unbalanced bilinguals (L1 English – different L2s) showed language switching cost only in the L1 but not in the L2 while performing a parity judgment task (i.e. classifying a digit as odd or even).

Why did we find language switching costly for the L1, but not for the L2? Before answering this question, it is important to understand the role of language strength of activation in word recognition.

As proposed by the BIA models, the speed with which a word is recognized depends on its baseline activation level, with more frequent word (such as L1 words) being recognized faster than less frequent words (such as L2 words). Additionally, the speed with which a word is recognized also depends on its relation to other words in the lexicon (e.g., van Heuven & Dijkstra, 1998). That being said, we assume that when a word from a weaker L2 is presented (e.g., the L2 word "BOTTON"), competing words from both the L2 and L1 will be activated. The activated words will send activation to the corresponding language node. In our example, the L2 word "BOTTON" will excite the L2 language node more than the L1 language node. The activated language nodes will inhibit words belonging to the competing language, so that L1 words will be strongly inhibited by the L2 language node. Therefore, when on a switch trials, an L1 word is presented, this needs to overcome the previously applied inhibition before being recognized, i.e. L1 switching costs. However, when a word from the stronger L1 is presented (e.g., the Dutch word "KNOP"), mostly words from the L1 will activate and act as competitors, but not so much words from the L2. This is because L1 words have a higher baseline activation level than L2 words, and therefore, when a L1 word is presented, L1 candidates are activated faster and more strongly than L2 ones. Therefore, if when L1 words are presented, competition from the L2 words is relatively weak, then the inhibition applied on the non-relevant L2 will also be relatively weak. Consequently, when on a switch trial the L2 becomes relevant again, the cost to reactivate this language will be small or absent, i.e. undetectable L2 switching costs.

This assumption has two main implications. Firstly, it assumes that the amount of inhibition exerted on words of the non-relevant language increases along with their activation level (more strongly activated words need to be inhibited more). This assumption is line with the IC model predictions, according to which the magnitude of inhibition depends on languages' activation strength. Yet, this is in contrast with the BIA model proposal that the L1 language node inhibits L2 words to a greater extent than vice versa (e.g., Dijkstra & van Heuven, 1998; van Heuven, et al., 1998), and with the idea that, because of their higher resting level, L1 words interfere more when switching into the L2 than the other way around, yielding larger switching costs for the L2 than for the L1 (Grainger et al., 2010). Secondly, it suggests that not only similar but also dissimilar words are activated and compete for selection. Specifically, despite the fact that no cognate, homograph or neighbor words were included in the present experiment, words from the non-relevant language were activated as indicated by the language switching costs measured in L1. This assumption is in line with the hypothesis that inhibition is exerted on the whole lexicon ("global control") and not on a restricted set of items ("local control"; Branzi et al. 2014). Finally, asymmetrical switching costs in advanced L2 speakers challenge the hypothesis suggesting that when the dominance difference between a stronger and a weaker language is relatively small, the amount of inhibition applied on the two languages is comparable, leading to symmetrical switching costs.

As described earlier, we found no significant switching costs for the L2. This result is in contrast with previous findings on bilingual lexical decision studies, in which the same amount of switching costs was found in L1 and L2 (e.g., Orfanidou & Summer, 2005; Thomas & Allport, 2000). The main difference between the present study and previous studies lies in the type of stimuli used. While the present study included only items with language-nonspecific orthography, previous studies entailed items with language-specific and language-nonspecific orthography. When items with language–specific orthography are used in the task, both words from the L1 and the L2 are recognized faster compared to a situation when items have language-non specific orthography (Orfanidou & Summer, 2005; Thomas & Allport, 2000). This effect seems to be greater for the weaker L2 than for the stronger L1 (e.g., L2 words benefit= 91ms and L1 words benefit= 26ms, in Thomas & Allport, 2000) and could be explained by the fact that language specific orthography is more helpful in a more complex situation (e.g., during L2 word recognition) than when the system is already relatively fast in recognizing the appropriate word (e.g., during L1 word recognition). Therefore, in case of language-specific orthography, for both L1 and L2 activation is mostly send to words of the relevant language, while words from the irrelevant language will be coactived very weakly. In support of this idea, Orfanidou and Summer (2005) showed that when only items with a language-specific orthography are used within a block, language switching costs are extremely reduced. However, the authors found that when language-specific and language-nonspecific stimuli are intermingled in the same block, the amount of switching costs for language-specific increases to the point that switching costs for language-specific and language-nonspecific stimuli become comparable. Moreover, replicating Thomas and Allport’s (2000) results, the authors did not find any significant interaction of language and switching costs, indicating overall symmetrical switching costs. Unfortunately, none of the two above-mentioned studies reported whether language switching patterns were modulated by orthography specificity (namely, no information was provided on the three-way interaction of language, switching costs and orthography specificity). This makes it difficult to compare results from those studies, in which language-specific and language-nonspecific stimuli were intermingled in the same experiment with the findings of the present study, which included only language-nonspecific stimuli. The present study is also difficult to compare with von Studnitz and Green’s (1997) study, despite the fact that, like in the present study, only stimuli with language-nonspecific orthography were used there. In their study, authors tested native speakers of German (L1) - highly proficient speakers of English (L2) in a bilingual lexical decision study. At the time of testing, all participants were living in the L2 environment, and reported being predominantly exposed to their L2 than to their L1. They reported being more proficient in the L1 than in the L2, but their performance on the language proficiency test (Meara test of word recognition; Meara, 1994) revealed that the difference between the two languages was very small. More precisely, in this proficiency test, for which the maximum score is 100 per language, participants scored in Exp.1, 89.7 and 87.4 for the

L1 and the L2 respectively, and in Exp. 2, 90.2 and 91.3 for the L1 and the L2 respectively. Therefore, it is possible that L2 lexical representation in this group of participants is not comparable to that of the participants of the present study. The same reasoning holds for Grainger and Beauvillain (1987) investigating bilingual word recognition in English (L1) - French (L2) speakers, who were pupils of bilingual schools in Paris and, therefore, equally exposed to the two languages.

Briefly, the aim of the present experiment was to assess whether language switching costs in bilingual language recognition are dominance related (the stronger the language the greater the cost; e.g., Green, 1998), or dominance-reversed (the stronger the language the smaller the cost; e.g., Dijkstra & van Heuven, 1998). We also considered the hypothesis that switching costs might not be influenced by language dominance given the relatively small dominance difference between L1 and L2 of our group of bilinguals. Results showed that switching costs were larger for the stronger than for the weaker language, indicating that language switching costs in bilingual comprehension are dominance-related. More precisely, we found that language switching was costly only for the L1, but not for the L2. We explained this result by suggesting that, in case of unbalanced bilinguals, when L2 words are presented, competitors from the L1 need to be suppressed. Therefore, when the L1 becomes relevant again, the previously applied inhibition has to be overcome, yielding L1 switching costs. However, when L1 words are presented, because of L2 words' lower baseline activation level, L2 candidates will not act as strong competitors. Therefore, the inhibition applied on L2 competitors will be weak or absent, leading to small or undetectable language switching costs when the L2 becomes relevant again.

Picture naming task

We turn now to the discussion of the picture naming task results. In the single-language block, we found that responses in the L1 were on average faster than in the L2, but that the difference was not significant. Comparable performance in the two languages might indicate that our participants were fairly balanced bilinguals. However, this assumption is in contrast with their performance in the lexical decision task, where participants responded faster on L1 words than on L2 ones.

Similar performance in L1 and L2 might be due to the order in which languages were presented in the single-language block. Recall that half of the participants named pictures first in the L1 and then in the L2 and the other half named pictures in the opposite order. Several studies investigating the effect of language order in language production have shown that, presented two separate language blocks, L1 naming is hampered if preceded by L2 naming. However, this negative effect is not found in the L2, where performance seems to benefit from previous naming in the L1 (Branzi et al., 2014; Levy et al. 2007; Misra et al., 2012). One way to interpret this asymmetry is by assuming that during

L2 naming, the L1 is strongly inhibited and that the inhibitory carry-over effect hampers L1 naming in the subsequent block. During L1 naming, however, weak or no inhibition is applied on the L2, leading to positive priming when the same pictures have to be named in a following L2 block (Misra et al. 2012). Therefore, negative priming on the L1 and positive priming on the L2 might explain, why we do not detect a reliable dominance effect in our group of bilinguals in the single-language block

Concerning language mixing costs, we found that trials of the single-language block were responded faster than repetition trials of the mixed-language block. The difference, however, was present only in the L1, but not in the L2, i.e. we found asymmetrical mixing costs. This effect is to be attributed to the fact that while in the single-language block, L1 and L2 were responded equally fast, in the mixed-language block, L1 responses became slower than L2 ones. The reversed dominance effect was mirrored by the accuracy data, showing similar error rates for the L1 and the L2 in the single-language block, and more errors for the L1 than for the L2 in the mixed-language block.

Better performance in the weaker than in the stronger language in the mixing language context is not a novel finding in language switching research (e.g., Christoffels et al., 2007; Costa & Santesteban, 2004; Costa et al., 2006; Schwieter & Sunderman, 2008; Verhoef et al., 2009; Verhoef, Roelofs & Chwilla, 2010). The paradoxical advantage of the weaker over the stronger language has been interpreted as speakers' unconscious strategy to help the weaker language, by lowering the activation level of the stronger language (e.g., Christoffels et al., 2007; Meuter, 2005). This effect can be attributed to proactive control that aims at anticipating and preventing interference before they occur (Braver, 2012). We suppose that, in the present study, interference from the stronger L1 was prevented by lowering its overall activation level (slower and less accurate performance in L1 than in L2). Differently from proactive control, reactive control is recruited as a late correction mechanism that activates only after a change has occurred (Jacoby, Kelley & McElree, 1999), such as a language switch. In this view, after a language change is detected, the non-relevant language is reactively inhibited, leading to switching costs when that language needs to be reactivated (Green, 1986).

Language switching costs. With regard to language switching costs, we found that repetition trials were named faster than switch trials and that this effect was the same for the L1 and the L2, i.e. symmetrical switching costs. Symmetrical switching cost for theee L1 and the L2 of highly proficient bilinguals have been extensively reported in the language switching literature (Costa & Santesteban, 2004; Costa et al., 2006; Declerck, et al., 2013; Fink & Goldrick, 2015). Within a less conservative dominance-related inhibitory account, it has been suggested that when the difference between L1 and

L2 dominance is small, then the amount of reactive inhibition applied on the two languages when they are non-relevant is similar, leading to symmetrical reactivation costs. With less conservative, it is meant that the relationship between language dominance and the amount of inhibition on that language is less "tight" compared to a more conservative inhibitory account, where language dominance and amount of inhibition on that language should be strictly related. Hence, in a less conservative account, asymmetrical switching costs are predicted only in cases of substantial dominance difference between two languages, while in a more conservative inhibitory account, any degree of dominance difference between two languages should lead to asymmetrical switching costs. Consequently, symmetrical switching costs between a stronger and a weaker language cannot be accounted for by a more conservative view of the IC model, according to which language switching costs are dominance-related (the stronger the language the larger the cost). In such a case, we should have found larger costs for the stronger L1 than for the weaker L2.

It could be argued that the symmetrical switching costs found in the production, but not in the recognition task, might be ascribed to the order in which the two tasks were presented (i.e., the lexical decision task preceded the picture naming task). In particular, prior practice of the items in the lexical decision task might have strengthened the weaker language to a greater extent than the stronger language, leading to a change in their dominance relation in the picture naming task. However, a comparison with a similar study suggests that the results obtained in the picture naming task are likely to be due to the nature of the task rather than to different practice effects for the two languages. Specifically, in a recent study by Mosca and Clahsen (2016), language switching costs were measured in a group of unbalanced bilinguals (L1 German – L2 English) performing a bilingual picture naming task. Participants were classified as highly proficient L2 speakers (C1 level of the CEFR), scoring 75.4% in the Oxford Placement Test. Despite the proficiency difference, results revealed symmetrical switching costs for the two languages and a tendency for reversed language dominance. In this picture naming task, there was no prior practice of the items. Based on this as well as previous evidence (e.g., Christoffels et al., 2007) it seems that in language production, symmetrical switching costs and reversed language dominance are to be expected in unbalanced bilinguals – highly proficient L2 speakers. Importantly, compared to Mosca and Clahsen (2016), the bilinguals tested in the present study were more proficient L2 speakers, scoring 84.4% at the Oxford Placement level (C1-C2 level of CEFR). This suggests that the pattern of results, namely symmetrical switching costs and reversed language dominance, found in the present study can be confidently attributed to the task itself together with the high proficiency reached in the L2 and less so by the fact that the weaker language might have benefited more than the stronger language from prior practice. Regardless, further studies are needed to assess the effect of practice on language control (see also Branzi et al., 2014; Declerck & Philipp, in press).

Yet, these results also indicate that, in the mixed-language block, the amount of language inhibition does not depend on its activation level. Indeed, if this was the case, given the reversed language dominance (faster and more accurate responses for the L2 than the L1), we should have found larger switching costs for the more activated L2 than for the less activated L1. A closer look at the data, however, revealed that switching costs in the L2 tended to be larger than in the L1, but that the difference did not reach significance. Therefore, higher activation for the L2 than for the L1, but comparable amount of switching costs for the two languages indicates the strength of activation of languages(regulated by a proactive control) does not directly modulate languages' strength of inhibition (regulated by a reactive control). This hypothesis is supported by the finding that low proficient L2 speakers can show a reversed dominance effect (faster responses in the L2 than in the L1) together with larger switching costs for the L1 than the L2 (Costa & Santesteban, 2004; Tarlowski, Wodniecka & Marzecová, 2013; Verhoef et al., 2009). In particular, this might indicate that when the proactive control lowers the activation level of the stronger language, the reactive control is not immediately affected by this adaptation and continues to perceive the stronger language as more active than the weaker language (but see Declerck et al., 2015).

Crucially, in a study with unbalanced bilinguals, Verhoef et al. (2009) reported reversed language dominance and larger switching costs for the L1 than for the L2 when speakers were given less time to prepare for the next trial (Cue to Stimulus Interval, CSI= 750ms), but symmetrical switching costs when they had more time to prepare (CSI= 1500ms). This result indicates that also unbalanced bilinguals can show symmetrical switching costs and that switching costs might depend on the task. In particular, it seems that in easier tasks (e.g., having more time to prepare), the difference between the stronger and the weaker language becomes more difficult to detect. Therefore, the fact that advanced L2 speakers often show symmetrical switching costs might indicate that the task is not subtle enough to determine the difference between the stronger and the weaker language, leading to overall ceiling effects. This leads to the question, why did we find ceiling performance in the production, but not in the recognition task?

The reason might lie in the difference between the processes supporting language production and recognition. Indeed, while language recognition is mostly supported by a bottom-up mechanism, language control in production is a mainly top-down process (see next session for further discussion on bottom-up vs. top-down control). As Yeung and Monsell (2003) observed, top-down control is effortful and, therefore, minimized where possible. In particular, Monsell and colleagues (2000) suggested that greater inhibition on the stronger task might be a useful strategy only when the activation strength of two tasks is extremely unequal, but not when the difference does not reach a certain threshold. Therefore, it might be the case that symmetrical switching costs between a stronger and a weaker language in language production are due to a process that aims at minimizing the effort.

Furthermore, Monsell et al. (2000) showed that more activation for the stronger than for the weaker task refers to languages' activation level at the beginning of the experiment, but that this state can be changed by practising the tasks during the experiment. Based on this, symmetrical switching costs could be attributed to a change in the amount of language suppression over time. Specifically, while at the beginning of the task the stronger language is suppressed more because it is perceived as more activated than the weaker, after a certain amount of practice, the actual strength of language activation might be perceived (the weaker language is activated more than otherwise), leading to larger switching costs for the weaker than for the stronger language. Such a modulation of language inhibition over time would yield overall symmetrical switching costs. To address this issue, however, future studies need to investigate whether the way languages are controlled remains unvaried for the duration of the task or whether the equilibrium between the languages is affected by practice.

Moreover, one might speculate that if language strength of inhibition can slowly adapt to language strength of activation, then symmetrical switching costs should always appear together with a similar or reversed dominance effect (faster responses for the weaker than the stronger language). This is because, if language strength of activation is reversed (the weaker language is activated more), then the amount of inhibition might be at first greater for the stronger than for the weaker language (default state irrespective of the actual languages' activation level) and afterwards greater for the weaker compared to the stronger language (language inhibition aligns with the actual activation level of the two languages), leading to overall symmetrical switching costs (probably with a tendency for greater L2 switching costs). If language strength of activation is similar, language inhibition is initially greater for the stronger than for the weaker, and then it adapts to the actual language strength of activation, yielding symmetrical switching costs. Support for this observation comes from multiple studies reporting symmetrical switching costs together with either same naming latencies for the stronger L1 and the weaker L2 (e.g., Calabria et al., 2011; Fink & Goldrick, 2015) or with faster responses in weaker than in the stronger language (e.g., Costa & Santesteban, 2004; Costa et al., 2006; Christoffels et al., 2007; Gollan & Ferreira, 2009; Mosca & Clahsen, 2016; Verhoef et al., 2010). However, this hypothesis is challenged by the finding that symmetrical switching costs can also occur together with faster responses in the L1 than in the L2 (Calabria et al., 2011; Declerck et al., 2012; Declerck, Koch & Philipp, 2015). Therefore, it remains to be determined whether language strength of activation and inhibition are two completely separate mechanisms or whether and to what extent they can influence each other.

Briefly, the goal of the picture naming task was to determine whether language inhibition in production is dominance-related as suggested by the IC model (Green, 1998) or whether this assumption is valid only in cases of great dominance difference between the stronger and the weaker language. Moreover, the study aimed at investigating whether language control in production and reception relies on the same mechanisms. Results of the picture naming task showed that, in a group of advanced L2 speakers, switching costs for the stronger L1 and the weaker L2 were symmetrical, indicating that the amount of language inhibition did not strictly depend on language dominance. In the lexical decision task, however, switching costs were larger for the stronger L1 compared to the weaker L2, suggesting that the amount of inhibition was influenced by language dominance. Moreover, in the picture naming task, we found that responses were faster in the L2 than in the L1. This is not what we found in the lexical decision task, where bilinguals showed a typical dominance effect (faster responses in the L1 than in the L2). Overall, our results suggest that bilingual language control in production and recognition might rely on different mechanisms.

Picture naming vs. Lexical decision task

In the production task, we found that L2 responses were named faster than L1 ones. This pattern has been interpreted as speakers' unconscious strategy to help the weaker language, by making the stronger language less available. This is not what we found in the recognition task, where words were responded to faster in the L1 compared to the L2. One might suppose that the different dominance effect found in reception and production might depend on task goals. While in the production task, the goal of the task was to name pictures in one or the other language, in the reception task, the aim of the task was to decide whether a string of letters was a real word or not irrespective of language membership. Therefore, it may be the case that bilingual language control relies on strategies to regulate languages' activation strength only when language membership is relevant for the task. However, previous bilingual lexical decision studies using a language-specific paradigm (i.e. deciding whether a string of letters is a word in a determined language) do not report faster responses in the weaker than in the stronger language (e.g., Thomas & Allport, 2000; von Studnitz & Green, 1997). This suggests that language membership is not the reason why reversed language dominance occurs. More generally, we are not aware of any recognition study with bilinguals reporting better performance in the weaker than in the stronger language. This indicates that the reversed language effect seems to be limited to bilingual production, and not to bilingual recognition.

Language production and language recognition are, indeed, two different processes. While in language production, information mainly flows in a top-down fashion (from the concept, to the lemma, down to the phonological and then to the articulatory level), language recognition is a predominantly bottom-up process (from letters, up to word recognition, to the phonological representation and then to the concept level). Yet, top-down and bottom-up attentions are commonly considered two different types of information processing (e.g., Carrasco, 2011; Pinto et al., 2013). Top-down attention is goal-oriented, meaning that attention is voluntary allocated to certain features (e.g., Beauchamp, Cox, & Deyoe, 1997). Bottom–up attention is stimulus-driven, indicating that certain stimuli can attract attention, even though the subject is not doing so intentionally (e.g., Schreij, Owens, & Theeuwes, 2008). Thus, the main difference between top-down and bottom-up attention is that the former is voluntary/non-automatic, while the latter is involuntary/automatic. Moreover, top-down attention is supposed to be more flexible than bottom-up attention (e.g., Shiffrin & Schneider, 1977). Specifically, the hypothesis is that since top-down attention is voluntary, resources can be strategically allocated depending on the task (e.g., Giordano, McElree & Carrasco, 2009; Jonides, 1981; Kinchla, 1980).

Because of this, one might suggest that the differences in the picture naming and the lexical decision task can be attributed to the different nature of their underlying processes. While in production, bilingual language control predominantly relies on a flexible and strategic top-down mechanism, in recognition, bilingual language control is mainly supported by a more rigid bottom-up process. The flexible nature of the top-down control could explain why we found reversed dominance effect (interpreted as a strategy to prevent interference from the stronger language) in the production, but not in the recognition task.

Similarly, the different language switching pattern found in the recognition and the production task could be ascribed to the difference between top-down and bottom-up mechanisms. In the recognition task, asymmetrical switching costs might be attributed to the fact that since words are more activated in L1 than in L2, L1 words need to be inhibited more than L2 words, leading to larger reactivation costs for the L1 relative to the L2. In the picture naming task, symmetrical switching costs for the stronger and the weaker language could be explained either by a strategic process of costs' minimization (i.e. similar amount of inhibition on L1 and L2) or by a modulation of language inhibition over time (stronger inhibition on L1 than L2 at the beginning of the task, but reversed inhibition pattern afterwards). However, more studies are needed to assess whether the difference between language control in production and recognition are to be attributed to the fact that they are mainly supported by two different mechanisms of information processing. A more systematic way to address this issue would be, for example, by including a single-language block for the recognition task, so to compare language mixing costs in recognition and production. Alternatively, it could be useful to use a language-specific recognition task, in which only words are used and where speakers have to decide whether a given word belongs to either L1 or L2. This type of task would be more comparable to a picture naming task, since only existing words are used and language membership is relevant for the task.

Conclusion

The aim of this study was to investigate bilingual language control in reception and production. In particular, we asked whether language control in bilingual perception relies on the same mechanism as in bilingual production and whether these two processes can be incorporated in a single bilingual language control model. To investigate this issue, we considered two prominent models of bilinguals language control, according to which, when one language is in use, the other is suppressed. The Inhibitory Control (IC) model (Green, 1986, 1993, 1998) suggests that the amount of inhibition applied on the non-relevant language depends on its dominance (the stronger the language the greater the inhibition). The Bilingual Interactive Activation (BIA) model (Dijkstra & van Heuven, 1998; Grainger & Dijkstra, 1992, van Heuven et al., 1998), proposes that inhibition from the stronger to the weaker language is greater than the other way around (the stronger the language the weaker the inhibition). Therefore, in case of language dominance difference, both types of models predict different amount of inhibition for the two languages and, consequently, different amount of reactivation costs (asymmetrical switching costs). For the IC model switching costs are dominance-related (the stronger the language, the greater the cost), whereas for the BIA model switching costs are dominance-reversed (the greater the language, the smaller the cost). Based on previous findings, we also considered the possibility that when the difference between L1 and L2 dominance level is relatively small, the magnitude of inhibition applied on the two languages might be comparable (symmetrical switching costs).

To investigate the mechanisms underpinning language control in production and recognition, we tested native speakers of Dutch (L1) - highly proficient speakers of English (L2) in a lexical decision and a picture naming task with language switching. Participants reported being more proficient and more exposed to the L1 compared to the L2. Results from the bilingual lexical decision task confirmed that L1 and L2 were processed differently. Specifically, we found that language switching was costly only for the L1 but not for the L2. We suggested that when L2 words are presented, L1 competitors need to inhibited. If during L2 processing, L1 words are suppressed, reactivating the L1 on a switch trial will lead to L1 switching costs. However, when L1 words are presented, the target word is fast recognized and words from the weaker L2 will not have enough time/strength to compete. In this case, the inhibition applied on the weaker L2 competitors will be relatively small or absent, yielding to undetectable L2 switching costs when the L2 becomes relevant again.

These results are in contrast with the BIA model, according to which switching costs are dominance-reversed, but they are in line with the dominance-related IC model, in that we measured switching costs in the stronger L1, and only a tendency towards switching cost in the L2.

In contrast to the lexical decision task, in the picture naming task we measured the same amount of switching costs for the L1 and the L2. Symmetrical switching costs between a stronger and a weaker language cannot be accounted for by a conservative dominance-related account of language control. We suggested that symmetrical switching costs in bilinguals with a relatively small difference between L1 and L2 proficiency level might be related to speakers' temporary adjustment to the task. Specifically, we assume that in the production task, the language control system strategically adapts to the task either by minimizing the effort of switching between languages (i.e. applying a comparable amount of inhibition on L1 and L2) or by adjusting the amount of language inhibition during the task (greater inhibition on L1 than L2 at the beginning of the task, and greater inhibition on L2 than L1 towards the end of the task). We proposed that the difference between language control in production and recognition might be due to the fact that they mostly rely on different mechanisms. While bilingual recognition is predominately supported by a more rigid bottom-up process, production mainly depends on a more flexible top-down control. Consequently, while in recognition the amount of language inhibition is strongly related to language default activation levels (L1 words more activated than L2 words), in production the amount of inhibition applied on the non-relevant language may be modulated by speakers' internal strategies. In conclusion, the present study suggests that bilingual production and recognition rely on two different mechanisms and cannot be accounted within one of the existing language control model.

Appendix A. *Participants*

All participants were native speakers of Dutch (L1), late learners of English (L2). Twenty-five of them reported knowing an additional language, nineteen had knowledge of two additional languages and fifteen knew three additional languages. German was one of those languages for seventeen participants, French for fifteen, Spanish for eleven, Swedish for five, Frisian for three, Italian for two, Greek, Latin and Low Saxon for one person respectively. Six of them reported being fluent in speaking an additional language and four of them in reading in an additional language. All participants considered themselves fluent speakers and readers of both L1 (Dutch) and L2 (English). Below are reported the information about the self-assessed language history and competence in L1 (Dutch) and L2 (English) from the Language Experience and Proficiency Questionnaire (LEAP-Q).

Table A1. Mean scores (standard deviation in brackets) about language age of acquisition (AoA) and age (in years) when participants became fluent in L1 (Dutch) and L2 (English).

	AoA	Speaking fluency	Reading AoA	Reading fluency
L1	0	5.74 (2.15)	5.11 (1.28)	8.66 (2.01)
L2	9.35 (2.34)	16.74 (2.37)	11.16 (2.36)	16.60 (2.04)

Table A2. Mean scores (standard deviation in brackets) of the self-rating task based on a ten-point scale (0= no knowledge, 10= perfect knowledge) for speaking, comprehension and reading skills in L1 (Dutch) and L2 (English).

	Speaking	Comprehension	Reading	Mean
L1	8.82 (0.98)	9.50 (0.69)	9.17 (0.90)	9.16 (0.85)
L2	6.96 (0.96)	8.06 (1.14)	7.93 (1.04)	7.65 (1.04)

Table A3. Mean scores (standard deviation in brackets) relative to the amount of current language exposure in L1 (Dutch) and L2 (English) based on a ten-point scale (0= never exposed, 10= always exposed) in different contexts (interaction with friends and family, during reading, watching TV, listening to the radio, self-instruction and language tapes).

	Friends	Family	Watching TV	Listening to radio/music	Reading	Language lab/ Self-instruction
L1	7.75 (1.71)	8.46 (2.45)	5.10 (1.98)	3.60 (2.24)	4.57 (2.48)	1.50 (2.44)
L2	4.65 (2.67)	0.79 (1.39)	7.42 (1.64)	6.75 (2.50)	7.41 (1.84)	2.89 (3.05)

Table A4. Mean scores (standard deviation in brackets) relative to the amount (in years) of prior language exposure in L1 (Dutch) and L2 (English) in different environments (in a country, in a family, at school/work).

	Country	**Family**	**School/Work**
L1	21.39 (4.57)	20.24 (1.49)	18.66 (5.44)
L2	0.38 (1.54)	0 (0)	2.55 (3.73)

Table A5. Mean scores (standard deviation in brackets) relative to the amount of contribution of different factors (interaction with friends and family, reading, watching TV, listening to the radio, self-instruction and language tapes) in learning L1 (Dutch) and L2 (English). Scores are based on a ten-point scale (0= not a contributor, 10= most important contributor).

	Friends	**Family**	**Watching TV**	**Listening to radio/music**	**Reading**	**Language lab/ Self-instruction**
L1	7.96 (1.23)	8.37 (1.89)	6.10 (1.51)	3.89 (2.56)	7.58 (1.23)	1.82 (2.46)
L2	5.20 (3.02)	1.46 (2.45)	7.93 (1.33)	4.89 (2.50)	7.96 (3)	3.63 (3.10)

Appendix B. *Materials (L1 Dutch, L2 English)*

List of the Words/Picture names (a) and Pseudowords (b) used:

a) *Bril, glasses; Borstel, brush; Citroen, lemon; Fles, bottle; Handschoen, glove; Horloge, watch; Jurk, dress; Kers, candle; Knop, button; Lepel, spoon; Potlood, pencil; Schommel, swing; Vogel, bird; Wolk, cloud.*
b) *Bris, smartes; Garstel, brunk; Cichoon, tewon; Snes, boddle; Handspleen, wrove; Melloge, wamps; Jark, pless; Hers, bantle; Blop, cunton; Remel, spean; Petloog, runcil; Scharmel, pring; Bosel, bime; Wolm, croud.*

Acknowledgements

The research presented in this article was supported by an Erasmus-Mundus Joint Doctoral scheme for ‘International Doctorate for Experimental Approaches to Language and Brain’ (IDEALAB) awarded to MM by the European Commission under grant no. <2012-0025/2013-1458-EMII EMJD>.

7 General Discussion

7.1 Summary

The present dissertation investigated the mechanisms underpinning language control in multilingual speakers and which factors might play a role during this process. The most influential model of language control, the Inhibitory Control (IC) model (Green, 1998), assumes that in both production and recognition the way in which languages are controlled depends on their relative dominance. Specifically, it proposes that in a mixed-language situation both relevant and non-relevant languages compete for selection and that the competition is solved by inhibiting the non-relevant language. Further, the degree of inhibition applied to the irrelevant language depends on its dominance level, that is, the stronger the language the greater the inhibition. The present dissertation is devoted to the investigation of this assumption. In particular, the work explores to what extent language control is affected by language dominance as formulated by the IC model and whether and in what way other factors might modulate this process. Principally, it assesses whether language control is affected by (i) the time speakers have at their disposal to prepare for an utterance, (ii) the type of languages (more vs. less similar languages) involved in the interactional context, and (iii) the modality (production vs. recognition) in which languages need to be controlled. Moreover, in addition to the IC MODEL, two alternative specifications regarding the mechanisms underpinning language control are considered within this dissertation, namely, the LANGUAGE SPECIFIC ACCOUNT for production and the BIA MODEL for recognition.

With regard to language dominance, Table 1 summarizes the main results obtained in the four publications presented in this dissertation. Specifically, it illustrates whether the way in which languages are controlled is determined by language dominance as asserted by the IC model (x= no; √= yes).

Table 1. The effect of language dominance on language control across the four publications.

	Publication I	***Publication II***	***Publication III***	***Publication IV***
	Bilinguals	*Trilinguals*	*Trilinguals*	*Bilinguals*
Production	x	x	x	x
Recognition	-	-	-	√

Briefly, Table 1 shows that for both bilinguals and trilinguals, language control in production is not determined by language dominance as suggested by the IC model. In fact, with regard to bilinguals, the data obtained reveal a comparable amount of switching costs for the stronger L1 and the weaker L2 (Publication I & IV). As for trilinguals, in two different groups (Publication II & Publication III, Exp. 2), language switching costs were found to be larger for the stronger native language (L1) compared to the weaker non-native languages (L2-L3), while the amount of switching costs did not differ between the stronger L2 and the weaker L3. In a third group of trilinguals (Publication III, Exp. 1), data showed that switching costs were smaller for the weaker non-native language (L3) compared to both the stronger native and the non-native language (L1-L2), whereas the switching costs incurred between the stronger L1 and the weaker L2 were similar. Overall, these results suggest that, in production, language switching costs are not strongly related to language dominance as assumed by the IC model. With regard to recognition, however, data showed larger switching costs for the stronger L1 than for the weaker L2 in bilingual speakers (Publication IV). This pattern suggests that in bilingual recognition, language control does seem to be regulated by language dominance as predicted by the IC model.

With respect to the three main factors considered in this thesis, namely (i) preparation time, (ii) language typology and (iii) processing modality, the studies reported in this dissertation clearly show that language control is strongly affected by these variables. Concerning the two alternative specifications of the IC model, namely (a) the language specific account for production and (b) the BIA model for recognition, results indicated that the way languages are controlled cannot be successfully predicted by these two accounts. The main findings concerning both the factors influencing language control and the alternative specifications to the IC model are discussed in the next sections.

7.2 Preparation Time

The first issue addressed in this dissertation is the role of preparation time on language control. Specifically, it was examined whether giving speakers more time to prepare for a specific language can affect the way languages are controlled in a mixed language context. This issue was explored in relation to both bilinguals (Publication I) and trilinguals (Publication II). To anticipate the main findings, the data obtained reveal that preparation time can drastically alter language switching patterns, while language dominance seems to play a less decisive role in determining language switching costs.

In Publication I, late unbalanced bilinguals were tested in a bilingual picture naming task involving preparation time manipulation (with vs. without preparation time conditions). The study showed that, when no preparation was given, language switching was a costly process. However, when speakers were informed in advance about the language to be used next, language switching became a cost-free process. This is an important finding, since it shows that the cost needed to reactivate a language strongly depends on preparation time, and therefore less so on language dominance. Crucially, this result also indicates that the cognitive system is able to fully prepare for a language switch, challenging the assumption that is impossible to eliminate switching costs (e.g., Roger & Monsell, 1995). In Publication II, late unbalanced trilinguals were tested in a trilingual digit naming task involving preparation time manipulation (shorter vs. longer preparation time conditions). Similar to the bilingual study in Publication I, the study showed that when preparation time was relatively short, language switching was costly. However, when speakers were given more time to prepare, language switching was cost-free. These results extend the finding discussed in Publication I, by showing that not only in bilinguals but also in trilinguals can language switching costs be dramatically altered by preparation time. Further, it shows that, like bilinguals, trilinguals are able to prepare for a language switch in advance as well.

Transient control. The ability to prepare in advance for a new task is linked to a type of temporary or transient control that takes place as soon as a sudden change is encountered (e.g., a cue indicating a language switch). Results from Publications I and II give us important information with regard to transient control of languages in multilingual speakers. Indeed, both studies show that multilingual speakers can prepare for a language switch in advance and that switching costs can be fully eliminated.

In Publication I, participants were given 800ms to prepare for the upcoming trial. Based on this, one may be tempted to think that a preparation time of 800ms is enough to allow full reconfiguration of the new language. However, previous studies have shown that a preparation time of 800ms or even longer is not sufficient to allow for advanced reconfiguration of the new task (e.g., (CSI= 800ms in Costa & Santesteban, 2004; CSI= 1000ms in Declerck, Koch, & Philipp, 2012; CSI= 1500ms in Fink & Goldrick, 2015; CSI= 800ms in Ma, Li, & Guo, 2016; CSI= 1000ms in Philipp et al., 2007; CSI= 1500ms in Verhoef, Roelofs, & Chwilla, 2009). The difference between the study presented in Publication I and previous studies lies in the interval between successive trials. While the inter-trial time here was fixed and relatively long (ITI= 2400ms), in previous studies the interval between two trials (as expressed by RSI or ITI) was either variable or relatively short (RSI= 1150ms Costa & Santesteban, 2004; RSI= 1400ms Declerck et al., 2012; ITI= 1000/1250ms in Fink & Goldrick, 2015;

ITI= 500ms in Ma et al., 2016; RSI= 1100ms in Philipp et al., 2007; ITI= 1500/2300ms in Verhoef et al., 2009). Hence, it is the “cooperation” between a relatively long interval between trials and preparation time that allow for switching cost elimination (see also Goschke, 2000). This raises the question, why are switching costs eliminated?

As mentioned in the introduction, switching costs might represent the effort of disentangling from the previous relevant task (Wylie & Allport, 2000), as well as the cost of preparing for a new one (Roger & Monsell, 1995). In Publication I, we have seen that, despite preparation time manipulation, the interval between one stimulus and the next was kept constant (3200ms). This means that in both preparation time conditions, the interval in which to recover from the previously applied inhibition was the same. However, it was only when the language cue was displayed beforehand that switching costs were eliminated, not when language cue and stimulus were presented simultaneously. These results suggest that is the display of the language cue that facilitates switching cost elimination. This might be due to two alternative explanations. The first is that longer intervals between trials allow for the dissipation of the persisting inhibition stemming from the previous trial, so that when the language cue is displayed beforehand, the new language can be prepared. The second is that longer intervals between trials facilitate recovery from the inhibition of the preceding trial and that this process is accelerated by the presentation of the language cue so that switching into that language can be prepared in advance.

This issue can be better explored in Publication II. Here, language switching costs were eliminated when speakers were given 1000ms preparation time (CSI), but only 500ms intervals between trials (RCI). Language switching, however, was a costly process when speakers had just 150ms time to prepare (CSI), but 1350ms intervals between trials (RCI). Considering that switching costs tended to be dominance-related, we might assume that switching costs mainly reflected persisting inhibition stemming from the previous trial (stronger languages are inhibited more than weaker ones) and not merely the cost of reconfiguring a new task, which is not expected to be modulated by language dominance. Thus, these data indicate that, paradoxically, the effect of persisting inhibition was detectable when speakers were given more time between their response and the display of the next language cue (1350ms), but not when the interval between their response and the new language cue was shorter (500ms). The reason for this might be that it is the display of the language cue that boosts the reactivation of a previously inhibited language, so that the longer the interval between the language cue and the stimulus, the more successful is the dissipation of the inhibition coming from the preceding trial.

Sustained control. In contrast to transient control, sustained control is believed to be a more prolonged type of control that affects a language as a whole (both repetition and switch trials). In particular, it has been suggested that sustained control aims at preventing interference between languages before it occurs by altering languages' activation level (making one language less or more available than the other). Results from both Publications I and II showed that, like transient control, sustained control can also be affected by preparation time manipulation. In particular, both studies indicate that languages' activation level is modulated by task demand, with stronger languages being more hampered than weaker languages when less time to prepare for the upcoming language is given (more demanding task).

In Publication I, sustained control was investigated by looking at mixing costs, that is, the difference in performance between trials in the single-language block and repetition trials in the mixed-language block. Results revealed that when preparation time was given, responses in the weaker L2 became faster in the mixed- compared to the single-language block (L2 mixing benefit). In contrast, performance in the stronger L1 was comparable in the single- and mixed-language blocks. Moreover, when no preparation time was given, the L2 was responded faster in the mixed- compared to the single-language block (L2 mixing benefit), while the opposite pattern was found for the L1. In this last case, responses were slower in the mixed- than in the single-language block (L1 mixing costs). These results indicate that compared to the single-language block, in the mixed-language block languages' activation level is adjusted. More precisely, L2 facilitation and L1 disadvantage can be attributed to speakers' strategies designed to help the weaker language in a language switching context by (i) enhancing the activation level of the weaker L2 and/or (ii) lowering the activation level of the stronger L1. Furthermore, these results show that languages' strength of activation in the mixed-language block is affected by preparation time. Specifically, L2 mixing benefit was found in both preparation time conditions, whilst L1 mixing costs were measured only when no time was given to prepare. This pattern can be attributed to the fact that the two conditions differed in complexity. Whilst a with-preparation time condition is a less demanding task since the language cue is processed before the stimulus, a condition in which no preparation is given is more demanding, in that language cue and stimulus have to be decoded simultaneously. Therefore, while in the with-preparation condition the level of uncertainty is relatively low, as information about the language to be used is given beforehand, in the without-preparation condition the uncertainty level is higher and might require additional controlling processes. This might be the reason why in the less demanding (with preparation time) condition only one controlling step was taken (L2 facilitated), and in the more demanding (without preparation time) condition an extra controlling step was required (L2 facilitated + L1 inhibited). This is a crucial finding, since it shows that mixing costs are not as fixed as previously assumed, but can be influenced by trial-specific manipulation, such as preparation time variation. More precisely, it suggests that mixing costs do not reflect the effort of keeping multiple tasks active in working memory as previously suggested (Braver et al., 2003; Kray & Lindenberger, 2000; Mayr, 2001; Roger & Monsell, 1995). Indeed, if this was the case, a similar amount of mixing costs should have been found in the two preparation time conditions. Rather, it indicates that mixing costs reflect a strategic adjustment of languages' activation level and that this adjustment depends on the task demand.

In Publication II, the attention was extended to trilinguals. Here, sustained control was examined by comparing languages' activation level in the shorter and longer preparation time conditions. In this respect, the results indicated that in all three languages, responses were faster in the longer compared to the shorter preparation time condition. This is not surprising, if we consider that a task becomes easier to perform if more time is given to prepare for it. In this sense, a condition with more preparation time is less demanding compared to a condition where less preparation time is provided. The disadvantage of being in a more demanding condition (shorter preparation time), however, was not the same for the three languages. In fact, it was greater for the stronger L1 than for the intermediately strong L2, and for the L2 compared to the weaker L3. The greater disadvantage for the stronger than for the weaker language was interpreted as a means of preventing interference from a stronger into a weaker language when the uncertainty level of the task is enhanced (e.g., shorter preparation time). The disadvantage seems to increase along with language dominance, with stronger languages being more taxed than weaker ones.

However, it should be noted that when comparing performance levels in the two preparation time conditions, both repetition and switch trials were included. Recall that switching costs were eliminated in the longer preparation time condition, but not when preparation time was shorter, where switching costs were larger for the stronger than for the weaker language. Therefore, the fact that the difference in performance between shorter and longer preparation time conditions was greater for the stronger compared to the weaker language, may be driven by the fact that switching costs in the shorter preparation condition tended to increase along with language dominance. A way to neatly see how languages' activation level was affected by the preparation time manipulation and disentangle it from the effect of language switching, is by looking at the raw naming latencies for repetition trials in the shorter and longer preparation time conditions illustrated in Table 1 of Publication II. As reported in Table 1, the difference between shorter and longer preparation times was numerically larger for the L1 (119ms) than for the L2 (56ms), and for the L2 compared to the L3 (39ms). The trilinguals tested in this study were proficient speakers of their native language (L1), but less proficient speakers of their two additional non-native languages (L2 and L3), where the difference in proficiency was relatively small. This considered, it seems that the difference between the two preparation time conditions tended to increase along with language dominance. A post-hoc analysis confirmed this hypothesis by showing that compared to the longer preparation time condition, in the shorter preparation time condition, the L1 responded more slowly compared to both L2 ($\beta = .14$, $SE = .01$, $t = 10.96$, $p < .0001$) and L3 ($\beta = .19$, $SE = .01$, $t = 13.94$, $p < .0001$), and that the difference between L2 and L3 was also significant ($\beta = .04$, $SE = .01$, $t = 3.10$, $p < .05$).

These results support the hypothesis that in more demanding tasks (shorter preparation time), stronger languages are more taxed than weaker languages. Importantly, they extend the assumption made in Publication I, by showing that, like bilinguals, languages' activation level in trilinguals is affected by task demand. Moreover, the fact that task demand affected the stronger and weaker non-native languages (L2 and L3) differently suggests that the way languages are controlled depends on their dominance (stronger vs. weaker language) and not so much on their status (native vs. non-native language). Indeed, if language status played a decisive role in language control, we should have found different effects of task demand for the native L1 compared to both non-native L2 and L3, but no differences between the non-native L2 and L3 (see next section for further discussion on the role of language status on language control).

7.3 Language Typology

The second issue addressed in this dissertation is the role of language typological closeness (more vs. less similar languages) in language control; this topic was examined in two groups of trilinguals (Publication III, Exp. 1 & 2). The data obtained reveal that language similarity affects the way languages are controlled and that language dominance alone does not suffice to explain language switching patterns. Before exploring the main findings with regard to the role of language similarity on language control, the next section will discuss the limited effect of language dominance on language switching patterns in relation to the four studies reported in this work.

The main goal of the present dissertation is to investigate whether, in a language switching context, the costs required to reactivate a language depend on its dominance level as asserted by the IC model, or whether other factors might be relevant during this process. The results discussed above clearly indicate that language switching costs in production are profoundly influenced by preparation time, but less so by language dominance. Precisely, when speakers were given longer time to prepare for the upcoming language, switching costs were eliminated in all languages irrespective of their dominance level. However, when speakers were given less or no time to prepare, switching costs seemed to be sensitive to language dominance. For example, when less preparation time was given, the unbalanced trilinguals tested in Publication II showed larger switching costs for the stronger language (L1) compared to the two weaker languages (L2 and L3). This result was replicated in Publication III (Exp. 2), where unbalanced trilinguals showed larger switching costs for the stronger L1 than for the weaker L2 and L3. Further, Publication III (Exp. 1) reported smaller switching costs for the weaker L3 relative to the stronger L1 and L2. Based on this, it can be assumed that language switching costs are sensitive to language dominance, in the sense that if language switching costs differ between languages, this is because they are larger for the stronger than for the weaker language, but not vice versa. Nevertheless, the experiments reported in this dissertation also show that language dominance is not strongly determining language switching patterns. Specifically, in both Publication II and Publication III (Exp. 2), trilinguals showed a similar amount of switching costs for the stronger L2 and the weaker L3. In Publication III (Exp. 1), switching costs were comparable between the stronger L1 and the weaker L2 in trilingual speakers. Finally, Publication I and Publication IV (Exp. 1) reported a similar amount of switching costs for bilinguals' stronger L1 and weaker L2. Altogether, these data suggest that language dominance alone does not suffice to explain language switching patterns. Thus, how can we explain symmetrical switching between a stronger and a weaker language?

In Publication II, we suggested that switching costs between the stronger L2 (French/German) and weaker L3 (German/French) were symmetrical, probably because the dominance difference between these two languages was not enough to yield an asymmetrical switching pattern. The conclusion that the dominance levels in L2 and L3 were comparable was based on the observation that when no preparation time was given, naming latencies in L2 and L3 did not differ. However, it is worth noting that when performance levels in the shorter and longer preparation times were merged (in the overall model), naming latencies in the stronger L2 were significantly faster than in the weaker L3. This difference in naming speed disappeared when data in the shorter preparation time condition were observed separately from those in the longer preparation time condition. This might have been due to the fact that, as already discussed above, in the more demanding shorter preparation time condition, stronger languages are more hampered than weaker languages. Hence, while in the longer preparation time condition the L2 was responded faster than the L3, in the shorter preparation time condition performance levels in the L2 and L3 became comparable. This was not the case for the L1 and L2 language pairing. Here, the difference in naming speed between the two languages was greater in the shorter preparation time condition and, therefore, even if reduced, was still detectable in the longer preparation time condition. Therefore, the L2-L3 symmetrical switching costs found in the shorter preparation time condition cannot be attributed to the fact that the two languages did not differ enough in their proficiency (since overall the L2 was named faster than the L3), but rather to the fact that in the more demanding, shorter preparation time condition, the activation level of the L2 was lower compared to the less demanding, longer preparation time condition. Thus, in the shorter preparation time condition, the more comparable activation level of the stronger L2 and the weaker L3 might have led to a comparable amount of switching costs for the two languages. In Publication III (Exp. 2), despite the symmetrical switching costs between the stronger L2 and weaker L3, naming latencies in both the single- and mixed-language blocks were faster in the L2 than in the L3. This indicates that the language activation level was consistently higher for the stronger L2 compared to the weaker L3 and that the symmetrical switching costs between these two languages cannot be ascribed to the fact that their activation levels were comparable. In this experiment, participants were native speakers of German (L1), highly proficient learners of English (L2), and lower proficient speakers of Italian (L3). Relying on Schepens and colleagues' study (2013) comparing the cognate distributions across six European languages (Dutch, English, German, French, Italian and Spanish), it was possible to determine the typological distance (on the lexical level) between the three languages included in the experiment. Specifically, by measuring the degree of orthographic and phonological similarity between language pairs (where 0= no overlap between languages and 1= identical languages), Schepens et al. (2013) showed that English and German are more closely related languages (.76) than both English and Italian (.47) and German and Italian (.44). Therefore, for this group of trilinguals, the L1 (German) and L2 (English) were typologically closer languages compared to the L1 (German) and L3 (Italian), as well as the L2 (English) and L3 (Italian). Here, symmetrical switching costs between the stronger L2 and the weaker L3 were explained with the suggestion that because of the typological closeness to the stronger L1 (German), naming in the weaker L2 (English) was facilitated by decreasing the amount of inhibition applied to this language. Specifically, it was suggested that the controlling system might have tended to prevent unwanted interference from the stronger and typologically closer L1 (German) into the weaker L2 (English) by facilitating L2 reactivation. This was achieved by decreasing the amount of inhibition applied to the L2 when it was irrelevant, so as to facilitate its reactivation when it became relevant again. This conclusion was drawn after observing that in the L2, language switching costs tended to be smaller compared to both the L1 and the L3 and that most of the errors speakers made in the L2 were coming from the stronger L1. Thus, similar to Publication I, it was hypothesized that during language switching speakers might rely on unconscious strategies to prevent conflicts between languages.

Moreover, it is worth noting that while language typology modulated switching costs (reflecting transient control), it did not seem to have a visible effect on mixing costs (indexing sustained control). Indeed, while language switching costs for the L2 were smaller compared to what was predicted by the IC model, the amount of mixing costs did not differ across the three languages. Impulsively, one might be tempted to think that in order to facilitate L2 naming in the mixed-language block, the activation level of this language would need to be enhanced. This would be achieved by reducing the amount of sustained control on the L2 and would be measurable in smaller mixing costs for this language. However, if the activation level of the L2 is enhanced, the amount of inhibition required to suppress this language would also increase, leading to larger instead of smaller reactivation costs. Therefore, enhancing the activation level of the L2 would not be an effective strategy to facilitate its reactivation.

The effect of language typology on language control was clearly pinpointed by another group of trilinguals tested in Publication III (Exp. 1). Here, participants were native speakers of Italian (L1), highly proficient learners of German (L2), and lower proficient learners of English (L3). For this group, the L2 and L3 were typologically closer languages, compared to the L1 and L3 or the L1 and L2. In this experiment, the results showed that language switching costs were smaller for the weaker L3 compared to the stronger L1 and L2. This pattern supports the hypothesis postulated by the IC model that weaker languages are more easily reactivated than stronger languages. However, the comparable amount of switching costs found for the stronger L1 and the weaker L2 of this group suggests that this is not always the case. In order to understand why the amount of inhibition applied to the stronger L1 and the weaker L2 was comparable, errors in each language were explored. The error analysis revealed that the L2 was the language which mostly suffered from unwanted cross-language interference and that interference mainly came from the weaker L3. In particular, it was suggested that, because of its typological closeness, the stronger L2 (German) might have greatly interfered during L3 (English) processing. In order to reduce disturbances from the stronger and typologically closer L2 into the L3, the L2 is “hampered” by the controlling system. This can be achieved by making the L2 less available, that is, more difficult to reactivate. This hypothesis might explain why the L2 is the language in which speakers made more errors, why those errors were coming especially from the weaker L3 and why the amount of inhibition applied to the L2 was greater than predicted by a dominance-related account of language inhibition. Thus, the greater inhibition of the L2 would explain why switching costs for the stronger L1 and the weaker L2 were comparable. Similar to what was discussed above, in this experiment as well language typology seemed to have a clear effect on language switching patterns, while its influence was less straightforward on mixing costs. Here, language mixing costs were larger for the stronger L1 compared to the weaker L2 and L3. The amount of mixing costs did not differ between the stronger L2 and the weaker L3. This pattern was interpreted as a strategy adopted to facilitate naming in the conflicting weaker languages (L2 German – L3 English) by enhancing their activation level (smaller mixing costs).

Overall, the results from Publication III indicate that language typology is an influential factor in language control. However, in order to be confident that the effects can be attributed to language typology only and not to biases in the type of stimuli used (e.g., a language pair sharing more similar item words than another language pair), it is important to make sure that the words used to name the pictures were phonologically equally distant across the three languages. To do this, the phonological edit distance between words was measured using the software Phonological Corpus Tool (PCT) version 1.1 (Currie, Allen, Fry, Mackie & McAuliffe, 2015). The PCT calculates the number of one-symbol changes between words and assigns to each change a weight depending on feature similarity. For example, the dental stop consonants /t/ and /d/ differ from each other in that the first is voiceless (-voice), whilst the latter is voiced (+voice). In this case, the difference between /t/ and /d/ is of one feature change (from -voice to +voice). However, the difference between the stop consonants /t/ and /b/ includes three feature changes, namely one for voicing (/t/= -voice; b= +voice) and two for the place of articulation (/t/ = +coronal and -labial; /b/ = -coronal and +labial). The analysis of the feature changes revealed that the degree of phonological overlap between words was equal across the three language pairs (p> .05 for Italian-German vs. Italian-English, p> .05 for Italian-German vs. German-English and p> .05 for Italian-English vs. German-English). This indicates that the effect found in Publication III can be ascribed to the typological closeness of the languages involved in the switching context and not to the fact that words in a language pair were more similar compared to words in another language pair. Table 2 illustrates the phonological transcription of the words used in the three languages.

Table 2. List of the words used in Publication III and their phonological transcription (IPA) for Italian, German and English.

ITEM	ITALIAN	GERMAN	ENGLISH
“arrow”	/f.r.e.tʃ.a/	/pf.aɪ.l/	/æ.r.əʊ/
“bell”	/k.a.m.p.a.n.a/	/g.l.ɔ.k.ə	/b.ɛ.l/
“belt”	/tʃ.i.n.t.u.r.a/	/g.ʏ.r.t.ə.l/	/b.ɛ.l.t/
“broom”	/s.k.o.p.a/	/b.ə.z.ə.n/	/b.r.u:.m/
“butterfly”	/f.a.r.f.a.l.l.a/	/ʃ.m.ɛ.t.ɐ.l.ɪ.ŋ/	/b.ə.t.ə.r.f.l.a.j/
“chair”	/s.e.d.j.a/	/ʃ.t.u:.l/	/tʃ.e.ə/
“door”	/p.o.r.t.a/	/t.y:.r/	/d.ɔ:/
“dress”	/v.e.s.t.i.t.o/	/k.l.aɪ.t/	/d.r.ɛ.s/
“grapes”	/u.v.a/	/v.aɪ.n.t.r.aʊ.b.ə/	/g.r.æ.p.s/
“leaf”	/f.o.ʎ.ʎ.a/	/b.l.a.t/	/l.i:.f/

“mushroom”	/f.u.ŋ.g.o/	/p.ɪ.l.ts/	/m.ʌ.ʃ.r.ʊ.m/
“necklace”	/k.o.l.l.a.n.a/	/k.ɛ.t.ə/	/n.e.k.l.eɪ.s/
“onion”	/tʃ.i.p.o.l.l.a/	/ts.v.iː.b.ə.l/	/ʌ.n.ɪə.n/
“pumpkin”	/ts.u.k.k.a/	/k.yː.r.b.ɪ.s/	/p.ə.m.p.k.ə.n/
“spoon”	/k.u.k.k.j.a.j.o/	/l.œ.f.ə.l/	/s.p.uː.n/
“tree”	/a.l.b.e.r.o/	/b.aʊ.m/	/t.r.iː/
“watch”	/o.r.o.l.o.dʒ.o/	/uː.r/	/w.ɒ.tʃ/
“wheel”	/r.w.o.t.a/	/r.aː.t/	/w.iː.l/

Importantly, Publication III not only shows that language control is affected by the type of languages (more vs. less similar) involved in the switching context, but it also gives us important suggestions as to how multiple languages might be represented within the mental lexicon. In particular, as mentioned in Publication III, the idea is that languages are represented in subsets and that the degree of overlap between language subsets depends on the number of shared features (de Bot, 2006). This notion recalls Paradis’ (1987) observation, according to which languages that have less in common are represented more separately than languages that share more features. For example, languages with more cognates (words with similar form and meaning in two languages) will overlap more than languages with fewer cognates. Moreover, when a specific language subset is activated, depending on the degree of overlap, activation might spread to other language subsets. Hence, as proposed by Albert and Obler (1978), more similar languages may be more difficult to keep apart than less similar languages. Altogether, results from Publication III support the hypothesis that lexically closer languages (or languages with a higher number of cognates) interfere more with each other compared to lexically more distant languages (or languages with fewer cognates). Crucially, the study also reveals that the controlling system is able to strategically prevent conflicts between more similar languages. Precisely, while in one case (Exp. 1) the disturbing language is hampered more in order to help process the “unstable” language, in the other case (Exp. 2) naming in the unstable language is facilitated so as to reduce interference from the disturbing language.

In Publications I and IV, language control was explored in bilingual speakers. Participants were native speakers of German (L1) and highly proficient learners of English (L2) in the former, and native speakers of Dutch (L1) and highly proficient speakers of English (L2) in the latter. Thus, in both studies the native and the non-native language were closely related languages with regard to lexical overlap (high number of cognates between languages). Concerning production, both studies showed that in the mixed-language condition, performance in the weaker L2 was faster overall compared to the stronger L1. This paradoxical language effect was explained with the hypothesis put forward by previous studies that in a mixed-language context, speakers might tend to help the weaker language by making it more available than the stronger one (Christoffels et al., 2007; Costa & Santesteban, 2004; Costa et al., 2006; Peeters et al., 2014; Martin et al., 2013; Schwieter & Sunderman, 2008; Verhoef et al., 2009; Verhoef, Roelofs & Chwilla, 2010). Based on what was discussed above, one might argue that the paradoxical language effect found in Publications I and IV was triggered by the fact that the weaker L2 (English) needed to be "protected" by the interference coming from the stronger and more closely related L1 (German or Dutch). However, results from previous research reveal that the L1 disadvantage in the mixed-language condition can also be found in bilinguals with less closely related languages (e.g., L1 English – L2 Spanish or vice versa, Kleinman & Gollan, 2016; Schwieter & Sunderman, 2008) and with unrelated languages (e.g., L1 Spanish – L2 Basque, Costa et al., 2006; L1 Chinese – L2 English, Jin, Zhang & Li, 2014; Liu, Fan, Rossi, Yao & Chen, 2016; Slevc, Davey & Linck, 2016). Therefore, the facilitation effect found in the L2 of the bilinguals tested in Publications I and IV can hardly be attributed to language typology. This hypothesis supports Costa et al.'s (2006) assumption that language control in bilinguals is not visibly affected by language similarity. More generally, the comparison between bilinguals' (Publications I & IV) and trilinguals' (Publications II & III) performance seems to suggest that language control in these two groups might rely on different processes. Specifically, while bilinguals showed a reversed language pattern (faster naming in the weaker L2 than in the stronger L1), this was not the case for trilinguals, where responses were faster in the stronger L1 and slower in the weaker L3. The assumption that switching between two languages might involve different controlling mechanisms than switching across three languages is supported by the fact that, to the best of my knowledge, there is no evidence in the language switching literature of a reversed language pattern (faster responses for the weaker than for the stronger language) in trilinguals. One intriguing case, however, is represented by de Bruin et al.'s (2014) study. In this picture naming study, a group of late unbalanced trilinguals showed faster responses in the stronger language (L1) compared to the weaker non-native languages (L2 and L3), and no reliable difference between the stronger and weaker non-native languages when pictures had to be named in one language at a time (single-language block). However, when all three languages were relevant for the task (mixed-language block), the dominance effect disappeared and pictures were named equally fast in the three languages. Interestingly, a closer look at the raw data reveals that, in the mixed-language block, responses tended to be faster in the weaker L3 (1535ms), followed by the intermediate L2 (1560ms) and then by the stronger L1 (1589ms). Despite this interesting pattern, it remains an open question whether trilinguals can show a paradoxical language effect just as bilinguals do, and therefore whether bilinguals and trilinguals rely on the same or different controlling mechanisms during language switching.

Understanding the paradoxical language effect found in bilinguals is of fundamental importance, considering that, according to the IC model, language switching costs are supposed to be modulated by languages' activation level. However, both Publications I and IV challenge this hypothesis by showing that switching costs in the more highly activated L2 (faster responses) and in the less activated L1 (slower responses) were similar. If language switching costs depend on languages' activation level as postulated by the IC model, we should have found larger switching costs for the more activated L2 compared to the less activated L1. In order to understand whether language switching costs might be related to languages' activation level, it is crucial to emphasize the fact that language control is composed of two main mechanisms, namely transient control (reflected by language switching costs), and sustained control (reflected by language activation level or language mixing costs). This complementarity is very important, since all the proposed accounts of language control have primarily focussed on transient, but fewer on sustained control (e.g., Costa et al., 1999; Costa & Caramazza, 1999; Costa & Santesteban, 2004; Finkbeiner et al., 2006; Green, 1986, 1993, 1998; La Heij, 2005; Verhoef et al., 2009). In Publication IV, we proposed that transient and sustained control of language might be two independent, but interconnected, processes. On the one hand, they are independent in the sense that the amount of transient control (reflected by switching costs) is not directly related to the amount of sustained control (indexed by mixing costs). For instance, in the two bilingual studies reported in this work, we found larger mixing costs for the stronger L1 than the weaker L2, but the same amount of switching costs for the two languages. On the other hand, they might be interconnected in the sense that the amount of inhibition exerted by the transient control might at first depend on language dominance and then slowly adapt to the language strength of activation as regulated by the sustained control. In our bilingual case, this would imply that language switching costs are at first larger for the stronger L1 than for the weaker L2 (irrespective of languages' actual activation level), and afterwards larger for the weaker L2 than for the stronger L1 (in line with their actual activation levels). A fluctuation over time of the transient control might explain why overall switching costs for the stronger and weaker languages in the bilingual studies (Publications I & IV) reporting paradoxical language effects turned out to be symmetrical.

The idea that transient control gradually adapts to sustained control resembles Suzuki and Shinoda's (2015) proposal that, during cognitive control, conflict resolution is associated with a transition from transient to sustained control. Furthermore, the observation that transient and sustained control are interdependent mechanisms is not novel in the cognitive control literature. In particular, Braver (2012) suggests that both costs and benefits are associated with transient and sustained control so that, computationally, a trade-off between the two mechanisms is required to optimize information processing. In this regard, a closer look at Publication III appears to pinpoint this kind of balance between the two processes. Precisely, in the first group of trilinguals tested in this study (Exp. 1), switching costs (reflecting transient control) for the L2 were larger than predicted by a dominance-related account of language inhibition, while mixing costs (indexing sustained control) were lower compared to the stronger L1 and comparable with those of the weaker L3. In contrast to Publication II, where the amount of sustained control seemed to be modulated by language dominance, here L2's mixing costs were smaller than expected based on its dominance. Moreover, in the second group of trilinguals tested in this study (Exp. 2), the L2 revealed smaller switching costs compared to those predicted by an inhibitory dominance-related account, while L2's mixing costs were greater (even though only numerically) than expected based on its dominance. These results suggest that transient and sustained control are "conversely linked", in the sense that when one increases the other decreases. To be more specific, based on what was discussed for the bilingual studies presented in this dissertation, one might argue that it is the level of sustained control (reflected by languages' activation level or mixing costs) that determines the amount of transient control (indexed by switching costs). This suggestion would extend language control to Braver's (2012) observation that there might be a complementary connection between transient and sustained control.

Alternatively, it could be argued that transient and sustained control are entirely dissociated mechanisms and the amount of transient control is solely based on the complexity of the task to be performed. Specifically, as observed by Yeung and Monsell (2003), switching between tasks is an effortful process, so that the controlling system aims at minimizing the effort when possible. This can be achieved by applying the same amount of inhibition to tasks that do not differ too much in their complexity. In this way, greater inhibition of the stronger than the weaker task would be an effective strategy only when the difference in complexity between the two tasks reaches a certain threshold. Therefore, given the relatively small dominance difference between the stronger L1 and weaker L2 of the bilingual speakers tested in this dissertation, the inhibition applied to these two languages might have been the same, leading to L1-L2 symmetrical switching costs. Nevertheless, while this explanation might be valid to explain language switching patterns in bilinguals, it would be difficult to apply to trilingual language control. Indeed, if this was the case, in the trilingual studies reported in this work, we should have found larger switching costs for the stronger than the weaker language in the case of conspicuous dominance difference (i.e., in the L1-L3 language pairing), but the same amount of switching costs when the dominance difference was less prominent (i.e., in the L1-L2 and the L2-L3 language pairings). However, as thoroughly discussed above, this was not the case.

Finally, the assumption that bilingual and trilingual language switching might rely on different strategies of language control is supported by the performance comparison of the bilinguals tested in

Publication I and the trilinguals tested in Exp. 2 of Publication III. In the former, participants were native speakers of German (L1) and highly proficient speakers of English (L2). In the latter, participants were native speakers of German (L1), highly proficient speakers of English (L2), and lower proficient speakers of Italian (L3). The same experimental procedure and stimuli were used in the two experiments. Concerning the L1-L2 language pairing, results revealed similar performance in the two groups with regard to the single-language-block, namely faster responses for the stronger L1 (German) than for the weaker L2 (English). However, the pattern in the two experiments dramatically changed in the mixed-language block. In particular, while in the bilingual group both naming latencies and switching costs in L1 and L2 were comparable, in the trilingual group L1 responded faster than L2 and switching costs were larger for L1 relative to L2. These results might indicate two things. The first is that language control in bilinguals and trilinguals relies on different processes. Therefore, despite the same experimental procedure, the same type of languages, and the comparable dominance levels of those languages, language control in bilinguals and trilinguals is not comparable. The second is that language control might be affected by the number and type of languages involved in the context. More precisely, the controlling system might allocate resources differently depending on the interactional context. While in a dual-language setting support is sent to the two languages relevant to the task, in a trilingual context, resources have to be distributed across three languages (see Publication III for further discussion). In this case, bilinguals and trilinguals might perform similarly when only two languages are involved in the interactional context, but differently when trilinguals are confronted with an additional language.

Importantly, the fact that resources can be allocated differently in accordance with the interactional context might make it difficult to detect more subtle processing differences between native and non-native languages. In fact, while in the bilingual setting a reasonable amount of cognitive support can be sent to the weaker non-native language to the point that it can become as or more available than the stronger native language (paradoxical language effect), in the trilingual context resources have to be distributed across three languages, so that qualitatively differences between native and non-native languages might become evident. This view might explain why in the bilingual setting the amount of switching costs between the stronger native and the weaker non-native language remained equivalent (Publications I & IV), whilst it was likely to differ in the trilingual context. In this latter case, two groups of trilinguals showed larger switching costs for the native L1 compared to both non-native L2 and L3 (Publication II and Publication III, Exp. 2), whilst only one group of trilinguals revealed larger switching costs for the native L1 and the weaker non-native L3, but a similar amount of switching costs for the native L1 and the non-native L2 (Publication III, Exp. 1). However, because of the tight correlation between language dominance and language status (the native language is usually the stronger language), as well as of the different types of languages included in the tasks, it is difficult to tease apart these factors and clearly identify to what extent language status might have affected language control in these two different interactional settings.

Comparing language control in the bilingual and trilingual settings not only reveals that the language switching pattern differs between the two conditions, but it also serves to verify the assumptions made by the alternative views of language control. As an extension to the IC model in production, the language-specific account of language selection was considered in the present dissertation. As described in the introduction, Costa and colleagues (2004, 2006) proposed that the way languages are controlled depends on the dominance level reached in the weaker language (e.g., the L2). In particular, they suggested that when the proficiency level in the weaker L2 is relatively low, both the relevant and the irrelevant language compete for selection during language switching and that selection of the relevant language is achieved by suppressing the non-relevant language. As for the IC model, the amount of inhibition applied to the non-relevant language depends on its dominance. However, the authors extended the predictions made by the IC model by proposing that when speakers reach a relatively high level in their L2, the way languages are controlled would not rely on inhibition anymore. This is because, in the case of more advanced L2 speakers, only words from the relevant language will be considered for selection, so that there will be no competing language to be suppressed. In this case, language switching costs would reflect the effort of reconfiguring the new language. In contrast to the cost required to recover from previous inhibition, the effort to reconfigure a language is not believed to be affected by language dominance. Therefore, in more advanced L2 speakers the amount of switching costs in the two languages is predicted to be the same. Briefly, according to Costa and his team, language selection is *language non-specific* (competition between relevant and irrelevant languages) in less advanced L2 speakers and is reflected by larger costs for the stronger than for the weaker language or “asymmetrical switching costs”, while language selection is *language-specific* (only the relevant language is consulted for selection) in more advanced L2 speakers and is indexed by a similar amount of switching costs for the two languages or “symmetrical switching costs”. Crucially, the authors argued that once the language selection process has shifted from a language non-specific to a language-specific mechanism, the latter will be used also in other cases of language switching (e.g., when switching between a stronger L2 and a weaker L3).

Results from the two bilingual studies reported in this thesis showed symmetrical switching costs between the stronger L1 and the weaker L2 in advanced L2 speakers. This pattern can hardly be explained by a dominance-related inhibitory account (predicting larger costs for the stronger than for the weaker language), but can be accounted for by a language-specific mechanism (expecting symmetrical switching in advanced L2 speakers). However, as seen above, there might be other reasons why switching costs between a stronger and a weaker language in bilinguals are symmetrical, without requiring a change in the selection mechanism. In this respect, results from the trilingual experiments described in this dissertation show that both symmetrical and asymmetrical switching costs can be found within a group of advanced L2 speakers, rejecting the hypothesis that symmetrical switching costs reflect the permanent shift to a language-specific mechanism. Moreover, the analysis of the errors reported in Publication III revealed that irrespective of the language switching pattern (symmetrical or asymmetrical), all three languages were subject to cross-language interference. This implies that language control is a language non-specific mechanism, in that also words from irrelevant languages compete for selection.

7.4 Processing Modality

The third and last issue explored within this dissertation is the role of processing modality (production vs. recognition) in language control. Specifically, it was asked whether or not the way languages are controlled during production and recognition is the same. To answer this question, the performance of a group of bilingual speakers was compared in both production and recognition tasks (Publication IV). The main results indicate that language control is profoundly affected by processing modality, and that language dominance has an effect on bilingual recognition, but less so on bilingual production.

According to the IC model, language control in production and recognition is supported by the same type of mechanisms. Consider, for example, a lexical decision task in which a speaker has to decide by pressing a key whether or not a presented word belongs to a specific language (e.g., if written in blue, decide whether or not the word is an L1 word). In order to perform this task, a task schema has to be created which connects a language cue (e.g., blue ink colour) and a stimulus (e.g., L1 word) to a response (e.g., press “YES” key). When the language cue (e.g., blue ink colour) and the matching word (e.g., L1) are presented, the relevant task schema is activated (e.g., press “YES” key) and the task can be executed. One of the main functions of the task schema is to alter the activation level of the representations in the bilingual mental lexicon. This is achieved by activating words belonging to the relevant language (e.g., L1) and inhibiting words from the non-relevant language (e.g., L2). On a switch trial, the new task schema has to be triggered and the old one has to be suppressed. In this framework, inhibition is exerted on two loci: at the schema level (outside the mental lexicon) and at the lemma level (inside the mental lexicon). Since inhibition is believed to be reactive, more active lemmas (e.g., L1 lemmas) will be inhibited more than less active ones (e.g., L2 lemmas). Because the cost required to reactivate a language depends on the amount of suppression applied to it, switching costs for a more highly activated language (e.g., the L1) are predicted to be larger compared to a less activated language (e.g., the L2). In contrast to this, the inhibition applied outside the mental lexicon (at the task schema level) is supposed to be the same for the two languages. The same is held for production. Imagine, for example, a bilingual picture naming task in which speakers are requested to name a presented picture in the language indicated by a language cue (e.g., name in L1 if the background is blue). Like the recognition task, in the production task task schemas for the L1 and the L2 also have to be created. The triggered task schema (e.g., name in L1) regulates the activation level within the mental lexicon so that words belonging to the relevant language (e.g., the L1) will be more highly activated and words of the non-relevant language (e.g., the L2) will be suppressed. Again, the amount of suppression is higher for more active than for less active words (e.g., L1 and L2 words, respectively), so that reactivating more strongly inhibited words will be more costly compared to less strongly inhibited words.

The bilingual study described in Publication IV examines the assumptions made by the IC model by measuring language switching costs in unbalanced bilinguals performing a recognition (Exp. 1) and a production (Exp. 2) task. In the recognition task, we found larger switching costs for the stronger L1 than for the weaker L2, as predicted by the IC model. However, while this result reinforces the predictions made by the model, it also questions the reliability of the detailed mechanisms described in the model. In particular, according to the IC model, words' activation level is modulated by the task schema, so that a triggered task schema can enhance the activation level of words belonging to the relevant language and inhibit words from the non-relevant language. On a switch trial, reactivating words of the previously inhibited language will lead to language switching costs. The task used in this experiment was a "generalized" lexical decision task, where participants had to decide whether a string of letters was an existing word irrespective of language membership (for the same kind of task, see also Grainger & Beauvillain, 1987; Thomas & Allport, 2000; von Studnitz & Green, 1997). In this case, both languages are linked to a single task schema (e.g., in the case of an existing word, press right key). Given the absence of L1-L2 task schemas and so of L1-L2 lemma competition, no language switching costs are expected in this case according to the IC model. However, as observed by Studnitz and Green (1997), who also obtained language switching costs in a generalized bilingual lexical decision task, it might be that the switching pattern depends not only on the nature of the decision task but also on the lexical properties of the stimuli. Similarly, Thomas and Allport (2000) argued that language switching costs in a generalized lexical decision task might indicate practice effects due to the nature of the stimuli. More precisely, the authors suggested that bilinguals might usually treat the two language systems differently, so that even when a distinction between the two systems is not required (like in the generalized lexical decision task), bilinguals cannot help considering them differently, leading to language switching costs. However, the way in which the properties of the stimuli affect their processing has not been fully specified. More explicitly, in this sense, is the Bilingual Interactive Activation (BIA) model of visual word recognition (Dijkstra & van Heuven, 1998; Grainger & Dijkstra, 1992; van Heuven et al., 1998). According to the BIA model, when a visual word is presented to a bilingual speaker, similar words from both the relevant and the non-relevant language will be activated and compete for selection. The activated words will send activation to their respective language nodes (representational layers in which language membership is specified). The main function of the language node is to collect activation sent by the words of that language and inhibit words from the other language. The amount of inhibition depends on the activation strength of the language node. This means that the more a language node is activated, the stronger is the inhibition applied to the words of the other language.

As with the IC model, in the BIA model, more frequent words (e.g., L1 words) also have a higher resting level of activation than less frequent words (e.g., L2 words). In the latter model, this implies that L1 words activate faster than L2 words so that the L1 language node will collect a greater amount of activation than the L2 language node. It follows that the amount of inhibition exerted on the words of the other language will be greater for the more activated L1 than for the less activated L2 language node. Overall, this indicates that within the BIA model language control is regulated by the stimulus itself (e.g., words' baseline activation level) and not by an external process (e.g., a task schema). This hypothesis provides a reasonable explanation for the source of language switching costs in situations where language membership is not relevant (e.g., language-inclusive lexical decision tasks). Nevertheless, problematic within the BIA model is how inhibition is accomplished. As remarked earlier, the amount of inhibition exerted by the language node of a stronger language on the words belonging to a weaker language is greater than vice versa. This means that the cost of reactivating words from a weaker L2 is expected to be larger compared to words from the stronger L1, i.e. *dominance-reversed* switching costs. This is, however, not what we found in the study reported in Publication IV, and not what previous studies on bilingual visual word recognition have shown (Jackson et al., 2004; Macizo et al., 2012; Orfanidou & Sumner, 2005; Thomas & Allport, 2000; von Studnitz & Green, 1997, 2002).

The larger switching costs for the stronger L1 than for the weaker L2 found in Publication IV were explained with the hypothesis that when a word is presented visually, words belonging to both languages might compete for selection and that selection is achieved by suppressing the non-relevant language. However, differently from the BIA model and in accordance with the IC model, inhibition is supposed to be reactive, namely greater on more active than on less active words. Therefore, the cost required to reactivate words of a stronger language will be larger compared to words of the weaker language, i.e. *dominance-related* switching costs. Overall, this hypothesis suggests that inhibition at the lemma level can be driven by the stimulus properties as indicated by the BIA model, and that the strength of inhibition is greater for the stronger than for the weaker language as suggested by the IC model.

As for the production task, we found a reversed dominance pattern (weaker L2 faster than stronger L1) together with a comparable amount of switching costs for the stronger and weaker languages. As elaborated in the previous section, this pattern is in contrast to what was predicted by the IC model, according to which language switching costs depend on languages' activation level.

Overall, the data obtained in Publication IV clearly show that the way in which languages are controlled is severely affected by processing modality. This finding challenges the implications of the IC model, according to which language control in recognition and production are supported by the same controlling mechanisms. In particular, in Publication IV, we highlighted the fact that different processes are involved between lexical decision and picture naming tasks. While in a lexical decision task, information about the stimulus is mainly spread from the bottom up (from letters up to words, and from words up to concepts), during picture naming information mainly flows in a top-down fashion (from the concept down to word retrieval, and from there down to its phonological encoding and then articulation). Importantly, bottom-up and top-down control of attention are believed to be two distinct types of mechanisms, with the former activated by the stimulus features and the latter determined by the subject's intentions (e.g., Corbetta & Shulman, 2002; Jonides, 1981; Posner, 1980). Therefore, bottom-up or "exogenous" control is considered an involuntary mechanism independent of the subject's intentions, whilst top-down or "endogenous" control is supposed to be voluntarily driven (e.g., Carrasco, 2011; Jonides, 1981). With respect to the two experiments reported above, we might say that during lexical decision tasks non-intentional processes are activated (since participants cannot avoid recognizing a word), whereas in the picture naming task voluntary actions are required (the intention to name a picture). Moreover, a very important difference between these two processes concerns their functionality. Indeed, whereas top-down control is believed to be flexible, bottom-up control is supposed to be a more stable mechanism (e.g., Jonides, 1981; Pinto, 2013; Giordano, McElree & Carrasco, 2009). More specifically, the idea is that top-down control can allocate resources differently according to the situation (e.g., Coull & Nobre, 1998), while bottom-up control lacks this flexibility, so that even when irrelevant information is presented it cannot be ignored (e.g., Pestilli & Carrasco, 2005; Pestilli, Viera & Carrasco, 2007). Interestingly enough, the diversity between these two controlling systems might be ascribed to the fact that they are, at least partially, mediated by different neural structures (e.g., Kim, Gitelman, Nobre, Parrish, LaBar & Mesulam, 1999; Rafal & Henik, 1994; Robinson & Kertzman, 1995, but see Peelen, Heslenfeld & Theeuwes, 2004).

Overall, with regard to Publication IV, it can be argued that during visual word recognition (Exp. 1), mainly supported by a bottom-up mechanism, language control was affected by the properties of the stimuli. In particular, although language membership of the words was irrelevant for the task, words belonging to the stronger L1 were activated more than words of the weaker L2, leading to larger switching costs for the L1 relative to the L2. In contrast, during picture naming (Exp. 2), mainly supported by a top-down process, resources could be allocated strategically. Specifically, in order to boost naming in the weaker language, the stronger L1 became slower than the weaker L2 (reversed dominance pattern) and switching costs for the stronger and the weaker language were symmetrical (either reflecting a gradual adaptation of the amount of inhibition to the current language activation level or as a way to minimize the effort when switching between two tasks). More generally, the data presented in this dissertation seem to support the hypothesis that language control in production is supported by flexible and strategic mechanisms aimed at optimizing performance. This was observable in the bilingual picture naming studies (Publication I and Publication IV, Exp. 2), which reported facilitation of the weaker compared to the stronger language together with a similar amount of switching costs for the two languages, and in the trilingual studies, where data showed that in more demanding tasks weaker languages were hampered less than stronger languages (Publication II) and that conflicts between typologically closer languages were reduced by altering languages' inhibition strength (Publication III). In contrast, in the bilingual lexical decision experiment (Publication IV, Exp. 1), the controlling system did not seem to rely on noticeable strategies to perform the task. In fact, here bilinguals' naming performance appears to be exclusively modulated by language dominance.

8 Conclusions and Future Perspectives

The present dissertation investigated the mechanisms underpinning language control in multilingual speakers. Specifically, it explored whether the way in which languages are controlled in a language switching context is determined by language dominance as asserted by one of the prevailing models of language control, namely the IC model (Green, 1998), or whether other factors might play a role in this process. Most of the previous studies examining language control have (i) primarily concentrated on the transient control of languages, as reflected by language switching costs, (ii) relied on language switching settings where only two languages at a time could be tested, and (iii) predominantly focussed on production. The studies presented in this dissertation extend existing research by (i) exploring both transient and sustained control of languages, (ii) using language switching settings in which two or more languages at a time are simultaneously involved, and (iii) comparing language control in recognition and and production.

With regard to the influencing factors involved in language control, three main issues were addressed in this work. The first two studies looked at the effect of *preparation time* on bilingual (Publication I) and trilingual (Publication II) language control. These were followed by a study (Publication III) focussing on the role of *language typology* on trilingual language control. Finally, the last study (Publication IV) aimed at assessing whether the way languages are controlled is affected by *processing modality*. The main results revealed that language control is profoundly affected by these three variables. Concerning preparation time, data showed that when (bilingual and trilingual) speakers were given no or less time to prepare for a specific language, language switching yielded processing costs. However, when the same speakers were provided with more preparation time, language switching became a cost-free process. This finding was taken as clear evidence that the cost required to reactivate a language is primarily modulated by preparation time, rather than by language dominance. In relation to language typology, results showed that during language switching, typologically closer languages interfered with each other to a greater extent than typologically more distant languages, so that conflicts between languages were avoided by hampering the disturbing language and/or by facilitating the disadvantaged language. In this respect, it was suggested that language control can be affected by language typology and, therefore, is not solely determined by language dominance. As for processing modality, data showed that while language control in visual recognition was modulated by language dominance, this was not the case for language production. This finding suggests that the way languages are controlled depends on the processing modality, which is to say that languages are controlled differently between production and recognition.

Altogether, the present work proposes that, in production, language control is a flexible and strategic mechanism able to allocate resources differently depending on the context, so as to reduce major conflicts between languages. Moreover, it suggests that, in contrast to language control in production, language control in visual word recognition is a less malleable process, in that it is principally affected by the properties of the stimuli and less so by speakers' unconscious strategies for optimizing performance. More generally, the findings presented in this dissertation extend our knowledge about language control in bilinguals and trilinguals by showing that language control is a considerably more dynamic system than previously postulated. In particular, the work emphasizes the urge to extend existing models of language control to include all the other factors influencing this process. For example, if language control in production and recognition is supported by different mechanisms, future models of bilingual language processing should aim at integrating those differences in a unique account. A first step in this direction is exemplified by "Multilink", a computational localist-connectionist model for word translation implemented by Dijkstra and Rekké (2010). Since word translation includes aspects of both word recognition and production, Multilink aspires to provide a simulation of their basic mechanisms in bilinguals. The model shares similarities with several existing models of bilingual language processing. For example, similar to the BIA and the BIA+ model (Dijkstra & Heuven, 2002) activation in Multilink is interactive (activation spreads from one element to the other). Like the BIA+ and IC models, it comprises a task/decision system. Finally, along with the Revised Hierarchical Model (RHM; Kroll and Stewart, 1994), it shares the assumptions that the L1 and L2 lexicons might differ in size (the more dominant language usually has a larger lexicon compared to the less dominant language) and that the link between a word's form and its meaning might be different for the L1 and L2. Although Multilink has shown some potential in trying to simulate human behaviour in lexical decision and language decision tasks, it is still an unrefined model lacking implementations with regard to language control. For instance, in order to simulate results from a lexical decision task where bilinguals were requested to decide whether or not a visually presented word belongs to a particular language, only non-words and words from a single language (the relevant language) were presented to Multilink (Rekké, 2009). With regard to production, the processes related to the phonological output have not been implemented in Multilink (but, see e.g., van Halem, 2016; Peacock, 2015 for unpublished work on Multilink simulation of word translation). Hence, future work is needed to provide a unified model of language control comprising both language production and recognition.

Despite providing numerous insights into the mechanisms underpinning language control in multilinguals, the present work is also limited in other respects. For example, with regard to production, while it broadens our understanding of "forced" language switching, where the language to be used is indicated by a language cue, this work does not pertain to spontaneous language switching, where speakers are free to choose which language to use. As discussed by Green and

Abutalebi (2013), different interactional contexts might involve different controlling processes. In particular, when the language choice is not free (e.g., use Languages A and B with persons A and B, respectively), language switching is supposed to be a competitive process, in which competition from the non-relevant language needs to be solved. In contrast, when language switching is voluntary (e.g., in bilingual communities), language control is believed to be a cooperative process involving opportunistic planning, that is, speakers make use of whichever language is easier to retrieve. Overall, authors suggested that the language control network, together with its controlling processes, adapts to the specific interactional context (for a similar argument, see also Green & Wei, 2014). Considering this, it is plausible to assume that language control in a voluntary language switching situation might require fewer strategic mechanisms in order to avoid conflicts between languages compared to a non-spontaneous language switching context, where languages are supposed to compete. A recent study showed, indeed, that in a quasi-voluntary switching task bilinguals can rely on lexical accessibility to switch cost-free between languages (Kleinman & Gollan, 2016). Additionally, it could be asked whether there are substantial differences between language control in early multilinguals (who have acquired their languages simultaneously during early childhood) and late multilinguals (who have acquired a native language first and the additional language(s) subsequently). In fact, if it is the case that language control, at least in production, aims at avoiding interference from a stronger disturbing language into a weaker one, then the mechanisms supporting language switching should differ between these two populations. Some support for this hypothesis can be found in a single study showing that early bilingual learners of an L3 might not control languages in the same way as late bilingual learners of an L3 (Martin et al., 2013). Additionally, two distinct studies revealed different neural activations for early (Garbin, Costa, Sanjuan, Forn, Rodríguez-Pujadas, Ventura, Belloch, Hernandez & Ávila, 2011) and late bilinguals (Wang et al., 2007) when switching between languages. However, given the absence of more systematic comparisons between early and late multilinguals, further research is needed to understand whether and to what extent the language age of acquisition affects the way languages are controlled.

With regard to language control in recognition, the present work suggests that this process might mainly be driven by the properties of the stimuli (e.g., their frequency) and less so by speakers' strategic mechanisms to reduce competition between languages. However, it is important to highlight that this observation is based on data obtained from a single experiment and that more studies are necessary to reinforce this hypothesis. In this respect, a more subtle way to determine whether language control in production and recognition relies on different processes would be to look at speakers' neurophysiological behaviour as revealed, for example, by event-related potentials (ERPs).

REFERENCES

Abutalebi, J., Annoni, J. M., Seghier, M., Zimine, I., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., & Khateb, A. (2008). Language control and lexical competition in bilinguals: an event-related fMRI study. *Cerebral Cortex, 18,* 1496–1505. doi: 10.1093/cercor/bhm182

Abutalebi, J., Brambati, S. M., Annoni, J. M., Moro, A., Cappa, S. F., & Perani D. (2007). The neural cost of the auditory perception of language switches: an event-related functional magnetic resonance imaging study in bilinguals. *The Journal of Neuroscience, 27,* 13762–13769. doi: https://doi.org/10.1523/JNEUROSCI.3294-07.2007

Abutalebi, J., Della Rosa P. A., Castro Gonzaga A. K., Keim R., Costa A., & Perani D. (2013). The role of the left putamen in multilingual language production. *Brain and Language, 125*, 307–315. doi: 10.1016/j.bandl.2012.03.009

Abutalebi, J., & Green, D. W. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, *20,* 242–275. doi: http://dx.doi: 10.1016/j.jneuroling.2006.10.003

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. doi: http://dx.doi.org/10.1080/01690960801920602

Abutalebi, J. and Green, D.W. (2016). Neuroimaging of language control in bilinguals: neural adaptation and reserve. *Bilingualism: Language and Cognition, 19,* 689–698. doi: 10.1017/S1366728916000225.

Abutalebi, J., Miozzo, A., & Cappa, S. F. (2000). Do subcortical structures control language selection in polyglots? Evidence from pathological language mixing. *Neurocase, 6,* 101–106. doi: http://dx.doi.org/10.1080/13554790008402757

Aglioti, S., Beltramello, A., Girardi, F., & Fabbro, F. (1996). Neurolinguistic and follow-up study of an unusual pattern of recovery from bilingual subcortical aphasia. *Brain, 119,* 1551–1564. doi: https://doi.org/10.1093/brain/119.5.1551

Aglioti, S., & Fabbro, F. (1993). Paradoxical selective recovery in a bilingual aphasic following subcortical lesion. *Neuroreport, 4,* 1359–1362. doi: 10.1097/00001756-199309150-00019

Agresti, A. (2002). Categorical data analysis (2nd ed.). New York, NY: John Wiley & Sons. doi: 10.1002/0471249688

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In Parzen, E., Tanabe, K., and Kitagawa, G. (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer, New York.

Albert, M., & Obler, L. (1978). *The bilingual brain: Neuropsychological and neurolinguistic aspects of bilingualism.* New York: Academic Press.

Allan, D. (2004). *Oxford Placement Test 1.* Oxford: Oxford University Press.

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

Altmann, E. (2007). Cue-independent task-specific representations in task switching: Evidence from backward inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 892–899. doi: 10.1037/0278-7393.33.5.892

van Assche, E., Duyck, W., & Hartsuiker, R.J. (2012) Bilingual visual word recognition in a sentence context. *Frontiers in Psychology, 3,* 174. doi: 10.3389/fpsyg.2012.00174

Baayen, R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511801686

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi: 10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3,* 12–28. doi: 10.21500/20112084.807

Babcock, L., & Vallesi, A. (2015) Language control is not a one-size-fits-all languages process: evidence from simultaneous interpretation students and the n-2 repetition cost. *Frontiers in Psychology, 6,* 1622. doi: 10.3389/fpsyg.2015.01622

Balota, D. A., Cortese, M. J., Sergent-Marshall, S., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316. doi: http://dx.doi.org/10.1037/0096-3445.133.2.283

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,* 255–278. doi: 10.1016/j.jml.2012.11.001

Bates, D. M. (2010). lme4: Mixed-effects modeling with R. URL http://lme4.r-forge.r-project.org/book.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint.* arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using S4 classes. [http://CRAN. R-project.org/package=lme4]. R package version 1.1–7.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10,* 344–380.

Baus, C., Branzi, F. M., & Costa, A. (2015). On the mechanism and scope of language control in bilingual speech production. In J. W. Schwieter (Ed.), *The Cambridge handbook of bilingual processing* (pp. 508–526). Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781107447257.022

Beauchamp M. S., Cox R. W., & Deyoe, E. A. (1997). Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *Journal of Neurophysiology*, *78,* 516–520.

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology, 25,* 547–567. doi: 10.1080/20445911.2013.812093

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.08, URL: http://www.praat.org/

Bobb, S. C., & Wodniecka, Z. (2013). Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. *Journal of Cognitive Psychology, 25,* 568–585. doi: 10.1080/20445911.2013.792822

Borkowski, J. G., Benton, A. L., & Spreen, O. (1967). Word Fluency and Brain Damage. *Neuropsychologia, 5,* 135–140. doi: http://dx.doi.org/10.1016/0028-3932(67)90015-2

de Bot, K. (1992). A bilingual production model: Levelt's Speaking Model Adapted. *Applied Linguistics, 13,* 1–24. doi: 10.1093/applin/13.1.1

de Bot, K. (2004). The multilingual lexicon: Modelling Selection and Control. *The International Journal of Multilingualism, 1,* 17–32. doi: http://dx.doi.org/10.1080/14790710408668176

de Bot, K., & Jaensch, C. (2015). What is special about L3 processing? *Bilingualism: Language & Cognition, 18,* 130–144. doi: 10.1017/S1366728913000448

Bouvy, C. (2000). Towards the Construction of a Theory of Cross-linguistic Transfer. In Cenoz, J., & Jessner, U. (Eds), *English in Europe: The acquisition of a Third Language* (pp. 143–156). Clevedon: Multilingual Matters.

Box, G., & Cox, D. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological), 26*, 211–252.

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. *Neuropsychologia, 52,* 102–116. doi: http://dx.doi.org/10.1016/j.neuropsycholog ia.2013.09.022

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16,* 106–113. doi: 10.1016/j.tics.2011.12.010

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39,* 713–726. doi: http://dx.doi.org/10.1016/S0896-6273(03)00466-5

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. doi: 10.1016/j.neuroimage.2013.12.049 [PubMed]

Bultena, S., & Dijkstra, T. (2012). Lexical access in bilingual visual word recognition. In C.A. Chappelle, M. Gullberg & J. Williams (Eds.). *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell. doi: 10.1002/9781405198431.wbeal0685

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Results, 33,* 261–304. doi: 10.1177/0049124104268644

Calabria, M., Hernández, M., Branzi, F., & Costa, A. (2012) Qualitative differences between bilingual language control and executive control: evidence from task-switching. *Frontiers in Psychology, 2*, 1–10. doi: 10.3389/fpsyg.2011.00399

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51,* 1484–1525. doi: 10.1016/j.visres.2011.04.012

Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in L3 acquisition. In Cenoz, J., Hufeisen B., & Jessner, U. (Eds.), *Cross-linguistic Influence in Third Language Acquisition* (pp. 8–20). Clevedon: Multilingual Matters.

Cenoz, J. (2003). Cross-linguistic influence in third language acquisition: Implications for the organization of the multilingual mental lexicon. *Bulletin VALS-ASLA (Vereinigung fürangewandte Linguistik in der Schweiz), 78,* 1–11.

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

Cnaan, A., Laird, N. M., & Slasor, P. (1997). Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in Medicine, 16,* 2349–2380. doi: 10.1002/(SICI)1097-0258(19971030)16:20<2349::AID-SIM667>3.0.CO;2-E

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45,* 721–736. doi: 10.1006/jmla.2001.2793

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: LEA.

Corbetta M., & Shulman G. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3,* 201–215. doi: 10.1038/nrn755

Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition, 2,* 231–244. doi: 10.1017/S1366728999000334

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41,* 365–397. doi: http://dx.doi.org/10.1006/jmla.1999.2651

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. doi: 10.1016/j.jml.2004.02.002

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1057–1074. doi: 10.1037/0278-7393.32.5.1057

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience, 18,* 7426–7435.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58. doi: http://dx.doi.org/10.1080/016909698386582

Council of Europe (2001). Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge: Cambridge University Press.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green, D. W., & Price, C. J. (2006). Language Control in the Bilingual Brain. *Science*, *312,* 1537–1540. doi: 10.1126/science.1127761

Currie, H., C., Allen, B., Fry, M., Mackie, S., & McAuliffe, M. (2015). Phonological CorpusTools, Version 1.1. [Computer program]. Available from PCT GitHub page.

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language, 49,* 119–132. doi: http://dx.doi.org/10.1016/S0749-596X(03)00008-1

Declerck, M., Koch, I., & Philipp, A. M. (2013). Digits vs. Pictures: The influence of stimulus type on language switching. *Bilingualism: Language and Cognition, 15,* 896–904. doi: https://doi.org/10.1017/S1366728912000193

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*, 377–394. doi:10.1037/xlm0000021

Declerck, M., & Philipp, A. M. (in press). Is there lemma-based language control? The influence of language practice and language-specific item practice on asymmetrical switch costs. *Language, Cognition and Neuroscience*. doi:10.1080/23273798.2016.1250928

Declerck, M., Thoma, A. M., Koch, I., & Philipp, A. M. (2015). Highly proficient bilinguals implement inhibition - Evidence from n-2 language repetition costs when switching between three languages. Journal of Experimental Psychology: Learning Memory and Cognition. doi: http://dx.doi.org/10.1037/xlm0000138

Dell, G. S. (1986). A spreading activation theory of retrieval in language production. *Psychological Review, 93,* 283–321. doi: http://dx.doi.org/10.1037/0033-295X.93.3.283

van der Meij, M., Cuetos, F., Carreiras, M., & Barber, H. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology, 48,* 44–54. doi: 10.1111/j.1469-8986.2010.01039.x

Dijkstra, A., Grainger, J., & van Heuven, W. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41,* 496–518. doi: 10.1006/jmla.1999.2654

Dijkstra, A., & van Hell, J. G. (2003). Testing the language mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism, 6,* 2–16. doi: http://dx.doi.org/10.1080/13670050308667769

Dijkstra, A. T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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

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

Dijkstra, T., & Rekké, S. (2010). Towards a localist-connectionist model of word translation. *The Mental Lexicon, 5,* 401–420. doi: http://dx.doi.org/10.1075/ml.5.3.08dij

Dijkstra, A. T., Timmermans, M., & Schriefers, H. (2000). Cross-language effects on bilingual homograph recognition. *Journal of Memory and Language, 42,* 445–464.

Dornic, S. (1978). The Bilingual's Performance: Language Dominance, Stress, and Individual Differences. In D. Gerver & H. W. Sinaiko (Eds.) *Language, Interpretation and Communication* (pp. 259-271). New York: Plenum Press. Dornyei, Z. & Scott. doi: 10.1007/978-1-4615-9077-4_23

Ecke, P. (2015). Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism: Language & Cognition, 18,* 145–162. doi: 10.1017/S1366728913000722

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8,* 1784–1790. doi: 10.1038/nn1594

Festman, J. (2008). Cross-language interference during trilingual picture naming in single and mixed language conditions. In M. Gibson, B. Hufeisen, & C. Personne (Hg.), *Mehrsprachigkeit: lernen und lehren, Multilingualism: Learning and Instruction, Le Plurilinguisme: apprendre et enseigner, O Pluilinguismo: aprender e ensinar* (pp. 109–119). Selected papers from the L3-Conference in Fribourg, Switzerland.

Festman, J. (2009). Three languages in mind. Saarbrücken: VDM.

Festman, J. (2012). Language control abilities of late bilinguals. *Bilingualism: Language and Cognition*, *15,* 580–593. doi: https://doi.org/10.1017/S1366728911000344

Festman, J., & Mosca, M. (2016). Influence of preparation time on language control: A trilingual digit naming study. In J. W. Schwieter (Eds.) *Cognitive control and consequences of multilingualism* (pp. 147–172). Canada: John Benjamins. doi: 10.1075/bpa.2.07fes

Festman, J., & Schwieter, J.W. (2015). Behavioural measures of language control: Production and comprehension. In *The Cambridge Handbook of bilingual language processing*, ed. By John W. Schwieter. Cambridge: Cambridge University Press, 527–547. doi: 10.1017/CBO9781107447257.023

Filippi, R., Karaminis, T., & Thomas, M. S. C. (2014). Bilingual language switching in production: Empirical and computational studies. *Bilingualism: Language and Cognition, 17*, 294–315. doi: http://dx.doi:10.1017/S1366728913000485

Filippi, R., Richardson, F. M., Dick, F., Leech, R., Green, D. W., Thomas, M. S. C., & Price, C. J. (2011). The right posterior paravermis and the control of language interference. *Journal of Neuroscience*, *31,* 10732–10740. doi: 10.1523/JNEUROSCI.1783-11.2011

Fink, A., & Goldrick, M. (2015). Pervasive benefits of preparation in language switching. *Psychonomic Bulletin and Review*, *22,* 808–814. doi: https://doi.org/10.3758/s13423-014-0739-6

Finkbeiner, M., Almeida, J., Janssen, N., & Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32,* 1075–1089. doi: 10.1037/0278-7393.32.5.1075

Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition, 9,* 153–166. doi:10.1037/0278-7393.32.5.1075

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35,* 116–124.

Garbin G., Costa, C., Sanjuan, A., Forn, C., Rodríguez-Pujadas, A., Ventura, N., Belloch, V., Hernandez, M., & Ávila, C. (2011). Neural bases of language switching in high and early proficient bilinguals. *Brain and Language, 119,* 129–135. doi: http://dx.doi.org/10.1016/j.bandl.2011.03.011

Gemba, H., & Sasaki, K. (1989). Potential related to no-go reaction of go/no-go hand movement task with color discrimination in human. *Neuroscience Letters, 101,* 263–268. doi: http://dx.doi.org/10.1016/0304-3940(89)90543-0

Giordano A. M. McElree B. Carrasco M. (2009). On the automaticity and flexibility of covert attention: A speed-accuracy trade-off analysis. *Journal of Vision, 9,* 1–10. doi:10.1167/9.3.30

Goldrick, M., Runnqvist, E., & Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science*, *25*, 1031–1036. doi:10.1177/0956797613520014

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. doi:10.1037/a0014981

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16,* 562–576. doi: 10.1037/0894-4105.16.4.562

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 331–355). Cambridge, MA: MIT Press.

Grainger, J. (1993). Visual word recognition in bilinguals. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 11–26). Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/sibil.6

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Quarterly Journal of Experimental Psychology*, *39A,* 295–319. doi: http://dx.doi:10.1080/14640748708401788

Grainger, J. & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R.J. Harris (Ed.) *Cognitive processing in bilinguals* (pp. 207–220). Amsterdam: North Holland. doi: http://dx.doi.org/10.1016/S0166-4115(08)61496-X

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d), In M. Kailand M. Hickmann (Eds.), *Language Acquisition across Linguistic and Cognitive Systems* (pp. 267–284). New York, NY: Benjamins. doi: http://dx.doi:10.1075/lald.52.18gra

Green, D. W. (1986). Control, activation and resource. *Brain and Language, 27,* 210–223. doi: http://dx.doi.org/10.1016/0093-934X(86)90016-7

Green, D. W. (1993). Towards a model of L2 comprehension and production. In R. Schreuder & B. Weltens (Eds.), *The Bilingual Lexicon,* 249–277. Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/sibil.6

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81. doi: 10.1017/S1366728998000133

Green D. W., & Abutalebi J. (2013). Language control in bilinguals: the adaptive control hypothesis. *Journal of Cognitive Psychology, 25,* 515–530. doi: 10.1080/20445911.2013.796377

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

de Groot, A. M. B. (1998). Retroactive or proactive control of the bilingual system. *Bilingualism: Language and Cognition, 1,* 86–87. doi: https://doi.org/10.1017/S1366728998000157

Grosjean, F. (1988). Exploring the recognition of guest words in bilingual speech. *Language and Cognitive Processes, 3,* 233–274. doi: http://dx.doi.org/10.1080/01690968808402089

Grosjean, F. (1997). The bilingual individual. *Interpreting, 2,* 163–187. doi: 10.1075/intp.2.1-2.07gro

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition*, *1*, 175–176. doi: https://doi.org/10.1017/S1366728998000285

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Eds.), *One Mind, Two Languages: Bilingual Language Processing* (pp. 1–22). Oxford: Blackwell

Guo, T., Liu, F., Chen, B., & Li., S. (2013). Inhibition of non-target languages in multilingual word production: Evidence from Uighur-Chinese-English trilinguals. *Acta Psychologica*, *143,* 277–283. doi: 10.1016/j.actpsy.2013.04.002

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese–English bilinguals. *NeuroImage*, *56,* 2300–2309. doi: 10.1016/j.neuroimage.2011.03.049

Guo, T., Ma, F., & Liu, F. (2013). An ERP study of inhibition of non-target languages in trilingual word production. *Brain and Language, 127*, 12–20. doi: 10.1016/j.bandl.2013.07.009

Guo, G., & Zhao, H. (2000). Multilevel modeling for binary data. *Annual Review of Sociology, 26,* 441–462. 10.1146/annurev.soc.26.1.441

van Halem, N. (2016). The Multilink model for word translation: Similarity effects in word recognition and word translation. Unpublished bachelor's thesis, Radboud University.

Hammarberg, B. (2001). Roles of L1 and L2 in L3 production and acquisition. In J. Cenoz, B. Hufeisen and U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition* (pp. 21–41). Clevedon: Multilingual Matters.

Hermans, D., Bongaerts, T., de Bot K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition, 1,* 213–229. doi: https://doi.org/10.1017/S1366728998000364

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish-English bilinguals: an fMRI study. *NeuroImage, 14,* 510–520. doi: 10.1006/nimg.2001.0810

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. doi: 10.1006/brln.1999.2278

van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighbourhood effects in bilingual word recognition. *Journal of Memory and Language, 39,* 458–483. doi: 10.1006/jmla.1998.2584.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition, 100,* 464–482. doi: 10.1016/j.cognition.2005.02.006

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*, 169–178. doi: 10.1017/S1366728901000268

Jackson, G. M., Swainson, R., Mullin, A., Cunnington, R., & Jackson, S. R. (2004). ERP correlates of a receptive language-switching task. *Quarterly Journal of Experimental Psychology, 2,* 223–240. doi: http://dx.doi:10.1080/02724980343000198

Jacoby, L. L., Kelley, C. M., & McElree, B. D. (1999). The role of cognitive control: Early selection versus late correction. In: Chaiken, S.; Trope, E., editors. *Dual Process Theories in Social Psychology* (pp. 383–400). Guildford; NY.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59,* 434–446. doi: 10.1016/j.jml.2007.11.007

Jin, Z. L., Zhang, J. X., & Li, L. (2014). Endogenous language control in Chinese-English sitching: an event-related potentials study. *Neuroscience Bulletin, 30,* 461–468. doi: 10.1007/s12264-013-1427-7

Jodo, E., & Kayama, Y. (1992). Relation of a negative ERP component to response inhibition in a Go/No-go task. *Electroencephalography and Clinical Neurophysiology, 82,* 477–482. doi: http://dx.doi.org/10.1016/0013-4694(92)90054-L

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye's movement. In Long, J. B. & Baddeley, A. D. (Eds.). *Attention and performance IX* (pp. 187–204). Hillsdale, NJ: Erlbaum.

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*, 5–14. doi: 10.1027/2151-2604/a000125

Kellerman, E. (1983). Now you see it, now you don't. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 112–134). Rowley, MA: Newbury House. doi: 10.1075/lald.5

Kerns, J. G., Cohen, J. D., MacDonald III, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026. doi: 10.1126/science.1089910

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42,* 627–633. doi: 10.3758/BRM.42.3.627

Khateb, A., Abutalebi, J., Michel, C. M., Pegna, A. J., Lee-Jahnke, H., & Annoni, J.-M. (2007). Language selection in bilinguals: A spatio-temporal analysis of electric brain activity.

International Journal of Psychophysiology, 65, 201–213. doi: 10.1016/j.ijpsycho.2007.04.008

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

Kim, Y., Gitelman, D., Nobre, A. C., Parrish, T. B., LaBar, K. S., & Mesulam, M., -M. (1999). The Large-Scale Neural Network for Spatial Attention Displays Multifunctional Overlap But Differential Asymmetry. *NeuroImage, 9,* 269–77. doi: 10.1006/nimg.1999.0408

Kinchla, R. A. (1980). The measurement of attention. In Nikerson R. S. (Eds.). *Attention and performance IX* (pp. 213–238). Hillsdale, NJ: Erlbaum.

Kleinman, D., & Gollan, T. H. (2016). Speaking Two Languages for the Price of One: Bypassing Language Control Mechanisms via Accessibility-Driven Switches. *Psychological Science, 27,* 700–714. doi: 10.1177/0956797616634633

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18,* 655–681. doi: 10.3389/fpsyg.2010.00238

Koch, I., & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory & Cognition, 34,* 433–444. doi: 10.3758/BF03193420

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching - A review. *Psychonomic Bulletin and Review, 17,* 1–14. doi: 10.3758/PBR.17.1.1

Koch, I., Prinz, W., & Allport, A. (2005). Involuntary retrieval in alphabet-arithmetic tasks: Task-mixing and task-switching costs. *Psychological Research*, *69*, 252–261. doi: 10.1007/s00426-004-0180-y

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging*, *15*, 126–147. doi: https://doi.org/10.1037//0882-7974.15.1.126

Kroll, J. F., Bobb, S. C., Misra, M. M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128,* 416–430. doi: http://dx.doi.org/10.1016/j.actpsy.2008.02.001

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition, 9,* 119–135. doi: http://dx.doi:10.1017/S1366728906002483

Kroll, J., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. *Journal of Memory and Language, 33*, 149–174. doi: http://dx.doi.org/10.1006/jmla.1994.1008

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 289–307). Oxford: Oxford University Press.

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition, 32,* 533–550. doi: 10.3758/BF03195845

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 38–75.

Levy, B. J., Mc Veigh, N., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science, 18,* 29–34. doi: http://dx.doi:10.1111/j.1467-9280.2007.01844.x

Lima, S. D., & Huntsman, L. A. (1997). Sequential dependencies in the lexical decision task. *Psychological Research, 60,* 264–269. doi: http://dx.doi:10.1007/BF00419412

Linck, J., Schwieter, J.W., & Sunderman, G. (2012). Inhibitory control predicts language switching performance in trilingual speech production. *Bilingualism: Language and Cognition, 15,* 651–662. doi: 10.1017/S136672891100054X

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologica, 38,* 30–42. http://dx.doi.org/10.1016/S0028-3932(99)00106-2

Liu, H., Fan, N., Rossi, S., Yao, P., & Chen, B. (2016). The effectmof cognitive flexibility on task switching and language switching. *International Journal of Bilingualism*, *20,* 563–579. doi: 1367006915572400

Liu, H., Rossi, S., Zhou, H., & Chen, B. (2014). Electrophysiological Evidence for Domain-General Inhibitory Control during Bilingual Language Switching. *PLoS ONE*, *9*, e110887. http://doi.org/10.1371/journal.pone.0110887

Logan, G., & 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. doi: 10.1037/0096-1523.29.3.575

Los, S. A. (1996). On the origin of mixing costs: Exploring information processing in pure and mixed blocks of trials. *Acta Psychologica, 94,* 145–188. doi: http://dx.doi.org/10.1016/0001-6918(95)00050-X

Luce, R. (1986). Response times. New York: Oxford University Press.

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture–word interference. *Canadian Journal of Psychology, 36,* 349–376. doi: 10.1037/h0080652

Lusetti, S. (2001). Standardizzazione di un test di fluenza verbale: taratura sui normali. Tesi di Diploma e Laurea in Logopedia. [http://www.logopedista.it/public/doc/201052-ECU-246-01.pdf]

Ma, F., Li, S., & Guo, T. (2016). Reactive and proactive control in bilingual word production: An investigation of influential factors. *Journal of Memory and Language*, *86,* 35–59. doi: http://dx.doi.org/10.1016/j.jml.2015.08.004

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288,* 1835–1838. doi: 10.1126/science.288.5472.1835

MacGregor L. J., Pulvermüller F., van Casteren M., & Shtyrov Y. (2012). Ultra-rapid access to words in the brain. *Nature Communication, 3,* 711. doi: 10.1038/ncomms1715

Macizo, P., Bajo, T., & Paolieri, D. (2012). Language switching and language competition. *Second Language Research, 28,* 131–149. doi: 10.1177/0267658311434893

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012) CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE, 7,* e43230. doi: 10.1371/journal.pone.0043230

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,* 940–967. doi: 10.1044/1092-4388(2007/067)

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language Competition. *Bilingualism: Language and Cognition, 6,* 97–115. doi: 10.1017/S1366728903001068

Martin, C. D., Strijkers, K., Santesteban, M., Escera, C., Hartsuiker, R. J., & Costa, A. (2013). The impact of early bilingualism on controlling a language learned late: An ERP study. *Frontiers in Psychology, 4,* 1–15. doi: 10.3389/fpsyg.2013.00815

Macnamara, J. (1967). The bilingual's linguistic performance: a psychological overview. *Journal of Social Issues, 23,* 58–77. doi: 10.1111/j.1540-4560.1967.tb00576.x

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. doi: 10.1037/h0026390

Mägiste, E. (1984). Stroop tasks and dichotic translation: the development of interference patterns in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 304–315. doi: 10.1037/0278-7393.10.2.304

Mägiste, E. (1985). Development of intra- and inter-lingual interference in bilinguals. *Journal of Psycholinguistic Research, 14,* 137–154. doi: 10.1007/BF01067626

Mariën, P., Abutalebi, J., Engelborghs, S., & De Deyn, P. P. (2005). Acquired subcortical bilingual aphasia in an early bilingual child: Pathophysiology of pathological language switching and language mixing. *Neurocase, 11,* 385–398. doi: 10.1080/13554790500212880

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. M. (2015). Balancing Type I Error and Power in Linear Mixed Models. arXiv 1511.01864.

Mayr, U. (2001). Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. *Psychology and Aging, 16,* 96–109. doi: 10.1037/0882-7974.16.1.96

Mayr, U. (2009). Sticky plans: Inhibition and binding during serial-task control. *Cognitive Psychology, 59,* 123–153. doi: 10.1016/j.cogpsych.2009.02.004

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology:General, 129,* 4–26.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1124–1140. doi: 10.1037/0278-7393.26.5.1124

Mayr, U., & Kliegl, R. (2003). Differential effects of cue changes and task changes on task-set selection costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 362–372. doi: 10.1037/0278-7393.29.3.362

McCallum, W.C., Farmer, S.F., & Pocock, P.K. 1984. The effects of physical and semantic incongruities on auditory event- related potentials. *Electroencephalography and Clinical Neurophysiology, 59,* 447–488. doi: 10.1016/0168-5597(84)90006-6

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86. doi: http://dx.doi.org/10.1016/0010-0285(86)90015-0

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part I. An account of basic findings. *Psychological Review, 88,* 375–407. doi: http://dx.doi.org/10.1037/0033-295X.88.5.375

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1423–1442. doi: 10.1037/0278-7393.22.6.1423

Meiran, N. (2000). Modeling cognitive control in task-switching. *Psychological Research, 63,* 234–249. doi: 10.1007/s004269900004

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41,* 211–253. doi: 10.1006/cogp.2000.0736

Meuter, R. (2005). Language selection in bilinguals: Mechanisms and processes. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 349–370). New York, NY: Oxford University Press.

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

Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology: General, 90*, 227–234. doi: 10.1037/h0031564

Misra, M., Guo, T., Bobb, S., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language, 67,* 224–237. doi: 10.1016/j.jml.2012.05.001

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences, 7,* 134–140. doi: 10.1016/S1364-6613(03)00028-7

Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research, 63,* 250–264. doi: 10.1007/s004269900005

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language, 80,* 188–207. doi: 10.1006/brln.2001.2588

Mosca M., & Clahsen H. (2016). Examining language switching in bilinguals: The role of task-demand. *Bilingualism: Language and Cognition*. doi: 10.1017/S1366728915000693

Nas, G. (1983). Visual word recognition in bilinguals: evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning and Verbal Behavior, 22,* 526–534. doi: 10.1016/S0022-5371(83)90319-5

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading* (pp. 264–336). Hillsdale, NJ: Lawrence Erlbaum Associates.

Ng, S., Gonzalez, C., & Wicha, N.Y. (2014). The fox and the cabra: an ERP analysis of reading code switched nouns and verbs in bilingual short stories. *Brain Research*, *1557,* 127–40. doi: 10.1016/j.brainres.2014.02.009

Orfanidou, E., & Sumner, P. (2005). Language switching and the effects of orthographic specificity and response repetition. *Memory & Cognition, 33,* 355–369. doi: http://dx.doi:10.3758/BF03195323

Paquette, L., & Kida, T. (1988). The effect of decision strategy and task complexity on decision performance. *Organizational Behavior and Human Decision Processes, 41,* 128–142. doi: 10.1016/0749-5978(88)90051-9

Paradis, M. (1987). *The Assessment of Bilingual Aphasia.* Hillsdale, NJ.: Lawrence Erlbaum.

Peacock, J. (2015). Modelling Lexical Effects with Multilink: Frequency, Cognate Status, and Translation Asymmetry. Unpublished master's thesis, Radboud University.

Peelen, M., V., Heslenfeld, D., J., & Theeuwes, J. (2004). Endogenous and exogenous attention shifts are mediated by the same large-scale neural network. *NeuroImage, 22,* 822–830. doi: http://dx.doi.org/10.1016/j.neuroimage.2004.01.044

Peeters, D., Runnqvist, E., Bertrand, D., & Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 40, 284–292. doi: 10.1037/a0034060

Penfield, W., & Roberts, L. (1959). Speech and brain mechanisms. Princeton, N. J.. Princeton University Press. doi: 10.1002/bjs.18004720433

Perea, M., & Carreiras, M. (2003). Sequential effects in the lexical decision task: The role of the item-frequency of the previous trial. *Quarterly Journal of Experimental Psychology, 56,* 385–401. doi:10.1080/02724980244000387

Pestilli F., & Carrasco M. (2005). Attention enhances contrast sensitivity at cued and impairs it at uncued locations. *Vision Research, 45,* 1867–1875. doi: 10.1016/j.visres.2005.01.019

Pestilli, F., Viera, G., & Carrasco, M. (2007). How do attention and adaptation affect contrast sensitivity? *Journal of Vision, 7,* 1–12. doi: 10.1167/7.7.9.

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography and Clininical Neurophysiolpgy, 60,* 423–434. doi: http://dx.doi.org/10.1016/0013-4694(85)91017-X

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. doi: 10.1080/09541440600758812

Philipp, A. M., Kalinich, C., Koch, I., & Schubotz, R. I. (2008). Mixing costs and switch costs when switching stimulus dimensions in serial predictions. *Psychological Research, 72,* 405–414. doi: 10.1007/s00426-008-0150-x

Philipp, A. M., & Koch, I. (2006). Task inhibition and task repetition in task switching. *European Journal of Cognitive Psychology, 18,* 624–639. doi: 10.1080/09541440500423269

Philipp, A., & 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. doi: 10.1037/a0016376

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-effects models in S and S-PLUS. *Statistics and Computing.* Springer, New York.

Pinto, Y., van der Leij, A.R., Sligte, I. G., Lamme, V. A. F., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. *Journal of Vision, 13,* 16. doi:10.1167/13.3.16

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. doi: http://dx.doi.org/10.1080/00335558008248231

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15,* 36–57. doi: https://doi.org/10.1093/applin/15.1.36

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

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

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). URL: http://www.pstnet.com

R Development Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing. [http://www.R-project.org].

R Development Core Team. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing. [http://www.R-project.org].

Rafal, R., & Henik, A. (1994).The neurology of inhibition: integrating controlled and automatic processes. In D. Dagenbachand T. H. Carr (Eds.), *Inhibitory Processes in Attention, Memory, and Language* (pp. 1–51). San Diego, CA: Academic Press.

Rekké, S. T. (2009). Multilink: a model for multilingual processing. Unpublished bachelor’s thesis, Radboud University.

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1303–1323. doi: 10.1037/0096-1523.32. 6.1303

Reynolds, M. G., Schlöffel, S., & Peressotti, F. (2016). Asymmetric Switch Costs in Numeral Naming and Number Word Reading: Implications for Models of Bilingual Language Production. *Frontiers in Psychology, 6,* 2011. doi: 10.3389/fpsyg.2015.02011

Ringbom, H. (1987). The Role of the First Language in Foreign Language Learning. Clevedon: Multilingual Matters.

Robinson, D. L., & Kertzman, C. (1995). Covert orienting of attention in macaques. III. Contributions of the superior colliculus. *Journal of Neurophysiology, 74,* 713–721.

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Noesselt, T., & Muente, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature, 415,* 1026–1029. doi: 10.1038/4151026a

Rodriguez-Fornells, A., van der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., & Muente, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and Functional imaging evidence. *Journal of Cognitive Neuroscience, 17,* 422–433. doi: 10.1162/0898929053279559

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231. doi: 037/0096-3445.124.2.207

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object set: The role of surface detail in basic-level object recognition. *Perception, 33,* 217–236.

Rothman, J. (2015). Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition, 18,* 179–190. doi: https://doi.org/10.1017/S136672891300059X

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1477–1491. doi: 10.1037/0278-7393.31.6.1477

Rubinstein, J., Meyer, D., & Evans, J. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 763–797. doi: 10.1037/0096-1523.27.4.763

Runnqvist, E., Strijkers K., Alario F. X., & Costa A. (2012). Cumulative semantic interference is blind to language: implications for models of bilingual speech production. *Journal of Memory and Language, 66,* 850–869. doi: 10.1016/j.jml.2012.02.007

Schepens J., Dijkstra T., & Grootjen F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition, 15,* 157–166. doi: https://doi.org/10.1017/S1366728910000623

Schepens J., Dijkstra T., Grootjen F., & van Heuven W.J.B. (2013) Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE,* 8, e63006. doi:10.1371/journal.pone.0063006

Schmid, M., & Jarvis, S. (2014). Lexical access and lexical diversity in first language attrition. *Bilingualism: Language and Cognition 17,* 729–748. doi: 10.1017/S1366728913000771

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. doi: 10.1037/0096-1523.29.1.92

Sdoia, S., & Ferlazzo, F. (2008). Stimulus-related inhibition of task set during task switching. *Experimental Psychology, 55,* 322–327. doi: 10.1027/1618-3169.55.5.322

Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, *16,* 304–324. doi: 10.1017/S1366728912000466

Schreij D., Owens C., & Theeuwes J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Attention, Perception, & Psychophysics*, *70,* 208–218. doi: 10.3758/PP.70.2.208

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information-processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190. doi: 10.1037/0033-295X.84.2.127

Schwieter, J. W., & Sunderman, G. (2008). Language switching in bilingual speech production*. The Mental Lexicon, 3,* 214–238. doi: 10.1075/ml.3.2.06sch

Schwieter, J. W., & Sunderman, G. L. (2011). Inhibitory control processes and lexical access in trilingual speech production. *Linguistic Approaches to Bilingualism, 4,* 391–412.

Schwieter, J. W. (2013). Lexical inhibition in trilingual speakers. In: Tirkkonen, J., Anttikoski, E. (Eds.), *Proceedings of The 24th Conference of Scandinavian Linguistics*. Publications of the University of Eastern Finland: Reports and Studies in Education, Humanities, and Theology (pp. 249–260). University of Eastern Finland Press, Joensuu, Finland. doi: 10.1075/lab.1.4.02sch

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature, 415*, 318–320. doi: 10.1038/415318a

Singleton, D. (1987). Mother and other tongue influence on learner French. *Studies in Second Language Acquisition*, *9*, 327–346. doi: https://doi.org/10.1017/S0272263100006719

Slevc, L. R., Davey, N. S., & Linck, J. A. (2016). A new look at "the hard problem" of bilingual lexical access: evidence for language-switch costs with univalent stimuli. *Journal of Cognitive Psychology, 28,* 385–395. doi: http://dx.doi.org/10.1080/20445911.2016.1152274

Spector, A., & Biederman, I. (1976). Mental set and mental shift revisited. *American Journal of Psychology, 89,* 669–679. doi: 10.2307/1421465

Stokholm, J., Jorgensen, K., & Vogel, A. (2013). Performance on five verbal fluency tests in a healthy, elderly Danish sample. *Aging, Neuropsychology, and Cognition, 20,* 22–33. doi: 10.1080/13825585.2012.656576

von Studnitz, R. E., & Green, D. W. (1997). Lexical decision and language switching. *International Journal of Bilingualism, 1,* 3–24. doi: http://dx.doi.org/10.1177/136700699700100102

von Studnitz, R. E., & Green, D. W. (2002). The cost of switching language in a semantic categorization task. *Bilingualism: Language and Cognition, 5,* 241–251. doi: 10.1017/S1366728902003036

Suzuki, K., & Shinoda, H. (2015). Transition from reactive control to proactive control across conflict adaptation: An sLORETA study. *Brain and Cognition*, *100,* 7–14. doi: 10.1016/j.bandc.2015.09.001

Tarlowski, A., Wodniecka, Z., & Marzecová, A. (2012). Language Switching in the Production of phrases. *Journal of Psycholinguistic Research, 42,* 103–118. doi: 10.1007/s10936-012-9203-9

Thomas, M. S. C., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of Memory and Language, 43,* 44–66. doi: http://dx.doi.org/10.1006/jmla.1999.2700

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. Nature 381, 520–522. doi: 10.1038/381520a0

Venables, W. N., & Ripley B. D. (2002). Modern Applied Statistics with S. Springer, New York, fourth edition. [http://www.stats.ox.ac.uk/pub/MASS4]. doi: 10.1007/978-0-387-21706-2

Verhoef, K., Roelofs, A., & Chwilla, D. (2009). Role of inhibition in language switching: Evidence from event related brain potentials in overt picture naming. *Cognition, 110,* 84–99. doi: https://doi.org/10.1016/j.cognition.2008.10.013

Verhoef, K., Roelofs, A., & Chwilla, D. (2010). Electrophysiological evidence for endogenous control in switching attention between languages in overt picture naming. *Journal of Cognitive Neuroscience*, *22*, 1832–1843. doi: 10.1162/jocn.2009.21291

Vogelgesang, P. (2003). Häufigkeit von Buchstaben. In: Peters, A. (2008). *Intelligenter Austausch von Daten in Private Area Networks*. GRIN Verlag.

Wang, X. (2015). Language control in bilingual language comprehension: evidence from the maze task. *Frontiers in Psychology, 6,* 1179. doi: 10.3389/fpsyg.2015.01179

Wang, Y., Kuhl, P. K., Chen, C., & Dong, Q. (2009). Sustained and transient language control in the bilingual brain. *Neuroimage, 47,* 414–422. doi: http://dx.doi.org/10.1016/j.neuroimage.2008.12.055

Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *NeuroImage, 35*, 862–870. doi: 10.1016/j.neuroimage.2006.09.054

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50,* 439–456. doi: http://dx.doi.org/10.1080/713755710

Weissberger, G., Wierenga C., Bondi M., & Gollan T. (2012). Partially over-lapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging, 27,* 959–974. doi: 10.1037/a0028281

Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: implications for a polyglot speaking model. *Applied Linguistics*, *19*, 295–333. doi: 10.1093/applin/19.3.295

Wylie, G., & Allport, A. (2000). Task switching and the measurement of “switch costs.” *Psychological Research, 63,* 212–233. doi: 10.1007/s004269900003

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 455–469. doi: 10.1037/0096-1523.29.2.455

Selbstständigkeitserklärung

Gemäß §4 (2) 7. der Promotionsordnung der Humanwissenschaftlichen Fakultät der Universität Potsdam vom 15. Mai 2013

Hiermit erkläre ich, Michela Mosca, dass ich die Arbeit selbständig und ohne unzulässige Hilfe Dritter verfasst habe und bei der Abfassung alle Regelungen guter wissenschaftlicher Standards eingehalten habe.

Potsdam, den ____________________ ______________________________________
Michela Mosca