BEHAVIOURAL AND ELECTROPHYSIOLOGICAL EVIDENCES FOR THE EFFECT OF BILINGUALISM ON SPEAKERS' COGNITIVE CONTROL ABILITY

BY

BEINAN ZHOU

A thesis submitted to

The University of Birmingham

for the degree of

Doctor of Philosophy

School of Psychology

The University of Birmingham

ABSTRACT

Bilingualism means more than speaking two languages, it also has cognitive consequences. Recently, the question whether or not bilingualism affects cognitive control abilities has raised fierce debates. In this thesis, the effect of bilingualism on executive control was tested in different domains. First, bilingual speakers showed enhanced attentional control abilities while performing non-verbal executive control tasks. This was evident in terms of their response distribution profiles, which showed less extreme distribution tails than those of monolinguals, suggesting less frequent lapses of attention. Second, bilingual speakers resolved intra-language lexical competition differently from monolingual speakers. Their response distribution profiles as well as their brain activities were differentially affected in a picture naming task that manipulated the level of lexical competition. These results were best explained by bilingual speakers having enhanced engagement of executive control while resolving lexical competition within a single language, even though this might not be reflected at behavioural level. Third, bilingual speakers demonstrated enhanced task shifting abilities at a latent factor level, while they did not differ from monolinguals with regard to inhibition and updating abilities. Results also suggested a more correlated network of executive control for bilingual speakers than for monolingual speakers. Therefore, this thesis has obtained converging evidence that bilingualism benefits executive control. Reasons for inconsistencies in the literature and absence of the bilingualism effects are discussed.

ACKNOWLEDGEMENTS

I have many people to thank, for their help and support in any means over the past four years. As the list of names becomes longer and longer, I feel blessed and spoiled.

My deepest gratitude must go to Dr. Andrea Krott, my supervisor, for her enthusiastic supervision. This thesis would not have been in its shape without your patient guidance, insightful advice, timely comments and feedbacks along the way. Thank you for your strong supports and warm encouragements in the troughs, as well as the occasional pushes and pulls through the lows. Thank you.

Thanks also go to Dr. Yang Zhang for his technical supports implementing testing programs; and Dr. Camillo Porcaro for his teaching on EEG data analyses. Data collection for the huge project would not have been possible without my lovely helpers: Emily, Ffion, Manuela, Wendy, Verity and Vivienne.

The psycholinguistic lab deserves a special mention. The many lab meetings that we shared were enjoyable. The inspiring discussions and gentle critics helped to improve the quality of my work. I am really grateful to be part of it.

I would like to take this opportunity to thank my dear friends, who make me believe that I am not alone on this journey. Thank you, Drew, Evi, Kareen, Kristian, Isabella, Johnny, Ju, Mel, Saheeda, Shan, Rui, Yishin, Yuki and Zheni.

Finally, I would like to thank my family for their constant support, encouragement, and understanding, without whom I would never have embarked on this PhD. Thank you, Mum and Dad for providing generous financial support so that I can pursue my goals free of worries. Thank you for believing in me when I doubt myself, helping me to regain self-confidence. I LOVE YOU, Mum and Dad! Thanks also go to Chen for being there for me, although thousands of miles apart, listening and supporting.

博学之，审问之，慎思之，明辨之，笃行之。

《中庸》

Study extensively, inquire thoroughly, think carefully, discern critically, and act resolutely.

Zhong Yong

致我挚爱的父亲周抗母亲张洁

致 青 春

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Statement of Research Question

Bilingualism is more pervasive than it might seem. It occurs in all age groups, at all levels of society, and in most countries (Grosjean, 2013): at least half of the world's population is bilingual (Grosjean, 1982). Therefore, bilingualism is the norm rather than the exception.

It has been well established, particularly in the aging literature, that cognitively demanding activities, such as education, leisure activities or professional occupations, are associated with a higher level of intellectual functioning (for a review, see Kramer, Bherer, Colcombe, Dong, & Greenough, 2004). Bilingualism may also be one of the various lifestyle factors that have a pervasive effect on cognitive functioning. For example, bilingual speakers outperform monolingual speakers in various domains, including meta-linguistic awareness, mental flexibility, attentional control and inhibitory control (for an early overview, see Bialystok, 2001; for recent reviews, see Bialystok et al., 2009; Kroll & Bialystok, 2013).

Given the ubiquitous nature of bilingualism, its cognitive consequences have critical implications for policy-making and classroom assessment. For example, knowing the benefits and disadvantages of bilingualism might help parents to make informed decisions about raising children bilingually. Knowledge about how typical bilingual and typical monolingual children perform in a range of verbal or non-verbal assessments is useful to determine whether a child is on a typical developing trajectory. This is not only relevant for child development and child language acquisition, but may also inform adults in their decision to learn a second language. Studies of people who became bilingual later in life demonstrate that late acquisition can also benefit speakers in terms of better auditory attention (e.g. Bak, Vega-Mendoza, & Sorace, 2014). Last but not least, bilingualism provides a good example of neural flexibility. Studies on how bilingualism shapes the brain thus deepen our understanding of how variations in life experience affect human cognition. The aim of this thesis is to contribute to the understanding of the impact of bilingualism on speakers' cognitive abilities.

Definition of Terms: Scope of the Thesis

This section defines the scope of this thesis. As the title suggests, this piece of work aims to add to our understanding of the relation between bilingualism and cognitive control. Therefore, we first define bilingualism and outline the features of the bilingual speakers studied in the research described in this thesis. Second, we define the scope of cognitive control relevant for the question at hand, namely executive control.

Bilingualism

Definition of bilingualism.

Bilingualism may be defined from different perspectives (Romaine, 1989). A speaker may be a 'coordinate' or a 'compound' bilingual, depending upon the acquisition context of their two languages. In coordinate bilinguals linguistic representations of the two languages have their uniquely associated concepts. This is common in speakers who learnt two languages in distinct settings/environments. In compound bilinguals, linguistic representations of both languages are related to the same concept. This is often found in fluent L2-speakers. A speaker might also be described as an 'early' or a 'late' bilingual, depending upon the stage in life during which they acquired their second language. Or one might be described as a 'balanced' or a 'dominant' bilingual, depending upon the relative strength of each language. Some definitions limit bilingual speakers to those who have 'native-like' control of both languages, whereas others embrace a minimal definition that recognises the initial stages of contact between two languages. Since this thesis focuses on the cognitive outcome of speaking more than one language, following Grosjean (2013), bilingualism is defined here as the use of two (or more) languages in daily life.

Scope of the study.

Bilingual speakers who use both languages daily are not a completely homogeneous group, they still vary in terms of their language use. The diversity of language experiences seems to have a profound impact on the cognitive consequences of bilingualism. First, the bilingual advantage has been most evident in people who use both languages regularly (Bialystok, Peets, & Moreno, 2014; Yow & Li, 2015). Second, it is suggested that the age before which a second language is acquired potentially affects the interaction between bilingualism and cognitive ability. For example, Luk et al. (2011) found that bilinguals who started using both languages before the age of ten demonstrated a cognitive advantage whereas those who started using both languages later did not (Luk et al., 2011). Third, proficiency has been shown to correlate with cognitive performance in bilinguals (Mishra, Hilchey, Singh, & Klein, 2012; Tse & Altarriba, 2012). Therefore, in order to compare the cognitive performance of monolingual and bilingual speakers, this thesis targets bilingual speakers who use both the two languages on a daily basis, who are fluent in both languages, who perceive themselves to be proficient in both languages, and acquired the two languages before the age of ten.

Executive Control

Executive control in the bilingual literature.

Executive control has been proposed to explain the superior performance of bilingual speakers compared to monolingual speakers in tasks such as conflict or switching tasks. Executive control is often used as an umbrella term encompassing a wide range of higher-order cognitive processes. Nevertheless, in the literature on the effect of bilingualism on cognition, the term is often poorly defined, underspecified and, most importantly, inconsistent. For example, executive control has been referred to as: 'fluid intelligence' (Bialystok, Craik, Klein, & Viswanathan, 2004), executive functioning or the conscious control of thought and action (Carlson & Meltzoff, 2008), selective attention (Bialystok, Martin, & Viswanathan, 2005), the ability to act according to goals (Morton & Harper, 2007), and the ability 'to monitor goal-setting cues, to switch attention to goal-relevant sources of information, and to inhibit those that are irrelevant or competing' (Paap & Greenberg, 2013). Several frameworks for executive control have been proposed in the literature (Miyake et al., 2000; Norman & Shallice, 1983; Posner & Petersen, 1990). Next, these frameworks are considered in order to set the scene for the current research.

Baddeley (1996) described executive control in his model of working memory. It was defined as a series of 'general purpose control mechanisms that modulate the operation of various cognitive subprocesses and thereby regulate the dynamics of human cognition'. Researchers since have proposed that executive control, instead of being a unified function, consists of different subprocesses. For instance, Miyake et al. (2000) identified three separable executive functions: inhibition (of prepotent responses), updating, and shifting. However, these subprocesses are not independent of one another. Using confirmatory factor analyses, the authors found a moderate correlation between these components, which points to a common component in the executive control network. They suggested that this component is 'controlled attention', a domain-free attentional capacity to actively maintain working memory representations.

Executive control has also been discussed as part of the attentional network proposed by Posner and Petersen (1990), along with an alerting network and an orienting attentional network. According to this framework, alerting concerns achieving and maintaining an alert state, orienting concerns selecting information from sensory input, and executive control concerns monitoring and resolving conflict.

Norman and Shallice (1983) proposed another framework for executive control, which featured a Supervisory Attentional System (SAS). According to their framework, thoughts and actions turn on the activation of a response or schema, and the role of the SAS is to oversee and control the activation of the schemas through prioritising certain schemas or increasing the activation levels of certain schemas. Though rarely discussed in the literature on bilingualism and cognition, the SAS has been used to explain how bilingual speakers resolve language competition. In his inhibitory control model of bilingual language production, Green (1998) proposed that successful production or comprehension in the target language must be achieved by some executive control mechanism that regulates or suppresses the activation of the non-target language, such as a SAS that inhibits the non-target language through language tags at lemma level.

The model by Miyake et al. (2000) is the most useful model for the present study. While the attentional network of Posner and Petersen and Norman and Shallice’s SAS focus on attentional processes, the framework put forward by Miyake et al. (2000) is more inclusive, featuring not only attentional processes but also control processes. The Miyake et al. model (as well as the Posner and Petersen model) is more structured, allowing each subprocess to be examined individually or as a whole. In contrast, the Norman and Shallice model is holistic, making it difficult to assess the processes with simple tasks. Finally, there is evidence that bilingual speakers not only outperform monolinguals in resolving conflict, but also in switching back and forth between task sets (Christoffels, de Haan, Steenbergen, van den Wildenberg, & Colzato, 2015; Prior, & MacWhinney, 2010). This pertains particularly to the experience whereby bilinguals switch between their different languages. The framework used to study the cognitive consequences of bilingualism should therefore include a component of switching, and this is only the case for the Miyake et al.'s model. In this thesis, executive control is therefore defined in terms consistent with the framework of Miyake et al. (2000). According to this framework, there are three subprocesses of executive control: inhibition, updating and shifting, as well as a common attentional function that actively maintains working memory representations.

Review of Literature

A Short History of the Study of Bilingualism and Its Cognitive Implications

Early interest in the cognitive consequences of bilingualism focused on general intelligence performance. Darcy (1953) conducted a review of studies on how bilingualism affected intelligence test performance. Findings suggested that bilinguals performed worse than age-matched monolingual speakers. For instance, Saer (1923) reported that a group of Welsh/English bilingual children scored lower on a Stanford-Binet test of intelligence than did age-matched monolingual children. However, the author noted a distinction between verbal and non-verbal measures of intelligence: whilst a bilingual disadvantage was consistently evident in measures of verbal intelligence, it was less so in non-verbal measurements. Bilinguals sometimes performed equally well as monolinguals. The author concluded that bilingual speakers were disadvantaged when their intelligence was measured verbally, possibly owing to a language handicap. Despite this subtle but critical distinction, the belief that speaking a second language is detrimental to a child's linguistic and cognitive development has dominated the field for decades. The inferior performance of bilingual speakers has been interpreted as mental confusion (Darcy, 1953).

However, early studies demonstrated methodological limitations (Hakuta & Diaz, 1985). For example, factors now known to affect cognitive abilities, such as socio-economic background, were not controlled for in these studies. Moreover, it is questionable how bilingualism was defined and selected in the early studies, for instance, whether bilingual speakers reached an appropriate proficiency level or not.

An important study in these early days is that by Peal and Lambert (1962), who evaluated age-matched English monolingual and English/French bilingual speakers from similar socio-economic backgrounds in Montreal. Contrary to expectation, bilinguals outperformed monolinguals on almost all of the tests, including tests on both verbal and non-verbal intelligence. This study is monumental, not only because of its unanticipated findings, but because it strictly controlled for the socio-economic and language background of participants. The authors distinguished between ‘balanced bilinguals’, who reached age-appropriate ability in both of their languages, and ‘pseudo bilinguals’ who did not reach age-appropriate ability in their less dominant language. Only the former group was included in the study. Furthermore, closer analysis revealed that the performances of monolingual and bilingual children on spatial-perceptual tests were comparable, but that bilinguals were better at tests requiring symbol manipulation and reorganisation. These observations led the authors to conclude that bilingual children had superior mental flexibility, possibly due to managing their two languages daily.

A large body of research on this topic has accumulated since then. Importantly, research over the last decades has shifted from holistic analysis (applying standardised intelligence tests to monolinguals and bilinguals) to a procedural-oriented approach (Bialystok, 2001). This has allowed more detailed investigation, by focusing on the impact of bilingualism on specific cognitive processes, which has gone some way towards illuminating an understanding of the effects of bilingualism. For example, Bialystok (1992) reported that bilingual children outperformed monolingual children in an Embedded Figure Test. Crucial to performing the task was the ability to focus attention on one configuration and to ignore the misleading configuration, locating the superior ability of bilingual children at selective attention and inhibition of irrelevant information (Bialystok, 2001; Bialystok et al., 2004).

Bilingualism and Executive Control in Nonverbal Tasks: Recent Debate

Studies using executive control tasks such as the Simon task (Simon & Rudell, 1967), the Spatial Stroop (Bialystok, 2006), and the Flanker task (Eriksen & Eriksen, 1974) indicated that bilingual speakers outperformed monolingual counterparts in a range of cognitive control tasks (for a review, see Bialystok, Craik, Green, & Gollan, 2009). All of these tasks shared a common feature: they all required focusing attention on particular cues or features of the stimuli and ignoring distracting information. However, despite consistent reports of the advantages of bilingualism, the exact result patterns have been inconsistent. Hilchey and Klein (2011) reviewed studies that investigated the effects of bilingualism on cognitive functioning, using three representative tasks. In their review, the authors distinguished between two patterns of results, which were attributable to different underlying functions. The first pattern is a smaller interference effect in bilinguals than in monolinguals, which has often been interpreted as bilinguals being better able to inhibit prepotent responses. The second pattern is an overall speed advantage for bilinguals compared with monolinguals, which has been related to a general advantage in attentional control. The authors also noted that, while elderly bilingual participants have consistently demonstrated superior inhibitory control, this effect has been significantly less evident in children and young adults (Hilchey & Klein, 2011). By contrast, findings on the overall speed advantage for bilinguals were more robust. The authors concluded that the advantages of bilingual cognitive control are not restricted to inhibition *per se*, but are of more global nature.

Paap and Greenberg (2013) published a comprehensive study, in which no group differences were found between monolingual and bilingual participants in a variety of executive control tasks, including the Simon task, the Flanker task, a switching task and an antisaccade task. This study raised disquiet because it questioned whether bilingualism affects cognitive control at all. Most importantly, it emphasised the importance of null results. Since then, various studies have reported comparable performances between monolinguals and bilinguals (Blumenfeld, & Marian, 2014; Gathercole et al., 2014; Kirk, Scott-Brown, & Kempe, 2013; Kousaie, Sheppard, Lemieux, Monetta, & Taler, 2014). This has led to claims that the perceived cognitive advantage of bilingualism is solely due to publication bias (de Bruin, Treccani, & Della Sala, 2015). Such articles have prompted serious reflection on how the evidence should be evaluated (Bialystok, Kroll, Green, MacWhinney, & Craik, 2015; Paap, 2014; Valian, 2015). Therefore, further endeavours to elucidate the causes for the inconsistent results reported in the literature are needed. They hopefully will bring us closer towards understanding the nature of the bilingualism effect.

The first two chapters of this thesis directly aim at understanding the reason(s) for inconsistent findings, by focusing on one particular potential cause, namely the treatment of long responses in the data. Using response time distribution analysis, some studies have found that bilingualism affects response distribution profiles (Calabria, Hernandez, Martin, & Costa, 2011; Tse & Altarriba, 2012): it does not only seem to affect the average speed of responses, but also the occurrence of long responses. Interestingly, the enhanced performance in the conflict condition seems to stem from fewer extremely long responses (Calabria, Hernandez, Martin, & Costa, 2011). Chapter 2 investigates the hypothesis that data trimming procedures have impacted the results, in that strong data trimming might have eliminated a bilingual advantage. In Chapter 3, this hypothesis is empirically investigated. It reports differences in response time distributions of well-controlled monolingual and bilingual speakers in three non-verbal interference tasks. An attempt is made to delineate the relative contributions of inhibition and attention components within the executive control network to the bilingual cognitive advantage.

Bilingualism and Executive Control in Verbal Tasks

Bilingualism is usually negatively associated with verbal ability outcomes. For example, bilingual speakers tend to have smaller vocabulary sizes than their monolingual peers, when compared in one language (Bialystok, Craik, & Luk, 2008; Bialystok, Luk, Peets, & Yang, 2010; Oller, Pearson, & Cobo-Lewis, 2007; Portocarrero, Burright, & Donovick, 2007). Bilingual speakers also tend to show impaired lexical retrieval, such as more frequent ‘tip of the tongue’ experiences (Gollan, Montoya, Cera, & Sandoval, 2008), and they perform more poorly on verbal fluency tasks (Gollan, Montoya, & Werner, 2002; Rosselli et al., 2002; Sandoval, Gollan, Ferreira, & Salmon, 2010). This has been attributed to various reasons. First, bilinguals tend to have smaller vocabularies than monolinguals when taking only one of their languages. This means that it is more difficult to retrieve as many exemplars as monolinguals. Second, bilingual speakers need to resolve competition with words from their ‘other’ language (Sandoval et al., 2010). Third, bilingual speakers’ exposure to a single language is more limited than that of monolingual speakers. This means that, on average, bilingual speakers encounter or retrieve each lexical item less often than monolingual speakers, leading to less efficient lexical retrieval (Gollan et al., 2008; Ivanova & Costa, 2008). Given the claim that bilinguals have enhanced executive control abilities, the question arises whether these abilities might actually help bilinguals resolve lexical selection difficulties within a single language.

Despite a large body of literature on how bilingualism affects nonverbal cognition, there has been limited investigation into how it affects the executive control of verbal tasks.

Chapters 4 and 5 presents work investigating the effect of bilingualism on a verbal task under the condition of high lexical competition. Chapter 4 tests the hypothesis that bilingualism enhances executive control which assists speakers in resolving conflicts during competitive lexical selection, by means of a semantic blocking paradigm. Chapter 5 aims at gaining electrophysiological evidence for this hypothesis.

Bilingualism and Executive Control Network

Executive control is not just one single function, but it refers to a system with several interrelated functions, which together support flexible and goal-directed control of behaviours. Bilingualism might have a holistic impact on the executive network: that is, instead of enhancing or impeding the performance of an individual function, bilingualism might modify networks or the connection between different networks. This hypothesis has been supported by a recent study using functional neuroimaging which reported that bilingual speakers demonstrated greater functional connectivity in some of the executive function networks (Grady, Luk, Craik, & Bialystok, 2015).

Previous work has predominantly focused on the impact of bilingualism on one or more of the executive subprocesses (updating, shifting or inhibition) separately. Inhibition has been studied most, followed by shifting, and then updating. There are two studies that addressed the executive control network as a whole, but only within the bilingual population (Soveri, Rodriguez-Fornells, & Laine, 2011; Yow & Li, 2015). There are no studies that directly compared monolingual and bilingual speakers in terms of their executive control network.

Work described in Chapter 6 aims to fill this gap. It reviews in detail literature on how bilingualism affects each of the executive control subprocesses. Confirmatory factor analysis is used to compare the executive control network of monolingual and bilingual speakers. A latent variable analysis allows to analyse the effect of bilingualism on the mean performance of each executive control subprocess, as well as the relationship amongst the subprocesses.

In summary, this thesis investigates the effect of bilingualism on speakers' executive control from three aspects: bilingualism and executive control in non-verbal tasks, bilingualism and executive control in verbal tasks, and bilingualism and the executive control network. At the end of the thesis, we will summarize the main findings. We will also revisit the scope of the thesis and its potential limitations. Last, but not least, suggestions for future research will be made, towards a better understanding of the amazing phenomena of bilingualism and its cognitive consequences.

CHAPTER 2

DATA TRIMMING PROCEDURE CAN ELIMINATE BILINGUAL COGNITIVE ADVANTAGE

Abstract

Bilingualism and its cognitive impacts have drawn increasing interest. Recently, inconsistencies in the findings have raised discussions on what might have caused such discrepancies and how evidence should be evaluated. This review tries to shed new light onto the reasons for the inconsistencies by taking a novel perspective. Motivated by the finding that bilingualism affects response time distribution profiles, particularly findings that suggest bilinguals have fewer long responses, we investigated the relation between maximum response times allowed/included in the analysis of an experiment and the finding of a bilingual advantage. We reviewed 68 experiments from 33 articles that compared monolingual and bilingual speakers' performance in three commonly used non-verbal interference tasks (Simon, Spatial Stroop and Flanker). We found that studies that included longer responses in their analysis were more likely to report a bilingualism effect. We conclude that seemingly insignificant details such as the data trimming procedure can have a potential impact on whether an effect is observed. We also discuss the implication of our findings and suggest the usefulness of more fine-grid analytical procedures.

Introduction

Recent years have seen an ever-increasing interest in bilingualism, especially in how the bilingual experience leads to cognitive changes (Kroll & Bialystok, 2013). One of the key discoveries was the finding that the two languages of a bilingual speaker are always active, even if the speaker is using only one language in a particular situation (BijeljacBabic, Biardeau, & Grainger, 1997; Colome, 2001; Spivey & Marian, 1999; Hoshino & Thierry, 2012; Wu & Thierry, 2010). Therefore, bilingual experience requires that the speaker constantly monitors and controls language choice. But the soaring interest in bilingualism rather stems from reports that the constant demand on control processes appears to lead to a bilingual cognitive control advantage (for a review see Bialystok et al., 2009). Possibly the most exciting finding was that the onset of dementia appears to be later for bilingual speakers than for monolingual speakers (Alladi et al., 2013; Bialystok, Craik, & Freedman, 2007), suggesting that bilingualism provides a cognitive reserve. Another reason for the interest in bilingualism is the fact that evidence regarding the bilingual cognitive advantage is inconclusive. Quite a large number of studies have reported a bilingual advantage in cognitive control. But recent developments in the field have suggested that the evidence concerning the bilingual advantage, especially in non-verbal inhibition tasks, is far from conclusive (Hilchey & Klein, 2011; Paap, 2014; Paap & Greenberg, 2013).

Evidence For and Against Bilingual Advantage

Early studies investigating speakers' cognitive control abilities have reported a bilingual advantage. These studies utilized the Simon task (Simon & Rudell, 1967), a task in which participants have to respond to a stimulus feature (e.g. respond according to the colour of a stimulus, left hand for red, and right hand for blue). The position of the stimulus could be either compatible or incompatible with the response hand. In this paradigm, responses are typically slower when stimulus position and response hand are incompatible. Bilingual speakers outperformed monolingual speakers in this task: they showed smaller stimulus congruency effects and/or overall faster response times (Bialystok, Craik, et al., 2005; Bialystok, Craik, Klein, & Viswanathan, 2004; Bialystok, Martin, & Viswanathan, 2005). Using other variants of the Simon task (e.g., Spatial Stroop task, Bialystok, 2006) or paradigms that also require interference inhibition/conflict resolution (e.g., Flanker task, Eriksen & Eriksen, 1974), such bilingual advantage over monolingual performance has been reported many times over the years since the first report (e.g. Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009; Costa, Hernandez, & Sebastian-Galles, 2008; Kapa & Colombo, 2013; Tao, Marzecova, Taft, Asanowicz, & Wodniecka, 2011).

But the evidence is not consistent. While a bilingual advantage is frequently reported for bilingual elderly, it is less consistent in children and young adults (for a review see Hilchey & Klein, 2011). Most strikingly, in a recent comprehensive study, Paap and Greenberg (2013) did not find any evidence for the bilingual advantage. They conducted a series of non-verbal conflict tasks that had commonly been used in previous studies, including the Simon task, with monolingual and bilingual college students. Bilingual speakers were neither less vulnerable in the conflict condition nor faster overall. On the contrary, the only group difference pointed to a bilingual disadvantage.

These inconsistent findings have raised serious discussions about the nature of the bilingualism effect and thought-provoking debates on how the evidence should be evaluated (e.g. Kroll & Bialystok, 2013; Paap, 2014; Valian, 2015). Several reviews have drawn attention to the published and non-published null results regarding the bilingualism effect, arguing that one should not over-evaluate the significance of positive findings, and should not under-evaluate the meaning of null results. The series of discussions provided a great chance to reflect on the current status in this research field, and more importantly on where the field is heading. Bearing that in mind, one constructive way to enhance our understanding of the consequences of bilingualism is to understand what factor(s) drives the divergence of results. The focus of this article is to contribute to the discussion from a novel perspective, i.e. the impact of seemingly trivial data trimming procedures.

Factors That Potentially Drive the Inconsistency

In order to shed light onto the reasons for the inconsistencies in the literature, it is important to understand how other factors might interact with speakers' cognitive control ability. Quite a number of factors have been pointed out. It was evident from the beginning that bilingual research is challenging due to the diversity of speakers' linguistic profiles and experiences (Bialystok, 2001; Grosjean, 1998). Depending on their life experience, one bilingual speaker can differ from another one in many ways. Such heterogeneity in linguistic experiences has been shown to have led to diverse cognitive consequences, such as level of language proficiency (e.g. Mishra, Hilchey, Singh, & Klein, 2012), stage of second language acquisition (early bilingual VS late bilingual, Kalia, Wilbourn, & Ghio, 2014), the degree of bilingualism (dominant VS balanced bilingual, Goral, Campanelli, & Spiro, 2015), pattern of language use, varying experience with frequent language switch (Soveri, Rodriguez-Fornells, & Laine, 2011), the similarity between a bilingual speakers' two languages (Coderre & van Heuven, 2014) and multilingualism (Poarch & van Hell, 2012). In addition, there are factors that are closely related to bilingualism or factors that drive the different language experiences, which at the same time are related to general cognitive performances. These include social and economic status (Morton & Harper, 2007), different cultural backgrounds (Yang, Yang, & Lust, 2011) and immigration status. Last but not least there are factors that affect one's general executive functioning and that probably affect monolingual and bilingual speakers in the same way, such as age, education, exercise, music training, active video game experience and others (for an overview see Valian, 2015). These latter factors emphasize that cognitive control can be trained in other ways than by being bilingual and that the populations of monolinguals and bilinguals can substantially overlap with regards to their performance in cognitive tasks. Such an overlap would also explain why the bilingualism effect has not been found in every study.

A New Proposal

Another reason for the inconclusiveness of the literature might be the nature of the bilingual cognitive effect, which is better described as a mixture of effects rather than a single one. For instance, Hilchey and Klein (2011) differentiated between two patterns of bilingual advantage, namely an inhibitory control advantage (i.e. bilinguals showing a reduced conflict effect) and an overall response speed advantage. This distinction suggests two routes through which bilingual experience could affect cognitive control: inhibitory control and attentional control. While enhanced inhibitory control ability should help to resolve conflict, resulting in a reduced conflict effect; enhanced attentional control should help to maintain task goals, leading to an overall speed advantage. Due to the general impurity of cognitive control tasks, a specific task does not provide a pure measure of a single control ability, but draws on many aspects of cognitive control. Some tasks might be more sensitive to participants' inhibitory control ability and some to their attentional control ability. Therefore depending on the task, one might observe a result pattern rather consistent with a bilingual inhibitory control advantage and/or bilingual attentional control advantage.

Tse and Altarriba (2012) utilized a novel analytical approach to the bilingual advantage effect. They performed an ex-Gaussian analysis to investigate response time distributions of bilingual speakers' performance in a Colour Stroop task. Response times in cognitive experiments typically present themselves in a positively skewed distribution, which can be approximated by an ex-Gaussian distribution (Heathcote, Popiel, & Mewhort, 1991), i.e. a convolution of a Gaussian distribution (the mean of which is captured by the parameter μ) and an exponential distribution (the mean and variance of which are captured by the parameter τ). While the Gaussian component (the parameter μ) can be understood as the main body of the distribution, the exponential component (the parameter τ) captures the tail of the distribution, i.e. extremely slow responses. Tse and Altarriba (2012) suggested that the parameters of ex-Gaussian models of response time distributions in a Colour Stroop experiment are differentially sensitive to inhibitory control and attentional control. They argued that inhibitory control ability modulates the Gaussian component (μ) because differences in inhibitory control would affect the ease with which one resolves the competition between the two conflicting responses, leading to an overall shift of the response distribution in the conflict condition as compared to the no-conflict condition. In contrast, attentional control ability modulates the tail of the distribution (τ) because lapses of attention should lead to extreme long responses independent of condition. They found that more proficient speakers in L1/L2 showed a smaller interference effect in μ, and also smaller τ independent of condition. This result is important, first because it proposes a way to disentangle the contribution of inhibitory and attentional control, and second because it suggests that what is usually treated as unwanted responses (τ) conveys important information.

One other study that has utilized the ex-Gaussian approach to examine response time distributions is Calabria, Hernandez, Martin, and Costa (2011). They re-analysed results from an Attentional Network Task (ANT) originally reported in Costa et al. (2008) and Costa et al. (2009). In the original studies, they tested participants' attentional networks using the ANT, which generates measurements for three attentional networks: an alerting network, an orienting network and an executive network. In their re-analysis, they focused on the executive network, or more specifically, on response times for trials with conflict stimuli versus trials with non-conflict stimuli (regardless of cue type). Results revealed an overall speed advantage for bilinguals in both the Gaussian (μ) and exponential (τ) components of the response distributions. Also, for an experiment that contained only 25% inconsistent trials, i.e. under high monitoring demands, monolinguals had significant longer distribution tails (larger τ) in the incongruent compared to the congruent condition. This congruency effect was absent for bilinguals. These results suggest that the conflict effect in interference tasks is at least partially located in the tail of response distributions.

These observations lead to a new proposal that we investigated in the present review, namely that one reason for observing or not observing the bilingual advantage might be how the data was handled with regards to slow responses. In a traditional central tendency analysis, the typical procedure is to trim extreme responses, treating them as outliers. This is problematic if long responses are most sensitive to the experimental manipulation and/or group differences, as in studies where group/condition differences only emerged in the tail of response distributions (e.g. Epstein et al., 2011; Hervey et al., 2006). In other words, if a difference resides in the tail of the response distribution, by trimming the tail one also trims the potential to observe an effect. For instance, Leth-Steensen, Elbaz, and Douglas (2000) investigated response time distributions in a four-choice reaction time task for a group of children diagnosed with ADHD and a matched group with typical developing children. Using an ex-Gaussian analysis, they found that the two groups' performances differed only in the τ parameter (the distribution tails), not in the main part of the response time distribution. The authors concluded that data trimming in this situation is equivalent to an artificial elimination of effects.

Meta-analysis

In what follows, we present a new perspective on previous studies that compared monolingual and bilingual performance in non-verbal inhibition tasks, investigating their data trimming procedures. If bilingual advantages are at least partly located in the tails of response distributions, i.e. in the slow responses as in Calabria et al. (2011), then one would expect that cutting off slow responses would reduce the chance of finding such advantages. We focused on three non-linguistic inhibition tasks that have been most intensely used to investigate the bilingual advantage: the Simon task, the Spatial Stroop task and the Flanker task (the latter sometimes embedded in an ANT). To ensure comparability, some variations of the tasks were excluded (e.g. the Simon task with a delay component in Martin-Rhee & Bialystok, 2008). We found 68 experiments taken from 33 articles (see Table 2.1). Within these 68 experiments, 23 (34%) reported data trimming procedures, 4 reported excluding very short responses but did not report trimming of response distribution tails, 1 stated explicitly that long responses were not trimmed and the remaining 40 did not mention whether or not they trimmed the data. For the purpose of our analyses, we treated the latter studies as ones that did not trim the data.

Table 2.1

Data Trimming Procedures and Results of Studies Using Non-verbal Interference Tasks in Bilingual Research

Study	Age Group (mean age)	Task	Trimming Procedure	Maximum Time Allowed	Results
Bialystok et al. (2004)	Adults (43)	Simon (study1)	-	Until Response	Overall/Effect[1]
	Elderly(71.9)	Simon (study 1)	-	Until Response	Overall/Effect
	Adults (42.6)	Simon (study 2)	-	Until Response	Overall/Effect
		Simon (4 colours)	-	Until Response	Overall/Effect
	Elderly(70.3)	Simon (study 2)	-	Until Response	Overall/Effect
		Simon (4 colours)	-	Until Response	Overall/Effect
	Adults (~39)	Simon	-	Until Response	Overall/Effect
Bialystok, Martin, et al. (2005)[2]	Children (5)	Simon (study 1)	-	5500	Overall
	Children (5)	Simon (study 2)	-	5500	Overall
	Young Adults (20-30)	Simon	-	Not mentioned	No group difference
Bialystok, Craik, et al. (2005)	Young Adults (29)	Simon	-	1800	No group difference
	Young Adults (29)	Simon	-	1800	Overall[3]
Bialystok (2006)	Young Adults (~22)	Simon (low switch)	-	1000	No group difference
		Simon (high switch)	-	1000	No group difference
		Spatial Stroop (low switch)	-	1000	No group difference
		Spatial Stroop (high switch)	-	1000	Overall
Morton and Harper (2007)	Children (6.8)	Simon	None	Until Response[4]	No group difference
Linck, Hoshino, and Kroll (2008)	Young Adults (~20)	Simon	-	Until Response[4]	Effect
Martin-Rhee and Bialystok (2008)[5]	Children (4.6)	Simon (study 1)	-	5000	Overall
	Children (4.2)	Simon (study 2)	-	5000	Overall
	Children (8)	Spatial Stroop	-	Until Response	Overall
Bialystok, Craik, and Luk (2008)	Young Adults (~20)	Spatial Stroop (high switch)	-	Until Response	No group difference
	Elderly(~68)	Spatial Stroop (high switch)	-	Until Response	Effect
Carlson and Meltzoff (2008)	Children (~6)	ANT	> 1700 ms	Not mentioned	N/A[6]
Costa et al. (2008)	Young Adults (22)	ANT	-	1700	Overall/Effect

Costa et al. (2009)	Young Adults (~20)	ANT (92% cong)	>1200 ms	1700	No group difference
	Young Adults (~20)	ANT (8% cong)	>1200 ms	1700	No group difference
	Young Adults (~20)	ANT (50% cong)	>1200 ms	1700	Overall
	Young Adults (~20)	ANT (75% cong)	>1200 ms	1700	Overall/Effect (Block1)
Emmorey, Luk, Pyers, and Bialystok (2008)	Adults (46-50)	Flanker	none	2000	Overall (for unimodal bilingual, not bimodal)
Bialystok and DePape (2009)	Young Adults (~23)	Spatial Stroop	>2500 ms	Until Response	Overall (for no music experience group)
Luk, Anderson, Craik, Grady, and Bialystok (2010)	Young Adutls (~21)	Flanker (with Go/No-Go element)	-	1300[7]	No group difference
Tao et al. (2011)	Young Adults (~20)	Lateralized ANT	>1200 ms	1820	Effect
Yang et al. (2011)	Children (~6.5)	ANT	-	Not mentioned	Overall
	Children (~6.5)	ANT	-	Not mentioned	Overall
	Children (~6.5)	ANT	-	Not mentioned	Overall
Salvatierra and Rosselli (2011)	Young Adults (~26)	Simon	-	800	No group difference
	Young Adults (~26)	Simon (4 colours)	-	800	No group difference
	Elderly (~64)	Simon	-	800	Effect
	Elderly (~64)	Simon (4 colours)	-	800	No group difference
Yudes et al. (2011)	Young Adults (21-25)	Simon	>1200 ms	2000	No group difference[8]
Engel de Abreu, Cruz-Santos, Tourinho, Martin, and Bialystok (2012)	Children (~8.5)	Flanker	>3 *SD*	5000	Overall/Effect
Kousaie and Phillips (2012)	Young Adults (~24)	Simon	-	750	No group difference
	Young Adults (~24)	Flanker	-	750	No group difference
Poarch and van Hell (2012)	Children (~7)	Simon (study1)	>2.5 *SD*	1000	No group difference[9]
Kapa and Colombo (2013)	Children (~9)	ANT	None	1700	Overall (Monolingual = Late bilingual > Early bilingual)
Nicolay and Poncelet (2013)	Children (~8.5)	Flanker (50% cong)	-	1700[10]	No group difference
Paap and Greensberg (2013)	Young Adults 1 (students)	Simon1	>2.5 *SD*	Not mentioned	No group difference
	Young Adults 2 (students)	Simon2	>2.5 *SD*	Not mentioned	No group difference

	Young Adults 3 (students)	Simon3	>2.5 *SD*	Not mentioned	No group difference
	Young Adults 3 (students)	Flanker	>2.5 *SD*	1700	No group difference
Kirk et al. (2014)	Elderly (~70.3)	Simon	>2.5 *SD*	1000	No group difference
			None[11]	1000	No group difference
Marzecova et al. (2013)	Young Adults (~21)	LANT	>1200 ms	2000	Effect
Gathercole et al. (2014)	Children (~4)	Simon	-	Not mentioned	Monolingual advantage
	Grade Schoolers (~8)	Simon	-	Not mentioned	No group difference
	Young Adults (~25)	Simon	-	Not mentioned	Monolingual advantage
	Elderly (~67)	Simon	-	Not mentioned	No group difference
Mohades et al. (2014)	Children (8~11)	Simon	-	Not mentioned	Monolingual advantage
Blumenfeld and Marian (2014)[12]	Young Adults (~22)	Spatial Stroop (study1)	>2.5 *SD*	700	No group difference
	Young Adults (~22)	Simon (up/down arrow in study1)	>2.5 *SD*	700	No group difference
	Young Adults (~22)	Spatial Stroop (study2)	>2.5 *SD*	700	No group difference
	Young Adults (~22)	Simon (up/down arrow in study2)	>2.5 *SD*	700	No group difference
Pelham and Abrams (2014)	Young Adults (~20)	ANT	>2.5 *SD*	1700	Effect (Monolingual > Late bilingual = Early bilingual)
Kousaie, Sheppard, Lemieux, Monetta, and Taler (2014)	Young Adults (~21.5)	Spatial Stroop	>2.5 *SD*	Not mentioned	No group difference
	Elderly (~72.5)	Spatial Stroop	>2.5 *SD*	Not mentioned	No group difference
Grady, Luk, Craik, and Bialystok (2015)	Elderly (~70.5)	Simon	-	2550	No group difference
Abutalebi et al. (2015)	Elderly (~62)	Flanker	None	1700	Overall

Note: 1. Overall = Bilingual advantage in overall response speed (in both congruent and incongruent condition). Effect = Smaller congruency effect for bilinguals. 2. Study 4 and 5 in this article were the same as in Bialystok (2004); therefore we do not duplicate report here. 3. Bilingual overall advantage was found in the Cantonese/English bilingual group but not the French/English bilingual group; this is represented as two separate comparisons here. 4. Not stated directly, but followed the procedure by Bialystok et al. (2004). 5. Another two variations of the Simon task were conducted, with a short or a long delay. These were not included here. 6. A bilingual advantage was reported for response accuracies. Information regarding response speed is not available. 7. Assuming that responses can be made during the blank interval of 300 ms after the stimulus presentation for 1000 ms. 8. Yudes et al. (2011) also tested a group of simultaneous interpreters. We only report the monolingual/bilingual comparison. 9. A group of second language learners and a group of trilingual children were also tested. We only report the monolingual/bilingual comparison. 10. Not stated directly, but followed the procedure by Costa et al. (2008). 11. Kirk (2014) tested three monolingual control groups. We only report results regarding Angalo-English monolingual group which does not speak another dialect. Kirk (2014) also reported analysis without removing outlying responses and found no group difference. 12. Blumenfeld & Viorica (2014) also reported reverse efficiency scores which were based on both response latencies and accuracy rates. This analysis yielded slightly different result patterns. We focus on response latencies for reasons of comparability.

Two issues need to be pointed out. First, for studies that did trim the data, the practices differed. While some studies rejected long responses using standard deviations (e.g. 2.5 *SD* in Paap & Greenberg, 2013, which was approximately 700 ms after stimulus onset), others used a specific time cut-off (e.g. response times above 1700 ms). For comparison purposes, we translated cut-offs based on standard deviations into time cut-offs (using reported means and standard deviations). Second, for studies that did not trim the data, there was a big variation in terms of the maximum time allowed for making a response. For example, in Kousaie and Phillips (2012) 750 ms were allowed for making a response, meaning that in this particular design it was not possible to observe responses slower than 750 ms. This is equivalent to trimming the data at 750 ms. For these reasons, we focused on maximum response times being included in analyses (Figure 2.1). For studies that did report a data trimming procedure, this is either the explicitly stated cut-off time or the RT calculated by the mean and *SD*. For studies that did not report a data trimming procedure, this was the maximum time allowed for making a response. For simplicity reasons, we will refer to both types of trimming as the maximum response time allowance. Some of the studies reviewed were not included into further analysis because there was not adequate information about either the maximum time allowed or how the data was treated (the 3rd study in Bialystok, Martin, et al., 2005, all studies in Gathercole et al., 2014, Yang et al., 2011, and Mohades et al., 2014). Carlson and Meltzoff (2008) was not included either because they did not report RTs. This led to 58 studies being included in our statistical analysis.

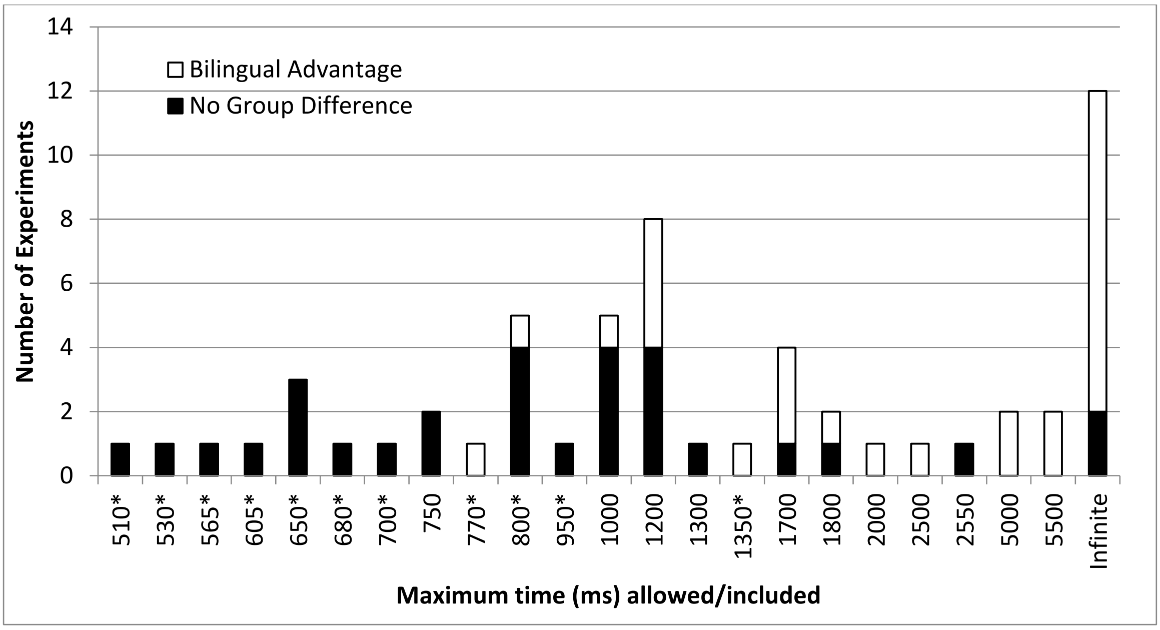


*Figure 2.1. Reported bilingual advantage for maximum response times included in the analyses in 58 non-verbal conflict experiments. * indicates that the cut-off time was estimated by the mean and SD.*

Figure 2.1 shows the number of experiments that did and did not find a bilingual advantage for the various maximum times allowed. A clear pattern emerges: the shorter the maximum response time allowance, the less likely a bilingualism effect was found. In order to statistically test whether observing a bilingualism effect depends on the maximum time allowed, we grouped the experiments into three types of maximum time allowance: short allowance (below 1000ms), medium allowance (1001ms – 3000ms) and long allowance (above 3001ms; see Table 2.2). A chi-square test of independence confirmed that the result patterns differed for the three allowance groups, χ^2 (2, N = 58) = 21.99, p < .001. Thus, consistent with our hypothesis, studies with short response time allowance were more likely to report no group difference whereas studies with longer allowance were more likely to report a bilingualism effect.

Table 2.2

Number of Experiments with Short, Medium or Long Cut-offs / Maximum Allowed Times and Findings of Bilingual Advantage

Cut-off Group	Children		Adults		Elderly		Overall	
	No Difference	Bilingual advantage	No Difference	Bilingual advantage	No Difference	Bilingual advantage	No Difference	Bilingual advantage
Short Allowance	1	0	16	2	3	1	20	3
Medium Allowance	1	2	5	8	2	1	8	11
Long Allowance	1	5	1	5	0	4	2	14
Column Total	3	7	22	15	5	5	30	28

As introduced, it has been pointed out that the age of the participants might play a role in whether a bilingualism effect can be detected or not. More specifically, bilingual elderly have been found to show the cognitive control advantage more consistently than other age groups. Therefore, we also tested the relationship between maximum response time allowance and the likelihood of finding a bilingualism effect in children, adults and elderly participants separately (see Table 2.2). A chi-square test of independence showed that while the relationship was not significant for children, χ^2 (2, N = 10) = 2.86, p > .05, it was significant for adults, χ^2 (2, N = 37) = 13.40, p = .001, and was marginally significant for elderly participants, χ^2 (2, N = 11) = 5.29, p = .071. The non-significant result for the children might be due to the fact that there was only one study that fell in the short allowance category. However the results pattern is clear: the short allowance study did not find any bilingualism effect whereas studies with medium and long allowances reported an effect more often. Similarly in the elderly group, studies with longer allowances are more likely to report a bilingualism effect descriptively. The marginal effect might be due to a small sample size. In summary, it appears that data trimming reduces the likelihood of observing a bilingualism effect regardless of age group.

Discussion

Based on the assumption that the effect of bilingualism on cognitive control resides at least partly in the tail of response distributions, we investigated a potential relationship between data trimming procedures adopted and the likelihood of observing a bilingualism effect by reviewing 68 experiments reported in 33 articles that compared monolingual and bilingual speakers using non-verbal interference tasks. We found that studies that included longer responses in their analysis were more likely to report a bilingualism effect, either in the form of overall response speed advantage or in the form of reduced interference effect.

This is consistent with earlier findings that the bilingualism effect emerges partially in the tail of response distributions (Abutalebi et al., 2015; Calabria et al., 2011; Tse & Altarriba, 2012). It appears that, when these prolonged responses were trimmed or not recorded, group differences might have also been eliminated.

A further analysis showed that the general result pattern was true for studies testing younger and older adults alike, suggesting that data trimming might be problematic independent of the age of the participants and therefore independent of the average response times or the cognitive abilities of the participants. There is unfortunately an insufficient number of studies and/or information about the participants in studies published to date to test whether data trimming procedures affect results independent of other factors that have been suggested to interact with the bilingual advantage (such as SES, immigrant status, language dominance, age of language acquisition, language usage etc.). Future studies will need to disentangle the relevance of and the potential interplay of all these factors.

This review provides some practical implications for future endeavours. Employing a more fine-grid investigation approach might be useful, particularly in situations where effects are subtle. As pointed out, one fruitful alternative to traditional approaches of analysis is the ex-Gaussian analysis of response time distributions (Abutalebi et al., 2015; Calabria et al., 2011). For instance, Abutalebi et al. (2015) reported that a group of bilingual elderly showed advantage in the τ component in the incongruent condition and the μ component in the congruent condition in a Flanker task, supporting the notion that bilingual speakers have enhanced attentional control. It has to be pointed out, though, that one needs to be cautious when interpreting the results from ex-Gaussian parameters in terms of underlying cognitive processes, because there is no one-to-one mapping between cognitive processes and those parameters (Matzke & Wagenmakers, 2009).

An alternative approach to the ex-Gaussian analysis is a delta plot analysis, which examines condition effects as a function of RT and which has been used in analysing conflict tasks such as the Simon task and the Colour Stroop task (Ridderinkhof, 2002a, 2002b; Ridderinkhof, Wildenberg, Wijnen, & Burle, 2004). In delta plot analyses, responses for each condition are grouped into bins according to their response times. Condition differences are then calculated and plotted for these bins. Delta plots prototypically have a positive slope due to effect sizes being larger for slower responses (Roelofs, Piai, & Garrido Rodriguez, 2011). If one condition requires more inhibition than the other, the difference between conditions does not linearly increase with RT, but is reduced instead, resulting in 'levelling-off' of the delta plot in longer RTs. The levelling-off has been explained by inhibition building up slowly (Ridderinkhof, 2002a). The extent to which the plots level off can effectively reflect the amount of inhibition involved. The stronger the inhibition is applied, the smaller the slope of the plot is. For example, Ridderinkhof (2002b) observed that in a Simon task, delta plots for participants with smaller Simon effect, who are believed to have more efficient inhibitory control, levelled off more than those with larger Simon effect. Because a delta plot analysis is most useful within the discussion of inhibitory control, this approach could be used to test whether inhibition was applied equally fast and to the same degree in monolinguals and bilinguals.

In conclusion, the current review adds to the discussion about the reality of the bilingualism cognitive advantage in that seemingly insignificant details such as data trimming and maximum time allowed for response might have a significant influence on the findings. Therefore, it is important to take these into account in order to fully judge the evidence for and against a bilingual cognitive effect. In addition, future studies are encouraged to report in detail how data were handled and possibly use more fine-grid analyses of RT data to shed light onto the effect of bilingualism on speakers' cognitive control abilities.

CHAPTER 3

BILINGUALISM ENHANCES ATTENTIONAL CONTROL IN CONFLICT TASKS – EVIDENCE FROM EX-GAUSSIAN ANALYSES

Abstract

Bilinguals have been found to possess cognitive advantages. But the nature of this advantage is unclear. While some evidence suggests that bilinguals have developed enhanced inhibitory control abilities, other evidence suggests that they possess enhanced attentional control abilities. In the current study, English monolingual and English/Chinese bilingual young adults were tested in three conflict tasks (the Flanker task, the Spatial Stroop task and the Simon task). Ex-Gaussian analyses were utilized to inspect response time distributions. The two participant groups showed comparable effects of stimulus-response congruency on the main part of response distributions (μ), but different effects on the distribution tails (τ). These results suggest that bilingual advantage mainly emerge from better attentional control. The usefulness of ex-Gaussian analyses was also discussed.

Introduction

This chapter strives to elucidate the nature of the bilingualism cognitive effect through exploring how different abilities contribute to it. The bilingual advantage in conflict tasks such as the Simon task (Simon & Rudell, 1967) has generally been related to bilinguals' enhanced executive control (EC) (Bialystok, 2001; Bialystok et al., 2009; Hilchey & Klein, 2011). As mentioned briefly in the previous chapter, the bilingualism effect might be a mixture of effects instead of a unified one. It potentially encompasses different aspects of executive control, such as inhibitory control and attentional control. Empirical evidence is mixed and inconclusive as to which of these control processes contribute to the advantage. By using the ex-Gaussian analytical approach, the current study aims to tease apart the contribution of inhibitory control and attentional control.

In what follows, we first introduce the evidence that enhanced inhibitory control and/or enhanced attentional control drives the bilingual advantage in non-verbal conflict tasks. We then introduce the ex-Gaussian analysis and explain why it has the potential to tease the two accounts apart.

Bilingual Advantage in Conflict Tasks

Chapter 2 (see Table 2.1) reviewed findings in three non-verbal interference tasks that have been used most often to investigate the bilingual cognitive advantage, namely the Simon task, the Spatial Stroop task and the flanker task (sometimes embedded in an Attentional Network Task or ANT). In the Simon task (Simon & Rudell, 1967), stimuli are presented either in a spatially compatible or incompatible way with the response hand. For instance, if a red square requests a right hand response, then a presentation on the right side of the screen is compatible with the response hand, but a presentation on the left side is incompatible. The classic Simon effect (congruency effect) refers to the finding that participants are slower when the position of the stimulus is not compatible with the response hand, suggesting that extra effort is required to resolve such spatial incompatibility and to overcome the conflict.

Result patterns in the Simon, Spatial Stroop, and Flanker task have been mixed. Bilingual speakers sometimes outperforme monolingual speakers in that they have smaller congruency effects, sometimes their responses are faster overall, sometimes both patterns are present, and other times no behavioural differences between the two groups are reported. When a bilingual advantage was found in non-verbal conflict tasks, it has been interpreted in various ways: as enhanced inhibitory control ability (e.g. Bialystok et al., 2004; Carlson & Meltzoff, 2008; Luk et al., 2010), as enhanced attentional control ability (e.g. Costa et al., 2008) or, very broadly, as enhanced executive functioning (Bialystok, Martin, et al., 2005; Bialystok et al. 2004). Here we focused on enhanced inhibitory control and enhanced attentional control as two candidates for the nature of the bilingual advantage. We next elaborate on these two terms before moving on to discuss evidence in support of each.

Inhibitory control is required when conflicting mental representations lead to different responses. Efficient inhibitory control would result in successful conflict resolution. We will use the term inhibitory control to refer to the processes involved in detecting and resolving conflict. Attentional control refers to a more general function that is involved in tasks requiring moderate focus of attention, not necessarily conditions that involve conflict. Within attentional control, two aspects appear to us of special importance for understanding bilingual advantage. The first is the alertness aspect of attention, proposed by Posner and Petersen (1990) as one of three aspects of attention (next to selection and orientation). Alertness refers to the sustaining attention or the keeping of vigilant attention (Robertson & Garavan, 2004). Here we focus on the ability to actively sustain or engage attention in task performance, or more concretely, the ability to maintain task goals in working memory. To put it differently, failing to sustain such attention would lead to a temporary lapse of attention or a temporary loss of task goals from the working memory. Increased attentional alertness to the task goal by bilingual speakers can, for instance, explain better performance in conditions of high cognitive demand (Costa et al., 2009). The second aspect of attentional control that appears to be of special importance here is attentional monitoring as discussed by Hilchey and Klein (2011), i.e. the ability to flexibly increase/decrease the degree of attentional engagement depending on the context.

One way that executive control might contribute to the bilingual advantage is through enhanced inhibitory control. The smaller congruency effect sometimes observed for bilinguals in interference tasks has been taken as evidence to support this notion (Bialystok et al., 2008; Linck et al., 2008; Tao et al., 2011). In the incongruent condition two representations are active, and participants need to overcome a prepotent response activated by the misleading information. Therefore, a reduced congruency effect in bilinguals suggests that they have superior ability to inhibit prepotent responses.

Another way in which executive control might contribute to the bilingual advantage is through enhanced attentional control. One major source of evidence for this hypothesis is that bilinguals sometimes show similar congruency effects to monolinguals in interference tasks but perform overall faster than monolinguals (Bialystok, 2006; Bialystok, Craik, et al., 2005; Bialystok & DePape, 2009; Emmorey et al., 2008; Engel de Abreu et al., 2012; Kapa & Colombo, 2013). The finding that bilinguals are faster in all conditions implies that the advantage lies within general cognitive processing, for example the ability to maintain task goals or to attend to goal-relevant information. In addition, there is evidence that bilinguals perform differently from monolinguals in tasks that tap attentional control abilities. For example, bilinguals have been found to show more rapid disengagement of attention (Colzato et al., 2008; Mishra et al., 2012). They have also been shown to benefit more from cues kept in working memory in a visual search task, suggesting bilingual enhanced top-down mechanisms of attentional control (Hernandez, Costa, & Humphreys, 2012). Furthermore, bilingual children have been found to be in general faster on a battery of tasks assessing alerting, auditory selective attention and divided attention (Nicolay & Poncelet, 2013). It is important to note that the two accounts of the bilingual advantage in non-verbal tasks are not mutually exclusive. In fact, sometimes both a smaller congruency effect and faster overall reactions have been observed (Bialystok et al., 2004; Bialystok, Martin, et al., 2005; Costa et al., 2009; Costa et al., 2008; Martin-Rhee & Bialystok, 2008), in line with contributions from both enhanced inhibitory control and attentional control. Furthermore, other findings suggest some interaction between general executive demand and the bilingual inhibitory advantage. Bilinguals sometimes only show a smaller interference effect in conflict conditions compared to monolinguals for tasks with elevated demand for controlled attention. For instance, bilingual speakers have been found to show a processing advantage only in a condition with high rate of response switches or with high monitoring demand (Bialystok, 2006; Costa et al., 2009).

Ex-Gaussian Analysis

The current study investigated the contribution of different executive control processes to the bilingual advantage in non-verbal interference tasks by using ex-Gaussian analyses of response time distributions (Heathcote et al., 1991; Schmiedek, Oberauer, Wilhelm, Suss, & Wittmann, 2007). Compared to a traditional analysis, this analysis not only provides a measure of the average level of processing speed, which is typically captured by the mean response time, it also produces a measure for extremely slow responses. As we will explain, these two measures can be mapped onto the two executive control processes discussed above (inhibitory control versus attentional control) and different results are to be expected depending on which of the executive control processes primarily contributes to the bilingual advantage.

When analysing response time (RT) as a processing index, it is very typical to focus on mean response times. The mean as the central tendency has been a convenient way to describe the overall performance of a participant group and to compare performance across participant groups or across experimental conditions. However, as pointed out by Balota and Yap (2011), such comparisons rely on the assumption that response times are normally distributed, which is often not the case. Especially in forced choice tasks, response distributions are typically positively skewed. Therefore, one commonly adopted procedure is to clean or to trim the raw data by removing outlying responses. The distribution tail is treated as ‘outlying responses’ because it is assumed to have abnormal underlying cognitive processes that deviate from those of average responses. However, given the pervasive existence of such long skewed tails, it is hard to deny that there is some commonality within the uncommonness. By ignoring information conveyed by the uncommon responses, the results or interpretations could be limited or at worst even misleading (Balota & Yap, 2011).

RT distribution analyses can improve this situation. There are various mathematical models available to describe RT distributions. The one that has repeatedly been found to produce excellent fit with empirical data and the one that has successfully been used for response conflict tasks is the ex-Gaussian distribution (Heathcote et al., 1991; Ratcliff, 1979; Schmiedek et al., 2007). The ex-Gaussian distribution results from the convolution of a Gaussian and an exponential distribution. Three parameters characterize the distribution: μ, σ and τ. The mean and variance of the Gaussian part are reflected by μ and σ, respectively; the mean and variance of the exponential part are reflected by τ. The overall mean of the ex-Gaussian distribution is the sum of μ and τ, the overall variance is the sum of σ^2 and τ^2. Figure 3.1 illustrates how the convolution of a Gaussian distribution (panel A) and an exponential distribution (panel B) creates a typical RT distribution (panel C). In what follows, we discuss what the parameters μ and τ represent with respect to response conflict tasks.

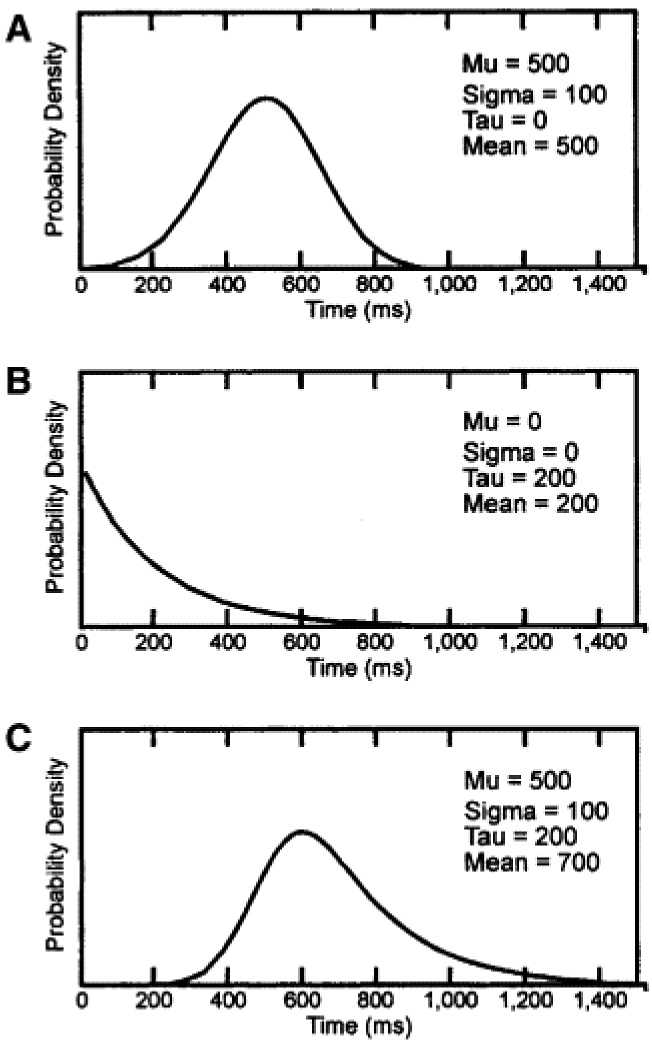


Figure 3.1. An example of a Gaussian (A) and an exponential (B) distribution and their convolution into an ex-Gaussian distribution (C). Figure taken from Balota & Spieler (1999).

The μ parameter.

The μ parameter captures the leading edge of an ex-Gaussian distribution. When τ is held constant, an increase in μ results in a positive shift of the distribution (Balota & Yap, 2011). See Figure 3.2 for an illustration (panels A and B or panels C and D). The μ parameter is directly affected by the experimental manipulation of a conflict task (Aarts, Roelofs, & van

Turennout, 2009; Heathcote et al., 1991; Hervey et al., 2006; Leth-Steensen et al., 2000; Spieler, Balota, & Faust, 2000; Tse, Balota, Yap, Duchek, & McCabe, 2010). In terms of a single response, it requires more processing time when conflict is present than when there is none. Therefore, in terms of the response distribution, prolonged response times in the conflict condition compared with the non-conflict condition should result in a positive shift of the distribution for the conflict condition without changing its shape.

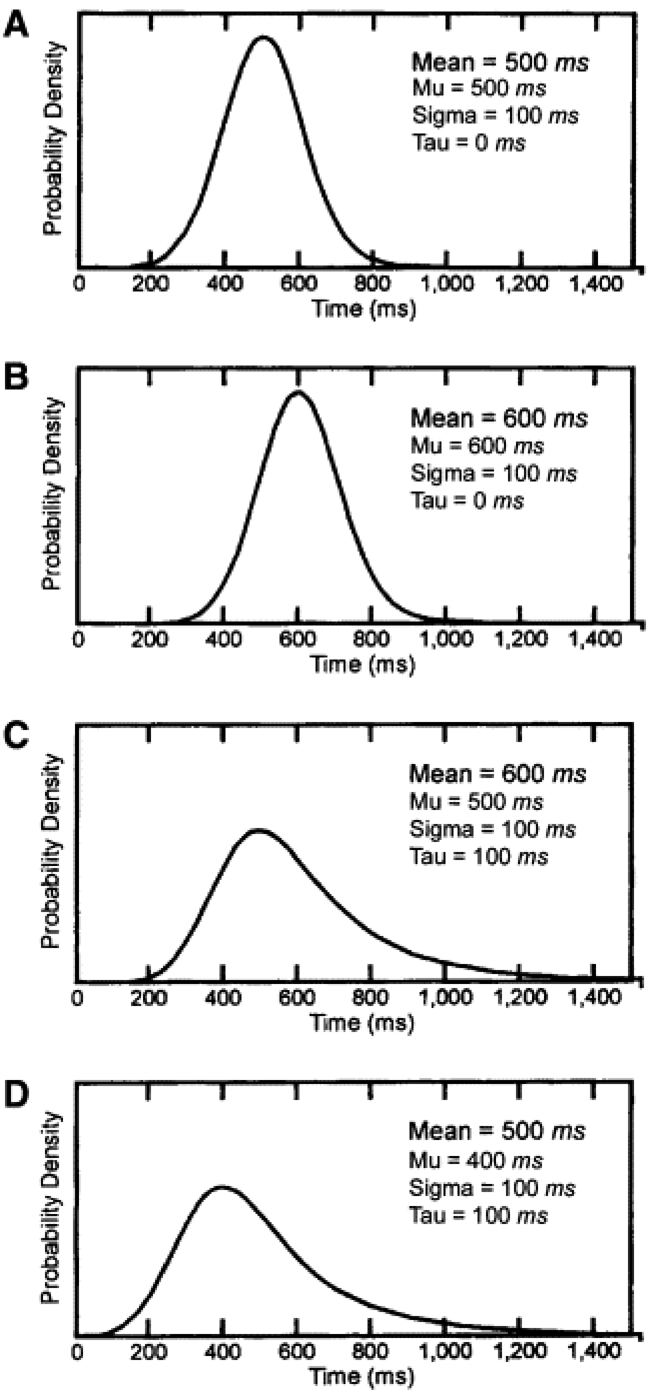


Figure 3.2. Possible changes in ex-Gaussian parameters and their effects on the distribution. Figure taken from Balota, Yap, Cortese, and Watson (2008).

Such a shift in distribution is exactly the pattern observed for conflict tasks. For example, for the response distribution of the Colour Stroop task, μ has been found to be significantly larger in the incongruent condition than in the congruent condition (de Zubicaray, McMahon, Eastburn, & Pringle, 2006; Heathcote et al., 1991; Spieler et al., 2000; Steinhauser & Hubner, 2009). The same pattern has also been observed for the Simon task (de Zubicaray et al., 2006), an adapted version of the Stroop task (similar to the Spatial Stroop task, Aarts et al., 2009), the Letter Flanker task (Blanco & Alvarez, 1994; Spieler et al., 2000) and an ANT adapted for children (Epstein et al., 2011). These all suggest that an increase in μ reflects the extra processing cost for majority responses in conflict conditions.

Following the logic that μ reflects the major delay in response when a subject encounters interference, one would expect that a participant group with superior conflict resolution ability would have a smaller congruency effect in μ. In other words, the distribution shift for a conflict compared to a non-conflict condition should be smaller for a group with superior conflict resolution ability. Therefore, if bilingual speakers have enhanced inhibitory control ability, we would expect that monolingual and bilingual speakers show similar μ in the congruent condition, but bilinguals should have a smaller μ in the incongruent condition than monolinguals.

The τ parameter.

The τ parameter, which reflects the tail of the response time distribution, has been found to be less affected by condition differences in conflict tasks. In the fore-mentioned studies, τ has been found not to differ between congruent and incongruent conditions. See the Colour Stroop task (Heathcote et al., 1991; Spieler et al., 2000), the adapted Spatial Stroop task (Aarts et al., 2009) and the Letter Flanker task (Spieler et al., 2000). Instead, τ is typically modulated by attentional control ability that is necessary to maintain the goal of a task. For example, significantly larger τ has been reported for individuals with attention deficit hyperactivity disorder (ADHD) compared to a healthy control group in various tasks, e.g. in the Conners' Continuous Performance Test (similar to a Go/No-go task) (Hervey et al., 2006), a button pressing task in response to a stimulus circle (Leth-Steensen et al., 2000), as well as in the Attentional Network task, Go/No-Go task, Stop-signal task and N-back task (Epstein et al., 2011). Similarly, individuals with very mild dementia of the Alzheimer's type have been found to show significantly larger τ than a healthy control group in a Colour Stroop task, a Simon task and a switching task (Tse et al., 2010).

An increased tail of the response time distribution has been argued to reflect poorer performance of the attentional control system that maintains task goals across time (Tse et al., 2010), or in other words, a momentary lapse of attention (Hervey et al., 2006; Leth-Steensen et al., 2000; Schmiedek et al., 2007; Shao, Roelofs, & Meyer, 2012; Unsworth, Redick, Lakey, & Young, 2010). This becomes clearer when zooming into the individual response level. A response is likely to be extremely slow when one temporarily loses track of the task goal, be it by attending to something outside the task or by being attracted to a non-relevant feature of the task. This might lead to a delayed processing of the stimulus and/or delayed initiation of a response (Unsworth et al., 2010). It might also lead to an initially wrong response decision that is made on the basis of the wrong information, which is subsequently corrected before the response is executed. When zooming out to the response distribution level, prolonged reactions are reflected in the tail of the distribution, i.e. in τ. It has been found that working memory measurements are positively correlated with the τ parameter (Schmiedek et al., 2007). This finding reinforces the interpretation of τ because working memory capacity can be conceptualized as an attentional control ability required for goal maintenance (Tse et al., 2010), and a lapse of attention occurs when such a control fails.

Since τ is driven by extreme responses, while capturing both the degree of extreme responses and the likelihood of such extreme cases, one would expect that people who are less likely to have temporary lapses of attention to have a smaller τ. Therefore, if bilinguals are better at attentional control, τ is expected to be smaller for bilinguals than monolinguals. And this should be the case regardless of the experimental condition (congruent or incongruent condition).

Ex-Gaussian Analysis and Bilingual Speakers

The meta-analysis in Chapter 2 suggests that at least part of the bilingual advantage in conflict tasks is carried by very slow responses or, more precisely, by bilinguals having fewer slow responses than monolinguals. This means that the bilingual effect should be visible at least partly in the tail of response time distributions, and it suggests that bilinguals might possess enhanced attentional control ability. To my knowledge, only two previous studies have investigated the performance of monolingual and bilingual speakers in conflict tasks with ex-Gaussian analyses: Calabria et al. (2011) and Abutalebi et al. (2015). Both studies showed that the bilingual advantage is at least partly carried by very slow responses and therefore supported the hypothesis of bilingual enhanced attentional control. But the picture is more complex.

Calabria et al. (2011) re-analyzed results of the Flanker component of the Attentional Network task originally reported in Costa et al. (2008) and Costa et al. (2009) by means of an ex-Gaussian analysis. Results showed that, in contrast to monolinguals, bilinguals had no congruency effect in τ when the experiment contained only 25% inconsistent trials (versus 33%). Therefore, bilinguals appear to sustain attention better when conflict situations are relatively rare. Their results also revealed an overall speed advantage for bilinguals, in both the Gaussian and the exponential part of response distributions. Calabria et al. (2011) speculated that the overall speed advantage in the Gaussian component might be due to advanced functioning of the monitoring system. That bilinguals were overall faster in the exponential component of the response distributions is in line with the assumption that bilinguals have better attentional control abilities. Because bilinguals did not show a reduced congruency effect in the Gaussian part of the response distributions, the results suggest that bilinguals might not have an advanced conflict resolution ability. However, while Calabria et al.'s (2011) data are very interesting, they did not investigate the role of attentional control. They were therefore not very concerned about very long responses and thus removed responses above 3 SD from the data. Although these are very few responses, it is these very long responses that can have a big impact on τ (since it is a reflection of the mean of the tail) and are therefore important if one wants to study the effect of attentional control.

Abutalebi et al. (2015) investigated the bilingual advantage in elderly using ex-Gaussian analyses, and they analyzed all responses without trimming the data. Using a Flanker task, they found a bilingual advantage in the τ component in the incongruent condition and the μ component in the congruent condition. However, in terms of statistical analysis, they focused on independent sample *t*-tests for each parameter and for each condition separately. Therefore, we do not have information about the main effect of Group as well as about the interaction between Group and Condition. Figure 2 in Abutalebi et al. (2015) showed that, descriptively, bilinguals had smaller τ than monolinguals overall, consistent with the enhanced attentional control hypothesis. The μ parameters were also descriptively smaller in bilinguals than in monolinguals for both conditions. The interaction between Group and Condition was not clear from the figure, therefore it is not possible to conclude whether evidence was present for enhanced inhibitory control.

While results from Calabria et al. (2011) and Abutalebi et al. (2015) are very promising, they used the same experiment paradigm. It is essential to investigate effects in various tasks to establish the generalizability of the findings. This is especially important because of the vast amount of contradictory findings in the literature with regard to a bilingual advantage in conflict tasks (see chapter 2). In addition, using different tasks can address the task impurity problem, that executive control tasks are always contaminated by other task demands.

Current Study

Therefore, in the current study English monolingual and English/Chinese bilingual young adults were tested in the Simon, Spatial Stroop and Flanker task. The two groups were compared in terms of accuracy rates, traditional condition means (with and without trimming very slow responses), and response distributions (μ and τ) in conflict and non-conflict conditions. This study focused on these three tasks because they are the tasks that have been most extensively tested in this area of research (see also Hilchey & Klein, 2011). In addition, they do not require any verbal responses (see Colour Stroop task). Because bilinguals have been found to be disadvantaged in naming tasks (e.g. bilinguals were slower than monolinguals when naming pictures in their dominant language, Gollan, Montoya, Fennema-Notestine, & Morris, 2005) their responses in verbal conflict tasks might not only be affected by their executive function abilities, but also their verbal abilities.

If inhibitory control underlies the bilingual advantage in interference tasks, bilinguals should have smaller congruency effects in the Gaussian components (μ) of the response distributions in all three tasks. If attentional control underlies the advantage, bilinguals should have shorter tails (τ) regardless of task. And this should be the case regardless of condition. However, as we have seen in Calabria et al. (2011), participants might be able to adjust their attentional control depending on task conditions. We might therefore see a stronger attentional control advantage in harder conditions, i.e. in incongruent conditions. Given the findings by Calabria et al. (2011) and Abulatebi et al. (2015), we might find both smaller congruency effects and shorter tails for bilinguals, which would imply that bilinguals possess both inhibitory and attentional contol advantages.

Method

Participants

Ninety-nine participants took part in the experiment: 51 monolingual native English speakers and 48 English/Chinese (Chinese/English) bilingual speakers. They were mostly undergraduate and postgraduate students at the University of Birmingham and participated either for course credits or cash. Apart from those, eight of the bilingual speakers were students of the Chinese University of Hong Kong and they were tested in Hong Kong and were paid for their participation.

For the analysis, participants were selected on the basis of their responses to a questionnaire about their language use history (see Appendix A-1), adapted from Silverberg & Samuel (2004). This questionnaire gathered demographic information such as age and education. Participants were also asked to rate their self-perceived proficiency in English and to list all the languages that they learnt or were able to speak, as well as the age at which they started to learn them. In addition, bilingual speakers were asked to rate their proficiency in Chinese. They also indicated their current language use pattern (e.g. using mainly one language or using both languages on a daily basis). To be classified as bilingual, the following criteria had to be met: the participant (a) learnt English and Chinese before age 10, (b) had more than 50% native-like proficiency in both languages, and (c) used both languages on a daily basis at the time of the experiment, either in the same setting or in different settings.

Monolingual English speakers were defined as follows: the participant (a) did not speak another language fluently (i.e. proficiency level of another language, if any, was below 50%), (b) did not speak another language on a daily basis, and (c) did not learn another language before age 10.

These criteria led to 29 English/Chinese bilingual speakers being included into the analyses. Twenty-two of those grew up with Chinese as their L1, two participants with English as L1, and five were simultaneous bilinguals. 29 monolingual English speakers were randomly selected from those that met the monolingual standard to match the bilinguals in age, t (50) = -.74, p = .46, and education, X^2 (1, N = 58) = 0, p = 1.0. In addition, bilingual speakers were equally proficient in English and Chinese, t (28) = .13, p > .05. See Table 3.1 for a summary of the demographic information of the two participant groups.

Table 3.1

Demographic Data of Monolingual and Bilingual Speakers

Variable	Monolingual	Bilingual
N (male/female)	29 (10/19)	29(12/17)
Mean Age (*SD*)	21.0 (3.0)	21.6 (3.2)
Undergraduate/Graduate[1]	25/4	25/4
Mean Age of English onset (*SD*)	Birth	3.3 (1.5)
Mean Age of Chinese onset (*SD*)	N/A	1.3 (1.8)
Speak L2 fluently	No	Yes
Speak L2 on a daily basis	No	Yes

Note. Undergraduate = students pursuing a bachelor’s degree or having been awarded a bachelor’s degree within the last 12 months. Graduate = students with a master’s degree or above.

General Design and Procedure

All participants went through the same sequence of tasks, namely Flanker task, Spatial Stroop task and Simon task. Participants then completed the language history questionnaire described above.

Flanker Task

Material.

Using the Erikson Flanker paradigm (Eriksen & Eriksen, 1974), the current study adapted the procedure by Costa et al. (2009). Each stimulus consisted of five arrows in a row, with the central arrow being the target and two arrows on each side being the flankers. Each arrow was approximately 0.55 degree in visual angle; distance between arrows was approximately 0.06 degree.

Procedure.

In this and the following tasks, participants were instructed to sit approximately 60 cm from the monitor. They pressed a left and a right button to indicate the direction of the central arrow using a Cedrus RB-834 response pad, which also measured response time. Each trial started with a fixation cross for 400 ms followed by the stimulus, which disappeared with the response or after 1700 ms in case of no response. Stimuli appeared randomly either above or below the fixation cross with a 50/50 chance of occurrence. In a congruent trial the central arrow and the flankers pointed to the same direction, in an incongruent trial they pointed to opposite directions. In order to increase the difficulty of the task, 75% of the trials were congruent and 25% were incongruent, which is equal to the high response monitoring condition in Costa et al. (2009). Twenty-four practice trials were followed by two blocks of 48 trials. The sequence of stimuli was randomized, with a different randomization for each participant.

Spatial Stroop Task

Material.

The Spatial Stroop task is a modified version of the Simon task. Adapting the design by Bialystok (2006), a single arrow was used as the stimulus, 6.5 cm in length with a tail of 0.5 cm in width. The widest point of the arrow was 1.5 cm.

Procedure.

Participants pressed a left or right button to indicate the direction of the arrow using a Cedrus RB-834 response pad. Each trial started with a fixation cross for 800 ms and a subsequent 250 ms blank screen. Then an arrow (pointing to the left or right) was presented 7 cm to the left or right of the fixation cross. The target disappeared with the response or after 1000 ms in case of no response, followed by a 500 ms blank screen. Each participant completed 24 practice trials and 64 test trials. Each combination of arrow and position had equal probability of occurrence and the stimuli were presented randomly, with a different randomization for each participant. In congruent trials, the arrow pointed to the same side as the presentation side on the screen (e.g. the arrow pointed to the right and was presented on the right side of the screen). In incongruent trials, the arrow pointed to the opposite side as the presentation side (e.g. the arrow pointed to the right and was presented on the left side of the screen).

Simon Task

Material.

Stimuli were red or blue squares (2.2 cm by 2.2cm).

Procedure.

The procedure was the same as that of the Spatial Stroop task, except that the stimuli were arranged into pre-determined pseudo-random orders so that each colour and spatial combination occurred with equal likelihood. Participants pressed a left (red) or a right (blue) button to indicate the colour of the stimulus with their index fingers.

Results

Even though extreme responses were left in the analyses, one participant was excluded from the monolingual group who had extreme RTs in the Simon task (i.e. above four standard deviations of the mean RT of all participants). This participant clearly performed the task in a different way from other participants and did this consistently during the experiment. Response accuracies, response speed and estimated ex-Gaussian distribution parameters were analysed using a 2 (Condition) x 3 (Task) x 2 (Participant Group) mixed design *ANOVA*, with Group being a between-group factor. Greenhouse-Geisser corrections were performed when appropriate. Bonferroni correction was applied when following up any interaction. When reporting results, we will focus on main effects of Condition and Group as well as on any interactions involving Group.

Accuracy

Accuracy rate was the percentage of correct responses and was arcsine-transformed for statistical analyses. For illustration purposes, Figure 3.3 shows the average accuracy for each group per condition per task.

The *ANOVA* showed a significant main effect of Condition, F (1, 55.0) = 106.45, p <.001, $\eta_p^2 = .66$. Participants were more accurate in congruent conditions than in incongruent conditions. The main effect of group was not significant, F (1, 55) = 2.16, $p > .05$, $\eta_p^2 = .04$. Similarly, the two-way interaction between Condition and Group was not significance, F (1, 55) = 1.07, $p > .05$, $\eta_p^2 = .02$. The two-way interaction between Task and Group was significant, F (2, 109.5) = 3.4, $p = .037$, $\eta_p^2 = .06$. Follow up tests revealed that the significant interaction was driven by bilingual speakers being more accurate overall in the Spatial Stroop task, F (1, 55) = 4.77, p = .032, η_p^2 = .08; but not the Flanker task, F (1, 55) = 2.6, p > .05, η_p^2 = .05, or the Simon task, F (1, 55) = .12, p > .05, η_p^2 = .002. The three-way interaction of Task, Group and Condition was not significant, F (1.9, 105.3) = 1.09, p > .05, η_p^2 = .02.

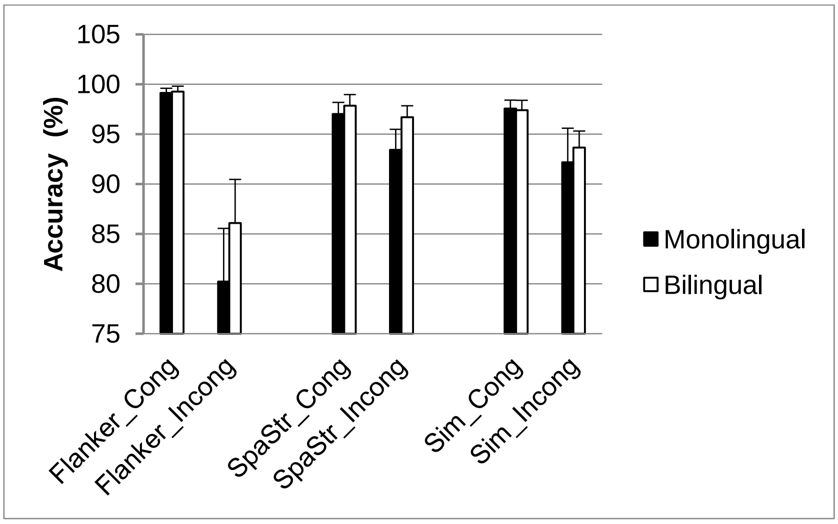


Figure 3.3. Accuracy for Flanker task (left), Spatial Stroop task (middle), and Simon task (right) for monolingual and bilingual speakers for congruent (Cong) and incongruent (Incong) conditions. Error bars represent 95% confidence intervals.

Reaction Times

Reaction times (RTs) were analysed in two ways: by a conventional analysis of condition means to allow comparisons to previous findings and by an analysis of the RT distributions. Given the results of the meta-analysis in Chapter 2, the analysis of condition means was done with and without trimming off very slow responses. For the RT distribution analysis, response times of accurate responses were fitted with ex-Gaussian distributions for each participant in each condition and each task. The ex-Gaussian distribution parameters μ and τ were estimated using the QMPE software, which uses the quantile maximum likelihood estimation method (Brown & Heathcote, 2003). Parameters were estimated for each participant under each condition using five quintiles. All ex-Gaussian parameters were successfully yielded with an average iteration of 14.7. Parameter estimations were all trustworthy according to the technical manual since the exit codes were all below 128. In addition to ex-Gaussian analyses, following Tse et al.'s (2010) suggestion, Quintile analyses (similar to the Vincentile analyses, but use five bins) were conducted to obtain converging evidence for the quality of fit of the RT distributions by the ex-Gaussian models.

Mean response time analyses.

When entering all data (see Figure 3.4) without taking out outliers, there was a significant main effect of Condition on response times, F (1, 55) = 544.4, p <.001, η_p^2 = .91, with incongruent conditions leading to longer response times than congruent conditions. The main effect of Group was not significant, F (1, 55) = 1.16, p > .05, η_p^2 = .02. Importantly, there was a significant Condition by Group interaction, F (1, 55) = 8.46, p = .005, η_p^2 = .13. Follow-up tests showed that the two participant groups did not differ on mean RTs for congruent stimuli, F (1, 55) = 0.97, p > .05, η_p^2 = .002, but there was a trend for a difference for incongruent stimuli, F (1, 55) = 3.07, p = .09, η_p^2 = .05. This pointed to a difference in the congruency effect (incongruent condition – congruent condition) that the two groups suffered. Such difference was confirmed by directly comparing the congruency effect of the two groups. Bilinguals showed significantly reduced congruency effects compared with monolinguals, F (1, 55) = 8.46, p = .005, η_p^2 = .13. Last but not least, the two-way interaction between Task and Group was not significant, F (2, 109.93) = 1.52, p > .05, η_p^2 = .03, neither was the three-way interaction, F (2, 91.58) = 1.37, p > .05, η_p^2 = .02.

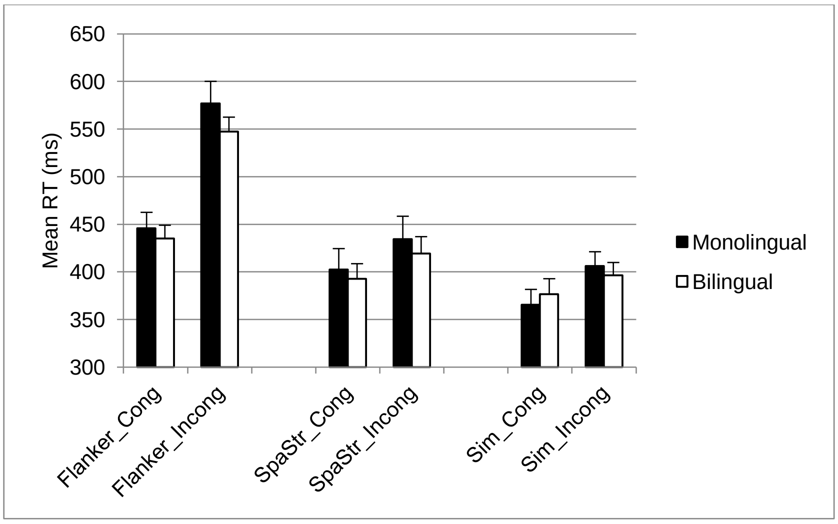


Figure 3.4. Mean RT without excluding slow responses for Flanker task (left), Spatial Stroop task (middle) and Simon task (right) for monolingual and bilingual speakers for (Cong) and incongruent (Incong) conditions. Error bars represent 95% confidence intervals.

To investigate the effect of data trimming, a traditional analysis of response time means was performed after removing responses above 2SD of the participant's mean RT (see Table 3.2 for means and SDs with and without removing outliers). Just as with outliers left in, there was a significant main effect of Condition, F (1, 55) = 626.55, p < .001, η_p^2 = .92, and no main effect of Group, F (1, 55) = .62, p > .05, η_p^2 = .01. Interestingly, the two-way interaction between Condition and Group was marginally significant, F (1, 55) = 4.05, p = .05, η_p^2 = .07, which had been highly significant and with a larger effect size when including responses above 2SDs. Post-hoc tests revealed no significant difference between the two participant groups in the congruent condition, F (1, 55) = 0.11, p > .05, η_p^2 = .002, nor in the incongruent condition, F (1, 55) = 1.41, p > .05, η_p^2 = .03. An analysis of the congruency effects (incongruent condition – congruent condition) showed only a marginally larger congruency effect for bilinguals than monolinguals across all tasks, F (1, 55) = 4.05, p = .05, η_p^2 = .07, which had been highly significant when leaving slow responses in. As before, there was no interaction between Task and Group, F (2, 109.5) = 1.53, p > .05, η_p^2 = .03, or a three-way interaction, F (2, 103.8) = 1.83, p > .05, η_p^2 = .03. In sum, this additional analysis shows that, in line with the conclusion in Chapter 2, trimming the data from slow responses can substantially reduce the bilingual advantage in a conflict task. This also means that the bilingual advantage might be at least partly located in the very slow responses. A detailed inspection of RT distributions as presented below will provide us with more information as to whether an effect is present in the Gaussian and/or the exponential component of the response distributions.

Table 3.2

Means of RT Before and After Removing Outliers

		Flanker			Spatial Stroop			Simon		
Variable	Language Group	Congruent (*SE*)	Incongruent (*SE*)	Effect	Congruent (*SE*)	Incongruent (*SE*)	Effect	Congruent (*SE*)	Incongruent (*SE*)	Effect
Mean RT (without rejecting outliers)	Monolingual	446 (8)	577 (12)	131	402 (11)	434 (12)	32	365 (8)	406 (8)	41
	Bilingual	435 (7)	547 (8)	112	393 (8)	419 (9)	26	377 (8)	396 (7)	19
Mean RT (after rejecting outliers beyond 2 *SD*)	Monolingual	442 (8)	551 (11)	109	396 (11)	421 (11)	25	357 (8)	392 (8)	35
	Bilingual	432 (7)	530 (7)	98	384 (8)	410 (8)	26	368 (8)	389 (7)	21

Ex-Gaussian analyses

Quintile analyses.

In order to investigate the quality of fit of the ex-Gaussian models to the data, Quintile analyses were conducted. In a Quintile analysis, response times are binned, averaged and plotted. In this way, no assumptions about the distribution of the data are made. Both empirical and theoretical Quintiles were calculated. The goodness of fit between the empirical and theoretical Quintiles reflects the extent to which the ex-Gaussian parameters capture the empirical RT distributions (see e.g. Andrews & Heathcote, 2001). Empirical Quintiles were calculated for each participant, each condition and for each task separately. Responses were first sorted and divided into five bins of equal number of responses. Five bins were used because the ex-Gaussian analyses were based on five bins. The average RT in each bin was then averaged across participants. Theoretical Quintiles were estimated according to the respective best-fitting ex-Gaussian distribution. This was done by line search on the numerical integral of the fitted ex-Gaussian distribution (using an R script originally developed by Dr. Andrew Heathcote and acquired from Dr. Chi-Shing Tse through personal communication. This had been used to estimate theoretical Vincentiles in, e.g., Yap, Tse and Balota, 2009). Table 3.3 shows the average empirical and theoretical Quintile bin values for each task, each condition and each group at bin level. Mixed design *ANOVA*s were conducted, with Group as a between-group variable and Estimation Method as a within-group variable, to test the effect of the fitting method. Results are presented in Table 3.3. For most bins, there was a significant main effect of Method, suggesting a difference between an empirical Quintile bin value and a theoretical bin value. Closer inspection revealed that such discrepancies were mostly small in values, with all of them being within 1 *SE* of the empirical values. This suggests that overall ex-Gaussian provided reasonably good fit to the data, despite some systematic differences between the empirical and theoretical values. Most important is the lack of an interaction between Group and Estimation Method, meaning that the Estimation Method affected both groups similarly. Therefore, we can rule out the possibility that any potential group differences in the ex-Gaussian parameters were due to model fitting.

Table 3.3

Empirical and theoretical Quintile bin values for each task, each condition and each bin. F-statistics and p-values for the Main effect of Estimation Method (Method), Group and interaction between Group and Method.

		Flanker Congruent					Flanker Incongruent				
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin1	Bin2	Bin3	Bin4	Bin5
Monolingual	Empirical	359.32	401.49	431.27	468.94	553.17	483.12	532.69	565.12	606.63	703.80
	Theoretical	339.97	399.34	430.99	468.46	556.17	481.00	527.92	561.37	601.26	687.28
	Discrepancy	-19.35	-2.15	-0.28	-0.48	2.99	-2.12	-4.77	-3.75	-5.36	-16.52
	SE	6.50	6.95	7.10	8.25	11.13	8.77	10.03	12.51	15.21	23.22
Bilingual	Empirical	362.79	402.79	430.59	461.82	528.32	462.92	516.31	552.06	584.55	642.38
	Theoretical	358.80	401.80	429.02	460.54	529.56	460.09	511.72	544.55	580.25	648.66
	Discrepancy	-3.99	-0.99	-1.57	-1.28	1.23	-2.83	-4.59	-7.51	-4.30	6.27
	SE	7.43	7.36	7.72	8.20	9.23	7.71	7.24	8.58	10.05	12.84
Method	*F*	5.01	13.32	6.39	3.49	0.69	4.33	22.32	36.67	31.79	0.13
	p	0.029*	0.001**	0.014**	0.07	0.41	0.042*	<.001**	<.001**	<.001**	0.73
Group	*F*	0.93	0.00	0.10	0.73	3.64	2.83	0.63	1.55	2.84	7.16
	p	0.34	0.99	0.76	0.40	0.06	0.10	0.43	0.22	0.10	0.01**
Group*Method	*F*	2.33	0.51	0.22	0.32	0.00	2.76	0.00	0.06	0.95	0.98
	p	0.13	0.48	0.64	0.57	0.96	0.10	0.96	0.81	0.33	0.33

		Spatial Stroop Congruent					Spatial Stroop Incongruent				
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin1	Bin2	Bin3	Bin4	Bin5
Monolingual	Empirical	318.27	360.95	395.19	433.31	518.59	343.68	387.98	423.68	463.20	556.06
	Theoretical	311.45	358.95	391.72	430.44	518.50	334.60	385.73	419.71	459.71	548.06
	Discrepancy	-6.82	-2.00	-3.47	-2.87	-0.09	-9.08	-2.25	-3.97	-3.49	-8.00
	SE	9.68	11.10	12.28	12.88	15.62	9.35	10.56	12.10	13.63	21.06
Bilingual	Empirical	308.36	351.48	390.72	429.81	515.24	342.72	385.03	415.20	446.39	525.61
	Theoretical	300.21	350.36	384.85	425.66	515.67	335.56	383.30	412.15	445.56	518.16
	Discrepancy	-8.16	-1.12	-5.87	-4.15	0.43	-7.16	-1.72	-3.05	-0.83	-7.45
	SE	6.08	7.65	9.05	10.28	14.10	7.26	8.59	9.79	10.83	16.93
Method	*F*	8.53	6.98	29.13	19.94	0.50	4.87	6.35	25.93	8.02	1.13
	p	0.005**	0.011*	<.001**	<.001**	0.49	0.032*	0.015*	<.001**	0.006**	0.29
Group	*F*	0.77	0.54	0.29	0.17	0.08	0.00	0.09	0.43	1.20	2.25
	p	0.38	0.47	0.59	0.68	0.78	0.97	0.77	0.51	0.28	0.14
Group*Method	*F*	0.03	0.19	0.90	0.46	0.02	0.11	0.33	0.20	2.37	0.13
	p	0.86	0.67	0.35	0.50	0.89	0.74	0.57	0.66	0.13	0.72

		Simon Congruent					Simon Incongruent				
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin1	Bin2	Bin3	Bin4	Bin5
Monolingual	Empirical	291.31	327.75	359.11	402.62	501.96	330.67	370.88	401.72	442.36	540.23
	Theoretical	292.04	327.16	357.24	398.21	500.87	322.18	368.00	399.84	439.03	532.03
	Discrepancy	0.73	-0.59	-1.87	-4.42	-1.09	-8.50	-2.89	-1.88	-3.33	-8.20
	SE	9.74	9.90	10.44	12.51	18.04	9.49	9.84	10.26	11.29	13.99

Bilingual	Empirical	293.60	334.55	369.55	413.37	498.42	323.81	366.46	393.61	421.72	493.23
	Theoretical	287.79	333.45	367.83	410.47	509.37	320.17	364.02	389.97	419.62	482.80
	Discrepancy	-5.81	-1.10	-1.72	-2.90	10.95	-3.64	-2.45	-3.64	-2.10	-10.43
	SE	6.32	7.70	8.81	9.81	13.41	7.32	7.64	7.80	8.30	10.65
Method	*F*	1.79	1.82	4.45	12.70	0.61	2.92	24.71	14.25	11.02	3.20
	p	0.19	0.18	0.04*	0.001**	0.44	0.09	<.001**	<.001**	0.002**	0.08
Group	*F*	0.04	0.52	0.52	0.49	0.02	0.20	0.12	0.66	2.65	9.03
	p	0.85	0.47	0.47	0.49	0.90	0.66	0.73	0.42	0.11	0.004**
Group*Method	*F*	3.63	0.24	0.05	0.19	0.55	0.30	0.52	0.23	1.41	0.24
	p	0.06	0.63	0.82	0.67	0.46	0.59	0.47	0.64	0.24	0.63

Note. * signifies that the effect was significant at .05 significance level. ** signifies that the effect was significant at .01 significance level.

The μ parameter.

Figure 3.5 shows the average μ for each group per condition per task. There was a significant main effect of Condition, F (1, 55) = 320.49, $p < .001$, $\eta_p^2 = .85$, with larger μ for incongruent conditions than congruent conditions. There was no main effect of Group, F (1, 55) = .56, $p > .05$, $\eta_p^2 = .01$, indicating that monolingual and bilingual speakers did not differ with respect to μ. The two-way interaction between Task and Group showed only a trend, F (1.9, 106.7) = 2.62, $p = .08$, $\eta_p^2 = .05$. Follow-up analyses showed that the two groups had very similar performance in the Flanker task, F (1, 55) = 0.006, $p > .05$, $\eta_p^2 = .009$, and the Spatial Strop task, F (1, 55) = 0.09, $p > .05$, $\eta_p^2 = .002$, while monolinguals were overall faster in the Simon task, F (1, 55) = 5.92, $p = .02$, $\eta_p^2 = .097$. Most importantly, there was neither a Condition by Group interaction, F (1, 55) = .02, $p > .05$, $\eta_p^2 = .00$ nor a three-way interaction, F (2,107.2) = .59, $p > .05$, $\eta_p^2 = .01$, meaning that the μ congruency effects were the same for the two participant groups, and this was the case for all tasks.

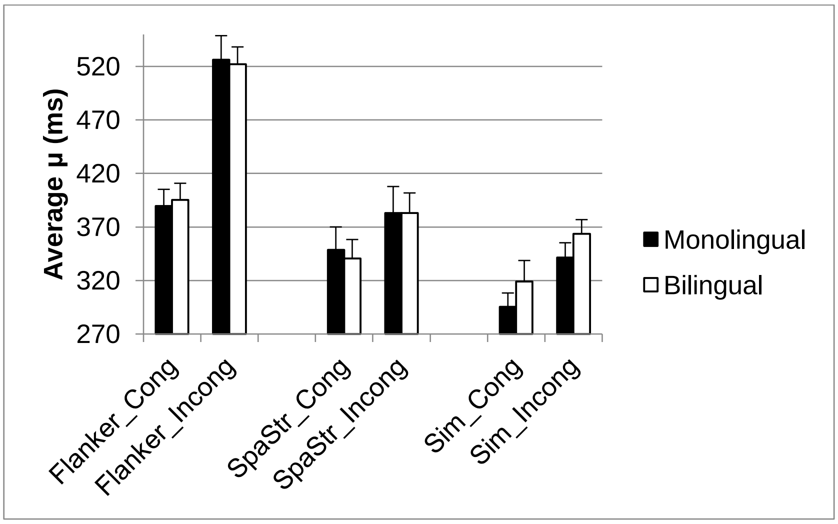


Figure 3.5. Means of the Ex-Gaussian parameter μ estimated from individual RT distributions for monolingual and bilingual speakers for congruent (Cong) and incongruent (Incong) conditions. Flanker task (left), Spatial Stroop task (middle) and Simon task (right). Error bars represent 95% confidence interval.

The τ parameter.

Figure 3.6 shows the average τ for each group per condition per task. There was a significant main effect of Condition, F (1, 55) = 10.33, p = .002, η_p^2 = .16, with the incongruent condition having smaller τ than the congruent condition. There was a significant main effect of Group, F (1, 55) = 17.91, p <.001, η_p^2 = .25, with bilingual speakers having a smaller τ than monolingual speakers. The two-way interaction between Condition and Group showed a trend, F (1, 55) = 3.51, p = .07, η_p^2 = .06. Follow-up tests revealed that bilingual speakers had significant smaller τ in the incongruent condition, F (1, 55) = 19.89, p < .001, η_p^2 = .26, while this difference was only a trend in the congruent condition, F (1, 55) = 3.35, p = .07, η_p^2 = .06. Therefore, the τ was more consistently smaller for bilinguals in the incongruent condition. There was no significant interaction between Task and Group, F (2, 109.5) = 1.28, p > .05, η_p^2 = .02, nor a three-way interaction, F (2, 109.4) = .38, p > .05, η_p^2 = .01, indicating that the τ pattern was consistent across the three tasks for the two participant groups, with bilinguals having a smaller τ in both conditions.

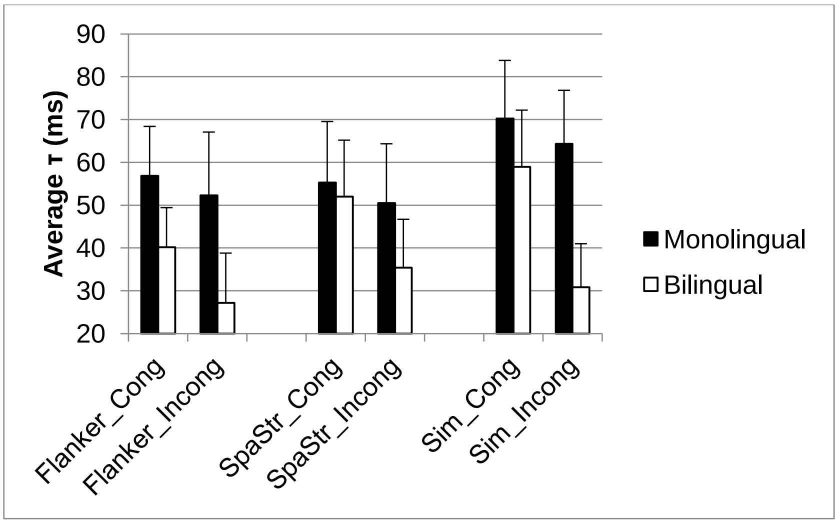


Figure 3.6. Means of the Ex-Gaussian parameter τ estimated from individual RT distributions for monolingual and bilingual speakers for congruent (Cong) and incongruent (Incong) conditions. Flanker task (left), Spatial Stroop task (middle) and Simon task (right). Error bars represent 95% confidence interval.

Discussion

The aim of the current study was to investigate the contribution of inhibitory control and/or attentional control to the bilingual advantage by investigating error patterns and utilizing ex-Gaussian analyses of response time distributions in three non-verbal interference tasks. We argue that the result patterns suggest enhanced bilingual attentional control, but not enhanced inhibitory control.

Response accuracies only showed a weak advantage for bilingual speakers in our tasks: bilinguals were overall more accurate than monolinguals in the Spatial Stroop task. Closer inspection showed that, descriptively, bilinguals had higher accuracy rates in the Flanker task, and the incongruent condition in the Simon task. However, it is not possible to tell whether this was caused by enhanced inhibitory and/or attentional control ability.

The analysis of the mean response times rather suggests an advanced bilingual attentional control ability than an enhanced inhibitory control ability. When very slow responses were included in the analysis, bilingual speakers showed a smaller congruency effect. However, this effect was reduced to a trend when very slow responses were removed from the analysis. Therefore, the difference between monolinguals and bilinguals appears to be rather driven by very slow responses, in line with a difference in attentional control.

Results of the distribution analysis strongly confirm this tentative conclusion. They do not support an enhanced inhibitory control ability in bilinguals. Across all tasks, μ was consistently larger in the incongruent condition than in the congruent condition. This means that μ (i.e. the leading edge of the RT distribution) was sensitive to interference, with average processing time being longer when interference was present. Importantly, the two participant groups did not perform differently in either the congruent condition or the incongruent condition. This suggests that bilinguals do not appear to resolve conflict better than monolinguals.

In contrast, results for the ex-Gaussian parameter τ are consistent with enhanced bilingual attentional control. The bilingual group had smaller τ values and therefore shorter RT distribution tails regardless of condition (congruent or incongruent) and regardless of task. The τ parameter reflects both the frequency and the degree of excessively long RTs. Therefore, the results mean that bilingual speakers had fewer excessively long RTs, and their extreme responses were not as extreme as those of monolingual speakers. Importantly, this was the case not only in the incongruent condition but also in the congruent condition (to a somewhat lesser degree). Therefore, the results of the τ parameter reflect enhanced performance of bilinguals in general, not restricted to situations that request dealing with conflicting information. Furthermore, this enhanced performance was consistent across all three interference tasks, confirming that we are dealing with an ability that is not restricted to a particular task, but rather domain-general. Given that our three interference tasks were relatively similar, the exact extent of the generalizability still needs to be established.

One somewhat unexpected result of our study was that incongruent conditions led to shorter RT distribution tails than congruent conditions. If a shorter tail reflects increased attentional control, then the results suggest that incongruent trials elevated the level of attentional control and that this was the case in both monolingual and bilingual speakers. The processing system therefore appears to be able to detect incongruent information and increase attentional engagement in the task accordingly. This increase of attentional engagement can be explained with the classic conflict-monitoring system proposed by Botvinick et al. (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999). A conflict-monitoring system detects conflict and modulates online shift of attentional control. This has been called upon to explain sequence effects in interference tasks (Gratton, Coles, & Donchin, 1992; Wuhr & Ansorge, 2005), i.e. that the congruency effect is smaller following an incongruent trial compared with a congruent trial. This suggests that attentional control is not a static mechanism or ability, but rather context sensitive. Once a response conflict is detected, attentional control is elevated. As a result one becomes particular vigilant and engage attention to a greater degree, and consequently reducing the occurrences of lapses of attention. Last but not least, the tendency for an interaction between group and condition suggests that bilinguals might be able to adapt the level of attentional control more swiftly and flexibly than monolinguals.

In contrast to our interpretation of bilinguals' shorter RT distribution tails, one might argue that bilinguals were more eager to respond quickly compared to monolingual speakers. But this cannot be the case because this should have led to a speed-accuracy trade-off. However, bilinguals had shorter response distribution tails while being similarly, if not more accurate than monolinguals.

Furthermore, more errors in the monolingual group could potentially have led to a more positively skewed response distribution due to post-error slowing (Danielmeier & Ullsperger, 2011; Laming, 1979). Responses following an error are typically longer than responses following accurate trials, especially when errors are infrequent. To exclude the effects of post-error slowing, an additional ex-Gaussian analysis was conducted, after removing all the error and post-error trials for both participant groups. The result pattern was almost identical to that reported in the results section. The only change was that the condition by group interaction for the τ parameter changed from marginal to significant. Thus, while both groups had a reduction of τ in the incongruent condition compared with the congruent condition, such reduction was larger for the bilingual than for monolingual speakers. In sum, this means that longer tails in the monolingual group were not due to post-error slowing.

Comparing these findings with those of Calabria et al. (2011) and Abutalebi et al. (2015), converging evidence for an attentional account of the bilingual advantage emerges, even though the results are not exactly the same. Results for the exponential component of the response distributions are consistent with Calabria et al. (2011), with both studies finding overall reduced tails for bilingual speakers, this was also true for Abutalebi et al. (2015) descriptively, although a statistical result for the main effect of Condition was not available. At the same time, results in Calabria et al. (2011) and Abutalebi et al. (2015) did not imply a strong role of inhibitory control either, due to the lack of an interaction between Group and Condition in the Gaussian component of the response distributions.

However, Calabria et al. (2011) and Abutalebi et al. (2015) reported additional effects, i.e. an overall advantage of bilingualism in the Gaussian component of the response distributions, which was not observed in the present study. There are two potential causes for this discrepancy, which are not mutually exclusive. First, the discrepancy could be caused by differences regarding the paradigm. While in the present study a ‘pure’ Flanker task was conducted (among other interference tasks), Calabria et al. (2011) conducted an ANT study where stimuli were preceded by cues. Further research would need to investigate whether this methodological difference can lead to different effects in the responses. Second, the discrepancy could be caused by the different populations tested. While Abutalebi et al. (2015) tested elderly the current study tested young adults. The former age group has seen rather consistent evidence for bilingual advantage (e.g. Bialystok et al., 2004; Bialystok, Craik, & Luk, 2008; Salvatierra, & Rosselli, 2011).

Because the evidence did not support enhanced inhibitory control, the question arises whether an overall speed advantage could be accounted for by an enhanced attentional control ability in bilinguals? This might be the case if an enhanced attentional control in the form of an enhanced alertness of the mental state led to enhanced processing speed. Such explanation is consistent with the proposal by Posner & Petersen (1990) that alertness can affect the rate at which a response is selected. This can then explain the faster responses overall in the Gaussian component observed in the two previous studies.

Our finding that bilingual speakers showed enhanced attentional control ability (smaller τ in all conditions) is also conceptually consistent with Tse and Altarriba (2014)'s finding, who investigated the modulation effect of L1/L2 proficiency on speakers' response distributions in a Simon task. They found that L1 proficiency modulated τ in both congruent and incongruent conditions, supporting enhanced attentional control for children with higher L1 proficiency. Their results also suggested enhanced inhibitory control for children with higher L2 proficiency, who showed smaller Simon effect in μ.

Similarly, using a linguistic Colour Stroop task, Tse and Altarriba (2012) found that the more proficient a speaker was in L1/L2, the smaller τ was, regardless of condition. In other words, more proficient individuals were better at maintaining task goals. In contrast to our finding, they also observed that language proficiency modulated the Stroop effect in μ, with more proficient individuals having a smaller Stroop effect in μ, suggesting that bilinguals with higher proficiency have developed enhanced inhibitory control compared with lower proficiency speaker. However, unlike tasks used in the present study, the Colour Stroop task involves verbal responses. It might be that bilingual's constant exercise of inhibitory control in language production enhances their conflict resolution ability in verbal tasks. Our results suggest that such ability does not necessarily transfer to non-verbal tasks.

What remains to be explained is why the bilingual advantage has sometimes materialized in previous studies in the form of an overall speed advantage, sometimes a reduced congruency effect, and sometimes both. By examining our data we have seen that effects in the tail of response time distributions can translate into effects on mean response times in a traditional analysis. This suggests that both overall speed advantages and reduced congruency effects can stem from response distribution tails. However, the relation between effects in tails and in mean response times is not a simple one-to-one translation. For instance, in the present study, overall effects in the tails were found, but a reduced congruency effect was found instead in the traditional analysis. The latter effect was descriptively present in the tails as well, but it did not reach significance. Therefore, what will show up in traditional analyses will depend on the combination of data trimming procedures and the effects present in the tails.

Last but not least, the present study confirmed the conclusion of Chapter 2 that trimming very slow responses can substantially reduce a bilingual effect in a traditional analysis of condition means. Therefore, the suggestion that the inconsistent findings for the bilingual advantage are (at least partly) due to different data trimming procedures appears to be a parsimonious one.

Conclusion

In conclusion, the present study found that monolingual and bilingual speakers did not differ in terms of the leading edge of RT distributions (μ) in three non-verbal conflict tasks. However, bilinguals had shorter distribution tails (τ) in both conflict and non-conflict conditions. Results support the conclusion that enhanced attentional control abilities, but not inhibitory control abilities, underlie the bilingual advantage in conflict tasks. It became apparent that ex-Gaussian analysis of RT distributions are very useful in providing in-depth knowledge beyond analyses of central tendencies. Thus, this type of analyses is recommended for future studies.

CHAPTER 4

BILINGUALISM ENHANCES LEXICAL SELECTION IN COMPETITIVE SEMANTIC CONTEXTS

Abstract

Bilingual speakers have been suggested to have developed superior executive control through life-long practice of speaking two languages. Such an advantage has been shown almost exclusively in the non-verbal domain by using interference tasks, in which bilingual speakers have been found to be faster overall and/or less affected by competing stimulus. The present study tested whether bilingual speakers are also less affected by a competitive semantic context when naming pictures in one of their languages. Monolingual, bilingual and proficient L2 speakers were tested using a cyclic semantic blocking paradigm. Traditional mean response times per condition as well as response time distributions were investigated. All participants suffered from semantic interference, but differed in their response time distribution profiles. While monolingual speakers showed an effect of the semantic context only in the Gaussian part of their response distributions, bilingual and proficient L2 speakers showed an effect only in the exponential part. These results are congruent with an enhanced top-down control mechanism that inhibits competitors, suggesting that bilinguals' enhanced executive control affects their lexical selection processes.

Introduction

Speaking more than one language is a demanding task because speakers have to constantly deal with the problem of parallel activation of languages regardless of the language in use (BijeljacBabic, Biardeau, & Grainger, 1997; Colome, 2001; Spivey & Marian, 1999; Wu, Cristino, Leek & Thierry, 2013; Hoshino & Thierry, 2012; Wu & Thierry, 2012). This poses extra monitoring demands on speakers. Nevertheless, bilingual speakers are successful in producing fluent speech without many intrusion errors or frequent involuntary switches between languages. This has led to proposals that bilingual speakers develop somewhat different attention and control procedures for speech production (e.g. Green, 1998) and that the daily exercise of such procedures might affect not only speech production, but also the general cognitive system (Abutalebi et al., 2012; Bialystok, Craik, & Luk, 2012). In the previous chapters we have introduced, discussed and cumulated empirical evidence that bilingualism leads to cognitive advantages (for a review see Bialystok et al., 2009). One such advantage is their enhanced executive control ability. Different aspects of executive control have been called upon to explain such an advantage, such as attentional control and inhibitory control (see Chapter 3, page 35).

Executive Control in Verbal Tasks

While such enhanced executive control in bilinguals has been tested extensively in the non-verbal domain (for a review, Hilchey & Klein, 2011, see also Chapter 2), it has seldom been discussed how enhanced executive control might contribute to the control within one of their two (or more) languages (but see Luo, Luk, & Bialystok, 2010). This chapter addresses this issue and investigates whether speakers who speak a second language fluently show enhanced lexical selection abilities in competitive semantic contexts when naming pictures in one of their languages. Literature in this area is sparse. We next review a few studies that compared the performance of monolingual and bilingual speakers in verbal tasks that require high executive demands.

Maybe the most classic task that involves response conflict within a single language is the Colour Stroop task (Stroop, 1935). In this task participants are instructed to name the ink colour of colour words (i.e. the word RED printed in black ink, and participants are required to say ‘black’). Responses are typically slower when the word meaning does not correspond to the ink (the word RED printed in black) than when the word meaning does correspond to the ink (the word RED printed in red). The difference in response times in the two conditions is called the Stroop effect or Stroop interference. Because the Stroop task requires that participants inhibit or override the tendency to produce a more dominant or automatic response, namely reading the word, the Stroop effect has often been related to the ability to resolve the conflict between parallel routes of activation or relevant/irrelevant dimensions of the stimuli (Macleod, 1991; Spieler, Balota, & Faust, 1996; Tse & Altarriba, 2012), arguably by inhibiting the representation of the irrelevant dimension of the stimuli. Evidence is mixed as to whether bilinguals perform differently in a Stroop task compared to monolinguals. Bialystok, Craik, and Luk (2008) reported a smaller Stroop effect for bilinguals, suggesting a greater degree of executive control, or more precisely inhibitory control. Tse and Altarriba (2012) tested bilinguals with various L1/L2 proficiencies on a Stroop task and found a negative relationship between language proficiency and the magnitude of the Stroop effect. In contrast, other studies found no differences between monolinguals and bilinguals (Kousaie & Phillips, 2012; Lee & Chan, 2000; Rosselli, Ardila, Santisi, et al., 2002). Regardless of the mixed evidence, what the Stroop task actually measures remains debatable. On the one hand, there are two colour words that appear to compete with each other in the conflict condition (black and red in the example above). The task might therefore tap into within-language lexical competition. On the other hand, the most prominent account of the Stroop effect suggests that the conflict lies between two tasks, namely word reading and colour naming (Macleod, 1991), rather than between two lexical representations. Therefore, results obtained from the Stroop paradigm might not tell us anything about how speakers resolve intra-language lexical competition, but rather about executive control when two tasks compete with one another.

A more appropriate verbal task for investigating within-language lexical selection and control is the verbal fluency task. In this task, participants are required to generate exemplars according to a given rule within a short time period. Two versions of the verbal fluency task are frequently used, the category fluency task, where participants are required to name words within a given semantic category (e.g. animals), and the letter fluency task, where participants are required to name words starting with a given letter (e.g. words starting with the letter 'l'). Executive control has been shown to contribute to the performance in the verbal fluency task (Shao, Janse, Visser, & Meyer, 2014) and to the occurrence of intrusion errors from the second language in bilingual speakers (Gollan, Sandoval, & Salmon, 2011). But despite bilingual speakers' assumed enhanced executive control, they typically perform worse in the verbal fluency task (Rosselli, Ardila, Salvatierra, et al., 2002; Sandoval, Gollan, Ferreira, & Salmon, 2010). This has been attributed to various reasons. First, bilingual speakers need to resolve competition of words from their 'other' language (Sandoval, Gollan, Ferreira, & Salmon, 2010). Second, bilingual speakers' exposure to a single language is more limited than that of monolingual speakers. This means that, on average, bilingual speakers encounter or retrieve each lexical item less often than monolingual speakers, leading to less efficient lexical retrieval (Gollan, Montoya, Cera, & Sandoval, 2008; Gollan, Montoya, & Werner, 2002; Ivanova & Costa, 2008). Third, the vocabulary sizes of bilingual speakers are typically smaller than those of monolingual speakers in one language (Oller, Pearson, & Cobo-Lewis, 2007; Portocarrero, Burright, & Donovick, 2007) meaning that it is more difficult to retrieve as many exemplars as monolinguals. Indeed, Luo et al. (2010) showed that when bilingual and monolingual speakers were matched for proficiency and vocabulary size, the bilingual disadvantage in the category fluency task vanished. And, interestingly, when matched for vocabulary size, bilinguals outperformed monolinguals in a letter fluency task. This finding is surprising because bilingual lexical retrieval has generally been viewed as being worse. For example, bilinguals have been found to name pictures more slowly than monolinguals in the Boston Naming Test, even in their dominant language (Gollan, Montoya, Fennema-Notestine, & Morris, 2005). In contrast, Luo et al.'s (2010) results suggest that bilingual speakers might be better at lexical selection within a single language, at least when experiencing increased lexical competition.

The Semantic Blocking Paradigm

We therefore investigated the potential bilingual benefit for lexical selection during increased lexical competition. We chose the cyclic version of the semantic blocking paradigm (Kroll & Stewart, 1994), in which participants name pictures that are blocked into homogeneous lists (i.e. pictures from the same semantic category) or heterogeneous lists (i.e. pictures from different categories). Participants typically name pictures from homogeneous lists more slowly than those from heterogeneous lists (e.g. Belke, Meyer, & Damian, 2005; Damian, Vigliocco, & Levelt, 2001; Maess, Friederici, Damian, Meyer, & Levelt, 2002). This finding has been referred to as the "semantic interference effect" or more specifically the "semantic blocking effect".

The semantic blocking effect has been proposed to be caused by lexical-semantic competition created by the cyclic presentation of stimuli in this paradigm (Abdel Rahman & Melinger, 2009; Belke, 2013; Damian et al., 2001). A few exemplars are presented cyclically several times in a block (e.g. duck, mouse, snake and fish). On the basis of the general assumption of connectivity and spreading activation in the mental lexicon, the target concept increases the activation levels of the lexical representation of the target as well as of semantically related representations, especially those within the same semantic category (Abdel Rahman & Melinger, 2009; Bloem & La Heij, 2003). This leads to higher activation levels of non-target alternatives in the homogeneous condition than in the heterogeneous condition. Assuming that lexical selection is a competitive process, in other words if the selection of an item depends on the activation level of alternatives, the homogeneous condition would thus lead to longer naming latencies than the heterogeneous condition. Although there is still debate as to the exact mechanism behind the effect, the locus of the effect is generally agreed to happen at lexical level (Belke, 2013; Damian et al., 2001).

In addition to competition among lexical items, a top-down control process has been proposed to be involved in the semantic blocking paradigm. Belke (2008) manipulated working memory load during a semantic blocking task and reported an interaction between semantic context and working memory load, with the semantic context effect being bigger when participants concurrently performed a working memory task. She proposed a biased selection mechanism that limits the scope of selection to a task-defined set (i.e. the items in a block) and uses category information to further facilitate selection in the heterogeneous condition. Once additional working memory load interferes with such control, the semantic context effect becomes exacerbated. As pointed out by Belke & Stielow (2013), such top-down control still needs to be implemented and specified by the current models of word production. Although the exact mechanism of such a top-down control mechanism is unspecified, its principal of operation makes it akin to executive control in that it maintains a task goal (e.g. a task set) and uses this goal to bias response selections, possibly through increasing the activation of target representations and/or reducing the activation of non-target representations (similar to inhibition).

Given that bilingual speakers typically have smaller vocabulary sizes than monolingual peers when compared in one language, one also needs to consider the potential role that vocabulary size might play in the cyclic naming task. The lexical cohort hypothesis (or the swinging lexical network proposal) suggests that naming a few exemplars from the same category activates a network of related items (Abdel Rahman & Melinger, 2009; Belke & Stielow, 2013). This means that the competition in the cyclic naming paradigm might not be confined to the small number of words being named, but might involve a larger set of semantically related words. Indeed, Belke et al. (2005) showed that the semantic blocking effect transfers to new members of a category. This suggests the possibility that vocabulary size modulates the semantic blocking effect. For example, a bigger vocabulary size might mean more activated competitors and thus a larger blocking effect. Since bilingual speakers typically have smaller vocabulary sizes than monolingual speakers, we needed to rule out in this study that any performance differences in the semantic blocking task were due to differences in vocabulary sizes. We therefore tested three groups of participants: a monolingual group, a bilingual group and a group of native English speakers with high second language proficiency (thereafter proficient L2 speakers). This group was well matched to monolingual speakers in terms of their target language profile (vocabulary size and self-rated proficiency) and resembled the bilingual group in terms of their second language experience (although to a lesser degree) (see Method).

Ex-Gaussian analysis

Results from the semantic blocking paradigm have to date been analysed by comparing error rates and mean reaction times between the homogeneous and the heterogeneous condition. However, reaction times are not normally distributed. They are positively skewed due to occasional extremely long responses. It is generally a common practice to exclude such long responses from any analysis because they are likely caused by 'anomalous' cognitive processes such as corrected errors or attentional lapses when compared to 'average' responses. However, long responses can be highly informative because they reflect effects of executive control. This is particularly pertinent for the present study since we are interested in executive control. We therefore analysed response time distributions and compared the results with those of a traditional analysis of condition means. Chapter 3 (page 35) presented an introduction to the ex-Gaussian distribution and its parameters. We next introduce what the parameters might mean within the context of the semantic blocking paradigm.

The μ parameter.

The μ parameter captures the leading edge of an ex-Gaussian distribution. It resembles the mean of responses after taking out extremely slow responses because it represents the majority of the responses. It is directly affected by the experimental manipulation such that a more demanding condition yields bigger μ than a less demanding one (Aarts, Roelofs, & van Turennout, 2009; Heathcote et al., 1991; Leth-Steensen, Elbaz, & Douglas, 2000; Spieler, Balota, & Faust, 2000; Tse, Balota, Yap, Duchek, & McCabe, 2010). For example, for the response distribution of the traditional Colour Stroop task, μ has been found to be significantly larger in the incongruent condition than in the congruent condition (de Zubicaray, McMahon, Eastburn, & Pringle, 2006; Heathcote et al., 1991; Spieler et al., 2000; Steinhauser & Hubner, 2009). Therefore an increase in μ reflects extra processing costs for majority responses.

A similar logic can be applied to the semantic blocking paradigm. A strong semantic competition in the homogeneous condition leads to longer processing times than in the heterogeneous condition. Therefore, prolonged reaction times in the homogeneous condition compared with the heterogeneous condition should result in a positive shift of the response distribution without changing its shape, meaning that response times should be uniformly longer.

The τ parameter.

The τ parameter, which reflects the tail of the reaction time distribution, captures the amount and the degree of long responses. We can see two possible causes for long responses in a semantic blocking experiment. First, long responses could be modulated by the attentional control ability that is necessary to maintain the goal of a task. Impaired attentional control ability could result in more frequent lapses of attention and/or temporary loss of task goals from working memory, thus leading to long responses. This idea is in line with the findings of a negative correlation between τ in forced choice reaction tasks (such as a word classification task) and working memory measures (such as scores of a reading span task) (Schmiedek et al., 2007; Tse et al., 2010). Also, populations with impaired attentional control ability have been shown to have longer RT distribution tails in various tasks (Hervey et al., 2006; Leth-Steensen, Elbaz, & Douglas, 2000), including the Attentional Network Test (ANT), a Go/No-Go task, a Stop-signal task and an N-back task (Epstein et al., 2011).

A second cause of the long responses, particularly relevant for the present paradigm, is covertly corrected errors, i.e. wrongly initiated responses that are corrected before articulation. This self-correction process can be separated into two stages. In the first stage, a selection is initiated which is not the target. In the second stage, the monitoring system detects an error and corrects the error by deselecting the incorrect candidate and instead selecting the correct target. Critically, it has been shown that competitors or non-selected items are actively inhibited (Tipper, 1985; Tipper & Driver, 1988). For example, in a Colour Stroop paradigm, naming speed was slowed down if the target was the distractor in the previous trial, suggesting that the distractor representation was actively inhibited and made less available when it needed to be subsequently named (Neill, 1977). If we assume that the same mechanism inhibits non-selected competitors in the first processing stage of a self-corrected error (i.e. during the initial selection of the non-target word), then the speed of recovering from the error in the second stage should depend on the degree to which competitors were inhibited: the stronger competitors were inhibited, the longer it should take to recover.

Predictions

Bilingual speakers' performance in previous studies predicts two possible outcomes of our experiment in terms of ex-Gaussian analysis. First, compared to monolinguals, bilinguals might show a reduced semantic blocking effect in μ due to a superior ability to handle competing information. This outcome is predicted by the finding that bilingual speakers were less affected than monolingual speakers by interfering information in non-verbal tasks (Bialystok et al., 2004; Linck, Hoshino, & Kroll, 2008; Tao, Marzecova, Taft, Asanowicz, & Wodniecka, 2011). They also have plenty of exercise to ensure that competing lexical items from a second language do not interfere when speaking in one language, which might enhance their top-down lexical inhibitory control not only across but also within one language. This might possibly come with a side effect in cases of covertly corrected errors. Since competitors of the initially selected words were inhibited stronger, it might take them even longer to recover from those covert errors. This would result in a bigger τ, i.e. longer response tails, in the competitive context. Such a result pattern might lead to no participant group differences in the traditional analysis of condition means because a smaller μ value and a larger τ value might cancel each other out.

Second, bilingual speakers' enhanced cognitive control ability might result in participants being more focused on the task, with fewer lapses of attention and therefore smaller distribution tails. Importantly, smaller distribution tails should be present in all experimental conditions (homogeneous and heterogeneous condition), because the enhanced control ability should benefit any responses. Note that a better focus on the task should not have consequences for the magnitude of the semantic interference effect, which should not differ between bilingual and monolingual speakers.

In addition, by comparing performances of proficient L2 speakers with that of monolingual and bilingual speakers, we can gain insights into the potential roles of vocabulary size and second language experience in the semantic blocking task. If vocabulary size was a critical modulator of the semantic blocking effect, then we would expect the proficient L2 group to behave similarly to the monolingual group. On the other hand, if bilingual experience was the critical modulator, we would expect the proficient L2 group to behave more similarly to the bilingual speakers.

Method

Participants

129 participants took part in the experiment: 81 native English speakers and 48 English/Chinese bilingual speakers. They were mostly undergraduate and postgraduate students of the University of Birmingham who participated either for course credits or cash. Eight of the bilingual speakers were students of the Chinese University of Hong Kong and were paid for their participation. For the analysis, participants were selected on the basis of their responses to a questionnaire about their language use history (see Appendix A-1), adapted from Silverberg and Samuel (2004). In this questionnaire, participants were asked to rate their self-perceived proficiency in English and to list all the languages that they were able to speak as well as the age at which they started to learn them. In addition, bilingual speakers were asked to rate their self-perceived proficiency in Chinese. To be classified as

bilingual, the following criteria had to be met: the participant (a) learnt both languages before age 10 and (b) was fluent in both languages with more than 50% native-like proficiency in both languages. Monolingual English speakers were defined as follows: the participant (a) did not speak another language fluently and (b) did not learn another language before age 10. The group of proficient L2 speakers consisted of native English speakers whose vocabulary knowledge and English proficiency were comparable to that of the monolingual English group (see Table 4.1) and with more than 50% native-like proficiency in a second language by self-report. The other language that the proficient L2 speakers spoke varied (French (9), German (5), Punjabi (3), Spanish (2), Yoruba (1), Urdu (1), Gujarati (1), Greek (1), Danish (1) and Cantonese (1)). In addition, we did not restrict the age at which another language was learnt, which means that the onset age varied and was on average later than that of the bilingual group (see Table 4.1). To ensure the homogeneity of groups, we targeted participants who were undergraduate students or who had just graduated from their undergraduate degree course.

To ensure that the proficient L2 group matched the monolingual group on vocabulary size, we measured participants' general vocabulary knowledge using the British Picture Vocabulary Scale, 2nd edition (Dunn, Dunn, Whetton, & Burley, 1997), a standardised vocabulary test for receptive vocabulary in the form of a picture pointing task. In addition, we measured participants' vocabulary knowledge in the four semantic categories used in the experiment with a Category Vocabulary List. For that, we assembled a vocabulary list of words from the four categories used in the experiment (see Material section below) with around 62 words in each category (see Appendix A-2). Participants were instructed to tick words that they knew.

These criteria led to twenty-five English/Chinese bilingual speakers and twenty-five age-matched native English speakers who met the proficient L2 standard described above being included into the analyses. See Table 4.1 for a summary of participants' demographic characteristics. The self-reported proficiency ratings revealed that bilinguals were equally proficient in English and Chinese, t (24) = .11, $p > .05$, and that they started to learn the two languages at roughly the same time, t (24) = 1.21, $p > .05$. Out of the participants tested, twenty-five native English speakers who met the monolingual standard were randomly selected to match the bilinguals and proficient L2 for age, F (2, 72) = .054, $p > .05$, η_p^2 = .001, and education, X^2 (4, N = 75) = 6.1, $p > .05$. As expected, vocabulary sizes (as measured by the BPVS) could not be matched across groups, F (2, 71) = 11.9, $p < .001$, η_p^2 = .25. This was due to bilingual speakers' vocabulary sizes being overall smaller than that of monolinguals, t (31.2) = 4.40, $p < .001$, and proficient L2 speakers, t (39.5) = 3.1, $p = .003$, with the latter two groups not differing from each other, t (41.8) = 1.4, $p > .05$. Vocabulary scores obtained by the Category Vocabulary List were highly correlated with the raw scores of the BPVS, r (74) = .79, $p < .001$, and showed the same pattern. Overall Category Vocabulary sizes differed across groups, F (2, 72) = 20.1, $p < .001$, η_p^2 = .40. Bilingual speakers' Category Vocabulary sizes were smaller than that of monolinguals, t (32.9) = 5.49, $p < .001$, and proficient L2 speakers, t (40.8) = 4.46, $p < .001$, with the latter two groups not differing from each other, t (42.6) = .91, $p > .05$. The same was true for self-ratings of English proficiency, F (2, 72) = 24.9, $p < .001$, η_p^2 = .41. Bilinguals had a smaller English proficiency than monolinguals, t (25.9) = 5.55, $p < .001$, and proficient L2 speakers, t (32.2) = 4.91, $p < .001$, with the latter two not differing from each other, t (33.9) = .69, $p > .05$.

Table 4.1

Demographic information and language background for monolingual, proficiency L2 and bilingual speakers.

	Monolinguals	Proficient L2	Bilinguals
N (Female)	25 (17)	25 (20)	25 (20)

Age (SD)	20.3 (2.7)	20.4 (4.5)	20.6 (2.8)
Median Education (Min: Max)	2 (2:4)	2 (2:4)	2 (2:3)
Mean Vocabulary Score (SD)	.68 (.08)	.66 (.12)	.46 (.19)
Mean RSVP Raw Score (SD)	143.4 (10.0)	138.3 (15.0)	121.1 (22.9)
Mean Proficiency Rating for English (SD)	.97 (.03)	.96 (.06)	.82 (.13)
Mean Proficiency Rating for Other Language (SD)	-	.67 (.21)	.82 (.21)
Mean Age of English Onset (SD)	0	0	2.68 (1.68)
Mean Age of Onset for Other Language (SD)	-	6.6 (4.9)	1.96 (2.21)
Fluent L2 speakers	No	Yes	Yes

General Design and Procedure

Informed consent was obtained from each participant. The order of the tasks was fixed, with the semantic blocking task first, followed by the language history questionnaire, the Category Vocabulary List and the BPVS.

Material.

We utilized the version of the semantic blocking paradigm developed by Damian, et al. (2001). Stimuli were taken from previous studies (Belke et al., 2005; Damian et al., 2001). They were controlled for visual similarity and phonological overlap within each presentation block. This led to sixteen black-and-white pictures of daily objects from four semantic categories, i.e. four pictures from each of the four categories: animals, clothing, furniture and tools. Table 4.2 shows the picture names arranged into a 4 x 4 matrix where each column constitutes a homogenous stimulus set (pictures from the same category), and each row constitutes a heterogeneous stimulus set (pictures from different categories). This resulted in four homogeneous stimulus sets and four heterogeneous stimulus sets. Eight such blocks were created out of eight stimulus sets. A test block was formed by repeating one stimulus set eight times. This resulted in 128 trials per condition.

Table 4.2

Pictures (see Appendix A-3) used in the semantic blocking task. Pictures of words in rows constitute heterogeneous stimulus sets, while words in columns constitute homogeneous stimulus sets.

Clothing	Animals	Tools	Furniture
Tie	Snake	Brush	Lamp
Coat	Duck	Saw	Chair
Boot	Mouse	Rake	Desk
Skirt	Fish	Drill	Bed

Procedure.

Each trial started with a fixation cross in the centre of the screen for 800 ms, followed by a target picture presented for 250 ms. Participants were instructed to name the picture aloud as quickly as possible while being accurate. The maximum response time allowed was 2200 ms from the onset of the picture. Responses were audio-recorded and response times were measured with a Cedrus SV-1 Voice Key. All participants saw all eight blocks. Block sequences were counterbalanced. Half of the participants started with a homogenous block, while the other half started with a heterogeneous block.

Results

Although ex-Gaussian analyses are of main interest, traditional analyses of RT were carried out for comparison purposes. Responses faster than 200 ms were excluded for both response time analyses. Arcsine-transformed percentages of correct responses, mean RTs and the ex-Gaussian distribution parameters μ and τ were submitted to 2 (Condition) by 3 (Group) mixed design *ANOVAs*, with Group being a between-group variable and Condition a within-group variable. Greenhouse-Geisser corrections were adopted when appropriate. Bonferroni corrections were applied for any follow-up comparisons. Because the semantic blocking effect is not seen in the first cycle of a block (Belke et al., 2005), all analyses were based on responses from the second cycle onwards.

Accuracy

Speech accuracy in the semantic blocking experiment was checked via audio recordings. Hesitations, repairs, stutters and incorrect object names (e.g. using ‘jacket’ instead of ‘coat’) were counted as errors. Most of the errors, though, were self-corrections of names of other pictures in the same experimental block. Due to a technical error, speech records from 13 of the proficient L2 group were not available. Nevertheless, speech errors were generally very rare and the mean and variance of error rates for proficient L2 speakers for which audio recordings were available was very similar to that of monolingual and bilingual speakers. We therefore conducted analyses with all three groups included.

For illustration purposes, Figure 4.1 shows the average percentage of correct responses for each participant group per condition (instead of arcsine transformed accuracy used in the statistical analysis). We found a significant main effect of Condition, F (1, 58) = 38.02, $p < .001$, $\eta_p^2 = .40$. The homogeneous condition led to more speech errors than the heterogeneous condition. The main effect of Group was not significant, F (1, 58) = .23, $p > .05$, $\eta_p^2 = .008$. The interaction between Group and Condition was significant, F (2, 58) = 4.48, $p = .015$, $\eta_p^2 = .13$. Follow-up paired sample t-tests showed significant semantic context effects for monolinguals, t (24) = 3.6, $p = .001$, and for proficient L2 speakers, t (10) = 4.34, $p = .001$. This effect was only marginal for bilinguals, t (24) = 2.0, $p = .06$. Follow-up one-way *ANOVA*s showed that accuracy rates for three groups did not differ for the heterogeneous condition, F (2, 58) = 1.04, $p > .05$, $\eta_p^2 = .04$, nor for the homogeneous condition, F (2, 58) =

1.64, $p > .05$, $\eta_p^2 = .05$. Therefore, the Condition by Group interaction was driven by bilinguals tending to suffer less from the semantic context effect.

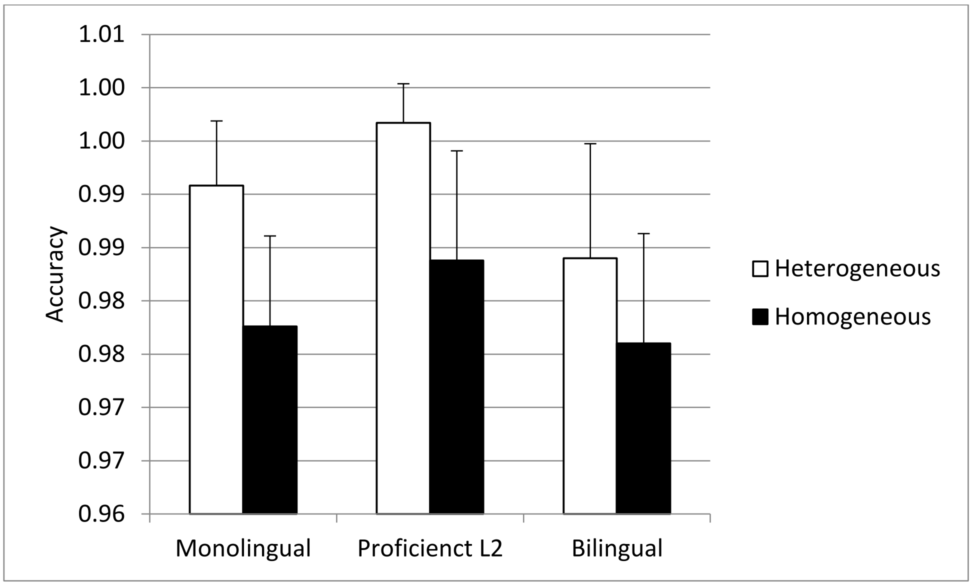


Figure 4.1. Accuracy for semantic blocking task for monolingual speakers, proficient L2 speakers and bilingual speakers in the heterogeneous and homogeneous conditions. Error bars represent 95% confidence intervals.

Reaction Times

Mean response time analyses.

Mean RT for each condition was calculated after excluding responses 2 *SD* above the mean of each participant. Figure 4.2 shows the average RT for the three participant groups for both conditions. The *ANOVA* showed a main effect of Condition, F (1, 71) = 89.55, $p < .001$, $\eta_p^2 = .56$, with slower responses in the homogeneous condition. There was no main effect of Group, F (1, 71) = 0.17, $p > .05$, $\eta_p^2 = .005$, nor a significant interaction between Condition and Group, F (2, 71) = 0.27, $p > .05$, $\eta_p^2 = .008$. Results from this analysis suggest that the competitive semantic context slowed down all three groups of participants equally. But, as will become apparent in the response distribution analyses, the picture is more complex.

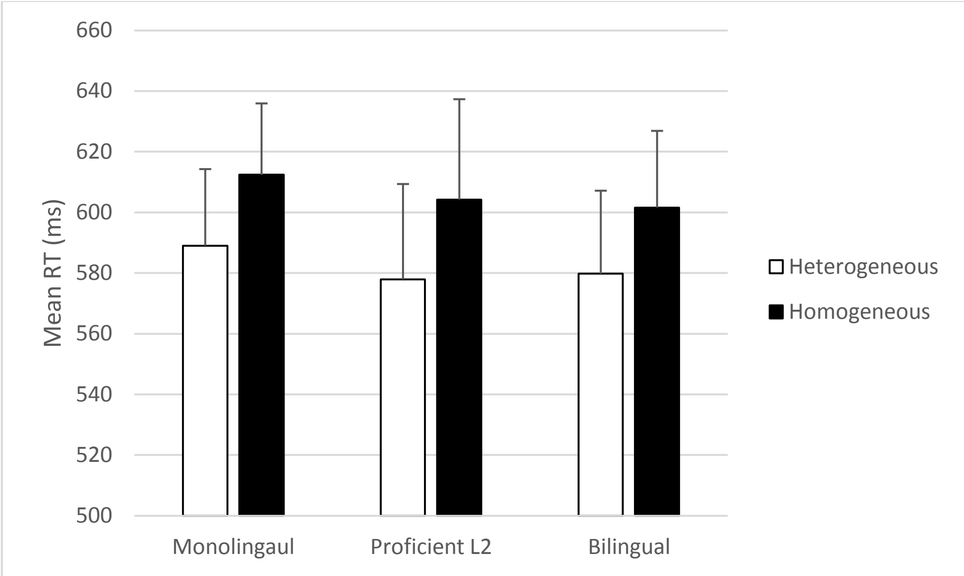


Figure 4.2. Mean RTs for monolingual, proficient L2 and bilingual speakers for heterogeneous and homogeneous conditions. Error bars represent 95% confidence intervals.

Ex-Gaussian analyses.

We analysed the distribution of RTs by fitting an ex-Gaussian distribution to response times on correct trials for each condition and each participant. We estimated the ex-Gaussian distribution parameters μ and τ using the QMPE software, which uses the quantile maximum likelihood estimation method (Brown & Heathcote, 2003). Parameters were estimated for each participant under each condition using 10 bins. All ex-Gaussian parameters were successfully yielded with an average iteration of 10.9. Parameter estimations were all trustworthy according to the technical manual since the exit codes were all below 128. One proficient L2 speaker was excluded from further analyses since the τ parameter estimate was more than 4 *SD* above the mean of all participants.

Decile analyses.

Additional Decile analyses (see Chapter 3, page 54) were carried out to make sure that the parameter estimations were reliable. Ten Decile bins were estimated because the ex-Gaussian analyses were based on ten bins. Table 4.3 summarizes the average empirical and theoretical Decile bin values for each condition and each group at bin level. 3 (Group) by 2

(Estimation Method) mixed design *ANOVAs* were conducted to test the effect of fitting method, with Group being a between group factor, and Estimation Method being a within group factor. Results are presented in Table 4.3. For half of the bins, there was a significant main effect of Method, suggesting a difference between an empirical Decile bin value and a theoretical bin value. However, such discrepancies were mostly small in value and within 1 *SE* of the empirical values, except for the last bin, for which the model predicted fewer very long responses. Closer inspection showed that numerical values of the discrepancies were very similar for all three participant groups. None of the *ANOVA*s showed a significant main effect of Group or interaction between Group and Estimation Method. Therefore, discrepancies between empirical and theoretical Decile bin values were equivalent across the three groups, ruling out the possibility that any significant group differences in the ex-Gaussian parameters could be artificially caused by the ex-Gaussian model fitting.

Table 4.3.

Empirical and theoretical Decile bin values for each condition, each participant group and each bin. F-statistics and p-values for the Main effect of Estimation Method (Method), Group and interaction between Group and Method.

		Heterogeneous condition									
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin6	Bin7	Bin8	Bin9	Bin10
Monolingual	Empirical	423.42	500.34	531.49	556.51	577.19	596.72	620.57	651.87	702.09	857.73
	Theoretical	445.05	499.28	527.51	551.22	573.78	597.15	623.33	655.60	702.13	819.91
	Discrepancy	21.63	-1.06	-3.98	-5.29	-3.41	0.43	2.77	3.73	0.04	-37.82
	SE	12.16	10.59	10.95	11.32	12.06	12.83	14.02	16.02	21.83	33.66
Proficiency L2	Empirical	411.65	485.25	518.17	541.32	564.06	586.37	609.85	640.30	711.88	861.41
	Theoretical	408.56	481.40	513.74	538.94	563.21	588.57	617.09	652.29	702.98	824.07
	Discrepancy	-3.09	-3.85	-4.42	-2.39	-0.84	2.20	7.24	12.00	-8.90	-37.35
	SE	16.17	14.24	14.09	14.09	14.75	15.62	16.44	17.67	39.90	53.24
Bilingual	Empirical	401.38	475.94	514.75	544.14	569.94	594.00	618.91	647.86	695.56	870.16
	Theoretical	393.93	454.54	510.49	538.06	563.27	589.11	617.63	652.04	700.17	814.30
	Discrepancy	-7.46	-21.40	-4.26	-6.07	-6.68	-4.89	-1.27	4.18	4.61	-55.87
	SE	16.08	13.55	12.57	12.49	13.47	14.51	15.17	16.16	18.94	30.62
Method	*F*	0.12	1.05	19.28	19.06	7.48	0.20	1.63	4.85	0.12	34.40
	p	0.73	0.31	<.001**	<.001**	0.008**	0.66	0.21	0.031*	0.73	<.001**
Group	*F*	1.68	1.90	0.53	0.37	0.22	0.11	0.08	0.05	0.04	0.02

	p	0.19	0.16	0.59	0.69	0.81	0.90	0.93	0.96	0.97	0.98
Group*Method	*F*	0.75	0.56	0.02	1.13	1.60	1.57	1.15	0.78	0.91	1.50
	p	0.48	0.58	0.98	0.33	0.21	0.22	0.32	0.46	0.41	0.23

		Homogeneous condition									
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin6	Bin7	Bin8	Bin9	Bin10
Monolingual	Empirical	426.88	515.43	549.56	576.24	602.38	628.03	654.52	685.54	740.43	922.30
	Theoretical	448.07	512.18	545.27	572.83	598.80	625.42	654.97	691.12	742.89	872.59
	Discrepancy	21.19	-3.24	-4.29	-3.41	-3.58	-2.61	0.44	5.58	2.46	-49.71
	SE	12.02	10.18	10.13	10.45	11.27	12.14	13.06	14.60	18.77	32.79
ProficiencyL2	Empirical	414.26	501.24	534.77	565.52	594.32	618.66	646.80	684.12	775.32	973.68
	Theoretical	395.30	475.15	530.81	562.14	590.62	620.58	654.60	697.11	759.19	913.08
	Discrepancy	-18.96	-26.09	-3.97	-3.37	-3.70	1.91	7.80	12.99	-16.12	-60.60
	SE	13.48	14.07	14.45	14.49	15.08	15.59	16.36	17.83	43.40	47.84
Bilingual	Empirical	411.64	491.79	530.49	562.74	590.32	619.68	649.98	691.44	756.77	969.60
	Theoretical	422.74	490.84	526.64	556.95	586.05	616.50	651.04	694.18	757.23	921.70
	Discrepancy	11.10	-0.94	-3.85	-5.79	-4.27	-3.18	1.06	2.74	0.46	-47.90
	SE	15.36	13.43	12.60	12.68	13.58	13.79	14.60	16.25	18.63	34.48
Method	*F*	0.22	1.64	11.66	20.93	17.37	0.92	2.03	7.08	0.48	27.95
	p	0.64	0.21	0.001**	<.001**	<.001**	0.34	0.16	0.01**	0.49	<.001**
Group	*F*	1.31	1.13	0.66	0.37	0.23	0.11	0.03	0.02	0.27	0.64

	p	0.28	0.33	0.52	0.69	0.79	0.89	0.97	0.98	0.76	0.53
Group*Method	*F*	1.57	1.02	0.01	0.77	0.05	1.41	1.16	1.29	0.85	0.09
	p	0.22	0.37	0.99	0.47	0.95	0.25	0.32	0.28	0.43	0.91

Note. Discrepancy = Theoretical - Empirical Decile means. *SE* = standard error of empirical response times. * signifies that the effect was significant at .05 significance level.

** signifies that the effect was significant at .01 significance level.

The μ parameter.

Figure 4.3 shows the average μ for each participant group in each condition. The *ANOVA* of μ showed no main effect of Condition, F (1, 70) = 2.68, $p > .05$, $\eta_p^2 = .037$, and no main effect of Group, F (2, 70) = .58, $p > .05$, $\eta_p^2 = .016$, but a significant interaction between Condition and Group, F (2, 70) = 3.50, $p = .036$, $\eta_p^2 = .091$. Follow-up paired sample *t*-tests for each participant group revealed that monolinguals had significant larger μ in the homogeneous than the heterogeneous condition, t (24) = 2.8, $p = .01$, while μ in the two conditions did not significantly differ for proficient L2 speakers, t (22) = 1.34, $p > .05$, or bilinguals, t (24) = .87, $p > .05$.

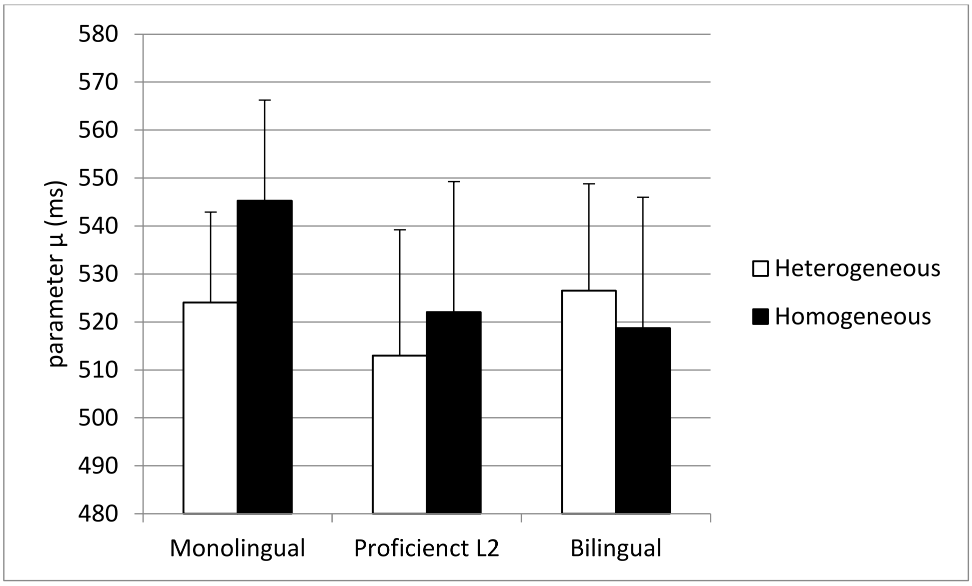


Figure 4.3. Means of the Ex-Gaussian parameter μ estimated from individual RT distributions for monolingual, proficient L2 and bilingual speakers for heterogeneous and homogeneous conditions. Error bars represent 95% confidence intervals.

The τ parameter.

Figure 4.4 shows the average τ for each participant group in each condition. The *ANOVA* of τ showed a significant main effect of Condition, F (1, 70) = 16.79, $p < .001$, $\eta_p^2 = .193$, with the homogeneous condition leading to longer tails. There was no main effect of Group, F (2, 70) = .06, $p > .05$, $\eta_p^2 = .002$. There was, however, a significant interaction between Condition and Group, F (2, 70) = 3.24, p = .045, η_p^2 = .085. Follow-up paired sample t-tests for each participant group revealed that for monolinguals, there was no significant difference in τ between the heterogeneous and homogeneous conditions, t (24) = .57, p > .05. In contrast, proficient L2 speakers had a significant larger τ and therefore longer distribution tails in the homogeneous than the heterogeneous condition, t (22) = 3.28, p = .003. The same was true for bilinguals, t (24) = 3.29, p = .003.

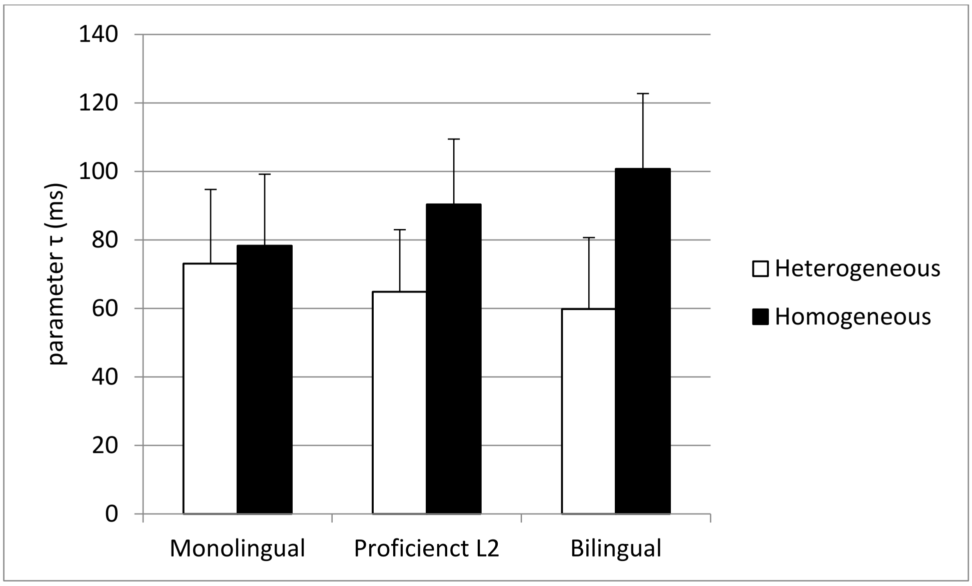


Figure 4.4. Means of the Ex-Gaussian parameter τ estimated from individual RT distributions for monolingual, proficient L2 and bilingual speakers for heterogeneous and homogeneous conditions. Error bars represent 95% confidence intervals.

Discussion

The aim of this study was to explore whether speakers who speak a second language fluently show enhanced lexical selection in a competitive semantic context when naming pictures in one language. We tested English monolingual speakers, English/Chinese bilingual speakers and native English speakers who are highly proficient in another language (proficient L2 group) in a cyclic semantic blocking paradigm. We inspected accuracies, mean RTs and reaction time distributions, the latter by focussing on μ and τ parameters obtained from ex-Gaussian analyses. In terms of accuracy, bilingual speakers did not suffer as clearly from the semantic context effect as the other two groups, supporting our hypothesis of an enhanced ability to resolve lexical competition in bilinguals. Results of the traditional analysis of mean RTs showed that all three groups suffered the semantic interference to the same extend. Results of the ex-Gaussian analysis, however, showed that the locus of the semantic blocking effect within the response distributions differed across the participant groups. Monolingual speakers showed a significant semantic blocking effect in μ but not τ, while bilingual speakers and proficient L2 speakers showed the reverse. In what follows, we first discuss in detail how the results support the hypothesis that participants who are fluent in a second language have better executive control in resolving lexical competition. We then dismiss a potential alternative interpretation and discuss the roles of other factors in this task. Last but not least, we discuss how the results could be explained by an incremental learning mechanism that does not assume competition during lexical selection.

Gaussian Part of Response Distributions: Enhanced Resolution of Lexical Competition

The results for μ suggest that monolingual speakers, but not bilingual speakers or proficient L2 speakers, were uniformly slowed down by the competitive semantic context in the homogeneous condition compared to the heterogeneous condition. Thus lexical selection appeared to be more effortful in the homogeneous condition, when closely related lexical items were highly active and competed for selection (Abdel Rahman & Melinger, 2009). This was not the case for bilingual speakers or proficient L2 speakers for whom μ did not differ across the two conditions. This is in accordance with the hypothesis that speaking a second language fluently promotes the ability to resolve lexical competition within a language.

This enhanced ability to resolve lexical competition within a language is arguably achieved through the constant exercise of managing the competition from the other language.

Thus, individuals who speak another language fluently might have developed a superior ability to monitor and resolve any competition between lexical items. As we have introduced, bilingual speakers have to constantly resolve the dual activation of both languages. One way to resolve such between-language competition is through a top-down control mechanism such as the inhibitory control model proposed by Green (1998). In his model, a supervisory attentional system inhibits the non-target language through language tags at lemma level. Evidence from picture-word interference, language switching, and the effects of cognate status on picture naming support the claim that bilinguals actively inhibit the non-target language during language production (for a review see Kroll, Bobb, Misra, & Guo, 2008). Our results suggest that such top-down modulation of lexical activation might not only regulate lexical competition across languages but also within a single language. Through the intense inter-language exercise in individuals that speak another language fluently, the inhibitory control might become more efficient and therefore might benefit speakers' intra-language resolution of lexical competition.

The involvement of inhibition ability in picture naming had also been suggested by Shao, Roelofs, and Meyer (2012), who found that RTs in a stop signal task positively correlated with RT distribution parameteres when naming objects and actions. More specifically, they found a significant correlation between inhibition ability and μ of the response distributions for the more demanding action naming task, while inhibition ability correlated with τ in the simple object naming task. This suggests that inhibition ability modulated the response speed on most of the trials in the action naming task, but only rarely in the object naming task. It therefore appears that inhibition ability might be particularly important in demanding contexts when the activation level of lexical competitors are high, which is the case for the semantic blocking task, particularly in the homogenous condition.

The results for μ are conceptually consistent with the finding that language proficiency of bilingual speakers modulates the degree of the Colour Stroop effect in μ (Tse & Altarriba, 2012). Tse and Altarriba (2012) reported that more proficient speakers in L1/L2 have smaller Colour Stroop effects than less proficient speakers. Our results are in line with this finding if we regard our participant groups as being taken from a continuum of second language proficiency: on the one end of the continuum bilinguals with very high proficiency and on the other end monolinguals with, if at all, very poor proficiency. As pointed out in the introduction, the Colour Stroop effect does not necessarily reflect the ability to resolve lexical competition. Nevertheless, Tse and Altarriba (2012)'s results do predict that high proficiency speakers should show reduced interference effect, be it the interference between tasks or lexical items.

Exponential Part of Response Distributions: Executive Control or Covert Errors?

The results for τ suggest that the semantic context affected extreme responses of bilingual and proficient L2 speakers, but not of monolingual speakers. The parameter τ reflects both the number and the degree of long responses, therefore monolingual speakers showed similar amount and/or degree of long responses in both conditions. In contrast, bilingual and proficient L2 speakers had a larger amount of extreme responses and/or more extreme responses when the semantic/lexical competition was strong.

As we have introduced, long responses can be due to poor general attentional control that results in more frequent lapses of attention or due to covert errors. In what follows, we will argue that our results for τ are rather consistent with the latter.

Attentional control.

In previous studies that used ex-Gaussian analyses in (non-)verbal conflict tasks, long responses have been related to working memory limitations (Schmiedek et al., 2007) or the attentional control mechanism that maintains the task goal in the working memory, the failure of which leads to lapses of attention (Hervey et al., 2006; Leth-Steensen et al., 2000; Unsworth, Redick, Lakey, & Young, 2010). According to this interpretation, the longer distribution tails in the homogeneous condition would mean that general attentional control was disrupted for proficient L2 speakers and bilingual speakers when there was strong lexical competition. We think this interpretation is inappropriate for the results at hand for three reasons. First, differences of general attentional control abilities between participant groups should affect both conditions equally and therefore should lead to a main effect of participant group, which is not what we found. Second, the implication that bilingual and proficient L2 speakers might have worse attentional control in the homogeneous condition is at odds with various claims that speaking a second language promotes attentional control ability (e.g. Hernandez et al., 2012; Mishra et al., 2012; Tse & Altarriba, 2012), see also Chapter 3. Third, the increase of distribution tails in a condition with strong competition is at odds with the conflict monitoring hypothesis (Botvinick et al., 2001; Botvinick et al., 1999). This hypothesis predicts that conflicts should elicit higher levels of cognitive control. For instance, it has been shown that the level of cognitive control can be elevated following a trial with interference (Botvinick et al., 2001; Botvinick et al., 1999). In other words, a more competitive semantic environment should make one more vigilant and should rather lead to fewer lapses of attention (shorter distribution tails) instead of more (longer distribution tails).

In addition to these arguments, we sought independent evidence that long responses in our experiment are not due to lapses of attention. For that we analysed the slowest 10% responses of each participant. If these responses were mainly driven by lapses of attention, we would expect them to occur equally likely across the 16 experimental items. On the other hand, if they were covertly corrected errors as argued below, we would expect them to occur more often for some items than for other items. Chi-square tests of equality suggested that long responses did not occur equally likely across the 16 words, neither in the heterogeneous condition, χ^2 (1, N=16) = 135.5, p < .001, nor in the homogeneous condition, χ^2 (1, N=16) = 168.4, p < .001. These results suggest that some words were more likely to lead to long responses than others. This speaks against the idea that the distribution tails were mainly driven by lapses of attention.

Covert errors.

After considering alternative explanations, we now turn to the option that longer response tails in the homogenous condition for bilingual and proficient L2 speakers were due to covertly corrected errors. Because τ reflects both the proportion and the degree of extreme responses, our result pattern for τ could either mean that bilingual and proficient L2 speakers exhibited more covert errors in the homogeneous condition than the heterogeneous condition or they took longer to correct them. It is unlikely, though, that they made more covert errors because that should have been reflected in their overt error rates. There was no indication that bilingual or proficient L2 speakers' responses were particularly erroneous in the homogeneous condition. Therefore, it is more likely that the long tails in the homogenous condition mean that bilingual and proficient L2 speakers took longer to correct their errors in a condition of strong lexical competition.

We propose that the reason for these very long responses is that bilingual and proficient L2 speakers very strongly inhibited non-target lexical items in the homogeneous condition. As mentioned in the introduction, there is evidence that competitors or non-selected lexical items of a previous trial are actively inhibited in a naming task (Tipper, 1985; Tipper & Driver, 1988). If this inhibition occurred very fast, i.e. during the preparation of the response, then the time to recover from an internal error during internal monitoring processes could depend on the degree of inhibition to the target lemma: the more the non-selected lemma was inhibited, the longer it would take to recover from such inhibition. The results for τ therefore most likely suggest that bilingual and proficient L2 speakers experience more difficulties recovering from covert errors. This implies a stronger inhibition to the competitors in the homogeneous condition than in the heterogeneous condition.

To conclude, after considering alternative explanations, we argue that the long response tails for bilingual and proficient L2 speakers in the homogenous condition are most likely due to covertly corrected errors. When taking the results of μ and τ together, the picture becomes clearer. The reduced semantic blocking effects in μ for bilingual and proficient L2 speakers suggest that they resolve lexical competition better than monolingual speakers, arguably through a more efficient top-down control mechanism that inhibits competitors more strongly than that of monolingual speakers, at least in the case of strong lexical competition. However, this stronger inhibition to competitors has the side effect that it elevates the difficulty to recover from an internal error, resulting in a longer response distribution tail in the homogeneous condition.

Alternative Interpretation: Impulsive Behaviour

One potential alternative interpretation of our results is that bilingual and proficient L2 speakers might be more impulsive, meaning that they have overall faster responses and, as a consequence, more covertly corrected errors reflected by longer distribution tails. There are two problems with this interpretation, though. First of all, if bilingual and proficient L2 speakers were more impulsive, they should have been faster (as reflected in their values for μ) in both homogeneous and heterogeneous conditions, which is not the case. Second, if bilingual and proficient L2 speakers made more covertly corrected errors, one would expect them to make more overt errors as well. But this was not the case either, neither in general, nor when only considering the homogeneous condition. Therefore, we conclude that it is not likely that the observed result pattern is due to bilinguals and proficient L2 speakers being more impulsive.

Potential Effects of Other Factors

Vocabulary size and language proficiency.

As mentioned in the introduction, vocabulary size could potentially play a critical role in the cyclic naming task, given that the semantic blocking effect transfers to new category members other than the few items named in a homogenous block (Belke et al., 2005). This predicts that participants with a larger vocabulary size, especially in the semantic categories used in the present experiment, should show a larger semantic blocking effect. In our study, monolingual speakers had on average a larger (category) vocabulary size than bilingual speakers, meaning that they potentially activated more competing items in the homogenous blocks than bilingual speakers. While this could be a potential explanation for the lack of a semantic blocking effect in the bilingual speakers in the Gaussian part of their response distribution, the results of the proficient L2 speakers rule this explanation out. In spite of having similar vocabulary sizes as the monolingual group, the proficient L2 group showed a response distribution profile resembling that of the bilingual group, not of the monolingual group. Therefore, vocabulary size does not seem to play a role in this task. Following the same logic, the result pattern also excludes the possibility that any group differences were due to differences in English proficiency.

Given the variability of their second language profile and generally later onset of their second language, it is interesting that the result pattern of the proficient L2 speakers resembles that of the bilingual group. This suggests that the practise of speaking two languages and being able to speak them fluently is sufficient for enhanced lexical retrieval ability in competitive semantic context to emerge, and this does not rely on very early exposure to a second language.

Cognates.

One might wonder whether the results have been affected by cognates in the experimental material. It is often found that bilingual speakers name cognate words faster than non-cognate words (Costa, Caramazza, & Sebastian-Galles, 2000). The English/Chinese bilingual group could not have been affected by cognates because there were no English/Chinese cognates in the experimental material. The proficient L2 group, on the other hand, spoke various languages and some of them have cognates for the experimental items (e.g. ‘lamp’ in English and ‘Lampe’ in German). Very similar result patterns for the semantic blocking effect on μ and τ for bilingual and proficient L2 speakers suggest that cognates did not affect the result pattern. However, to investigate a potential effect of cognate knowledge on the semantic blocking effect, we conducted a correlation analysis of the number of cognates in the stimuli for each participant and their semantic blocking effect in μ, r (24) = -.09, $p > .05$, and τ, r (24) = .08, $p > .05$, which suggested no relationship.

Interference from the other language.

Given the parallel activation of two languages in bilingual and L2 speakers, we also considered two ways in which interference from the second language might have affected task performance: lexical competition and increased executive control demands. First, results by Runnqvist and colleagues suggest that lexical competition can occur across languages (Runnqvist, Strijkers, Alario, & Costa, 2012), meaning that lexical items of the second language might compete for selection with the target items in our experiment. However, the effect from the ‘other’ language was, if at all, minimal. We did not observe a single intrusion error for bilingual or L2 speakers. Also, if the competition from the second language had been strong, it should have led to slower responses in both bilingual and proficient L2 speakers and in both experimental conditions.

Second, the need to manage and monitor language production when a second language is active might pose extra cognitive demands resulting in less efficient executive control. Belke and Stielow (2008) found that posing extra working memory load onto the cyclic blocking paradigm led to an overall increase in response time and a larger semantic blocking effect. Thus, if the demand for juggling two languages was significant in our experiment, bilinguals and proficient L2 speakers should have responded more slowly overall than monolinguals, and the additional cognitive demand should have led to a larger semantic blocking effect in bilinguals and L2 speakers. Neither was the case.

Not observing any significant effects of the other language in our task is maybe less surprising than it first looked. It can be explained, for instance, by the biased selection system proposed by Belke (2008). If this biased selection system is able to bias selection towards a few exemplars from a single semantic category, then it is not unreasonable to assume that such a top-down system is also able to bias selection towards a single language, thus minimizing the effect from the non-target language. Alternatively, the excellent top-down control of lexical competition in our bilingual and fluent L2 speakers might have been so efficient in inhibiting activation from the 'other' language that no effect was visible.

Lexical Selection without Competition

The account of the semantic blocking effect discussed so far assumes that lexical selection is competitive. In many word production models, selection by competition is assumed, such as in the WEAVER++ model (Levelt, Roelofs, & Meyer, 1999; Roelofs, 2003), the connectionist model by Starreveld and la Heij (1996) and the speech-activation theory by Roelofs (1992). Furthermore, Howard, Nickels, Coltheart, and Cole-Virtue (2006) proposed an elegant computational model to capture cumulative semantic interference, i.e. naming a picture slows down the subsequent naming of other words from the same semantic category. In their model, lexical competition is essential to predict experimental outcomes.

However, alternative views of word production hold that lexical selection can be achieved without competition (Dell, 1986; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007; Miozzo & Caramazza, 2003). Oppenheim, Dell, and Schwartz (2010) proposed a computational model to account for semantic interference effects without incorporating any lexical competition. In their model, a learning mechanism modifies the weights of lexical-semantic links after each successful lexical selection. It increases the connection weights from active semantic features to the target word and decreases the connection weights to co-activated non-target words. In a semantic blocking experiment, only words within a homogenous block, not within a heterogeneous block, share some semantic features. Therefore, the weights of non-target words in a homogeneous block are reduced each time a word within the block is named, while this is not the case in the heterogeneous condition. Since lexical selection in this model is non-competitive, the selection depends solely on the absolute level of activation of a word. Because the weights of lexical-semantic links to words that are semantically related are decreased to a stronger degree in the homogenous condition than in the heterogeneous condition, it takes longer for words to reach the activation threshold in the homogeneous condition, leading to the semantic blocking effect.

We next consider how our results might be explained by such an incremental learning model (Oppenheim et al., 2010). Two features of this model appear to be relevant: first, an incremental learning process modifies the semantic-to-lexical connections within the lexical network at the end of each trial. It strengthens connections between the semantic features and the target word, while it weakens connections between the semantic features and other activated words. Second, the model features a booster process, which positively boosts the activation level of all candidates until a candidate reaches the selection threshold.

We can see two options of how our results might be accounted for with respect to each of the two features. One option draws on the learning rate parameter η, which is used to calculate the weight change for the connection from semantic nodes to lexical nodes after each trial; it adjusts the magnitude at which connection weights change. This parameter might have a value close to zero for bilingual speakers and L2 speakers, meaning that the semantic-to-lexical connections of the targets were not strengthened much and those of the non-targets were not weakened much. Consequently, one would not see a semantic blocking effect in μ. It is, however, questionable whether it would make sense to have a learning mechanism that has almost no effect on the system as a whole. In addition, if the learning rate was close to zero, this would mean that the homogeneous condition was no longer different from a heterogeneous condition. It would therefore be difficult to explain the longer response distribution tails in the homogeneous condition that we found for bilingual and proficient L2 speakers.

Alternatively, it could be that all three participant groups have similar learning rates but bilingual and L2 speakers have a more efficient lexical booster. A lexical booster is proposed to aid lexical selection when there are highly active alternatives. However, in order to explain our results, additional features of the booster would need to be assumed. To explain the results for μ, the boosting factor would need to be flexible in that it could vary depending on the semantic context: it would need to boost items harder when competition is stronger. This is in line with the proposal by Oppenheim et al. (2010) that a booster operates more strongly when selection is more difficult. Therefore, if bilingual and proficient L2 speakers had a more efficient booster, which means they were able to boost items more strongly than monolingual speakers in the homogeneous condition, their homogeneous targets might reach the activation threshold faster than those of monolinguals'. This could mean that bilinguals and proficient L2 speakers responded to homogeneous targets as quickly as to heterogeneous targets, as found in our experiment. In addition, in order to account for our results for τ, the lexical booster seems to need the feature of anti-boosting (inhibiting) activation of the non-targets (for a similar idea in relation to inhibition across two languages see Runnqvist et al., 2012). Again, if bilingual and proficient L2 speakers had a more efficient booster which means they anti-boosted the non-selected items more strongly, then they would need more effort to reboost the target than monolingual speakers in case of a covert error.

Oppenheim and colleagues have linked the booster to the function of the left inferior frontal gyrus (Oppenheim et al., 2010), which has been associated, among others, with the selection of information among competing alternatives (Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). In other words, the booster can be understood as a top-down modulation mechanism that directs activation towards lexical items defined by a task goal and that aids selection of appropriate representations. Therefore, regardless of the approach one takes (lexical selection by competition or not), the current result pattern suggests that through managing more than one language, speakers develop a more efficient top-down control mechanism, which facilitates selection when there are competing alternatives. Importantly, such ability derives from controlling more than one language and can be used to resolve competition within a language. Critically, lexical inhibition or anti-boosting (inhibition) of non-targets is an essential feature of such a top-down control mechanism in order to fully account for the current results.

Conclusion

In conclusion, we tested monolingual, proficient L2 and bilingual speakers in a cyclic semantic blocking naming task. All three groups suffered from the semantic interference effect while their response distributions were differently affected. Monolingual speakers showed the effect in the Gaussian part of the response distribution, meaning that they were universally slowed down by the competitive semantic context. In contrast, bilingual and proficient L2 speakers showed semantic interference effects in the exponential part of their response distributions in that they had longer extreme responses in the homogeneous than heterogeneous condition. We have argued that these results suggest an enhanced ability to resolve lexical competition for bilingual and proficient L2 speakers within a competitive lexical selection account or a more efficient lexical booster for bilingual and proficient L2 speakers within a non-competitive lexical selection account. But whatever the underlying processes, the enhanced performance in competitive semantic contexts suggests a top-down control mechanism that guards against lexical intrusions and prevents responses from being slowed down when alternatives are highly active. As a side effect, it elevates the difficulty to recover from an initial wrong lexical selection.

Last but not least, we would like to point out that the response distribution analysis shows that instead of giving rise to an overall advantage or disadvantage, bilingualism causes changes in the distribution profile, which cannot be observed through a traditional central tendency analysis. Therefore, ex-Gaussian analyses can provide richer information than analyses of central tendencies and we suggest using this analytical approach more often.

CHAPTER 5

DIFFERENCES IN RESOLVING SEMANTIC INTERFERENCES BETWEEN MONOLINGUAL AND BILINGUAL SPEAKERS - ELECTROPHYSIOLOGICAL EVIDENCE

Abstract

The present study investigated electro-physiological evidence for the enhanced engagement of executive control in bilingual speakers to resolve lexical competition within a language. English monolingual speakers and bilingual speakers of English and another language were tested using the cyclic semantic blocking task while electroencephalograms (EEG) were recorded. Behaviourally, the two groups performed similarly, but their EEG showed divergent results. While both monolingual and bilingual speakers showed a posterior ERP effect between two experimental conditions around the time of lexical selection, this effect lasted longer for monolingual speakers. More importantly, only bilingual speaker showed a slightly earlier central-frontal ERP effect, and only monolingual speakers showed a significant later effect. We argue that these results suggest a stronger engagement of executive control for bilingual speakers, which facilitates the resolution of intra-language lexical competition.

Introduction

It is a generally accepted fact that the two languages of bilingual speakers are both active regardless of the language in use (e.g. BijeljacBabic, Biardeau, & Grainger, 1997; Thierry & Wu, 2007; Wu & Thierry, 2010; Hoshino & Thierry, 2012). This makes speech production a more demanding task for bilinguals since, in contrast to monolinguals, they need to select a target word among related lexical items in the same language and the selection might be affected by alternatives in their other language. This is in line with their performance on tasks that require lexical access. Bilingual speakers have been found to name pictures more slowly (Gollan, Montoya, Fennema-Notestine, & Morris, 2005). They also usually perform worse on verbal fluency tasks (Rosselli, Ardila, Salvatierra, et al., 2002; Sandoval, Gollan, Ferreira, & Salmon, 2010), where participants are given one minute to produce as many words as possible that start with a particular letter.

Despite showing disadvantages in verbal tasks, bilingual speakers have been suggested to have some cognitive advantages. As a result of the constant exercise of language monitoring and selection, bilinguals have been argued to have developed superior executive control abilities (for a review, see Bialystok, Craik, Green, & Gollan, 2009). Various studies have found that bilingual speakers outperform monolingual speakers in various non-verbal conflict tasks (e.g. Bialystok, 2006; Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernandez, & Sebastian-Galles, 2008; but see Paap & Greenberg, 2013). This superior performance has presented itself either as a reduced interference effect or overall faster responses (for a review, see Hilchey & Klein, 2011), and these have been attributed to stronger inhibition ability of distracting information (Bialystok et al., 2004; Luk, Anderson, Craik, Grady, & Bialystok, 2010), better attentional control ability (e.g. Costa et al., 2008) or, very broadly, as enhanced executive control (Bialystok, Martin, et al., 2005; Bialystok et al. 2004).

Executive control and self-monitoring are not only important in non-verbal interference tasks, they also play a central role during language production (e.g., Wheeldon & Levelt, 1995; Shao, Roelofs, & Meyer, 2012). Therefore, if managing more than one language in bilingual speakers leads to enhanced executive control, we would expect that bilingual speakers benefit from this ability when resolving lexical competition within a language. On first sight, this hypothesis might sound counter-intuitive, given bilinguals' worse performance on picture naming and verbal fluency tasks. However, enhanced processing demands due to the control of two languages might mask their enhanced competition resolution ability within one language. Furthermore, Luo et al. (2010) have shown with verbal fluency tasks, the worse performance in verbal fluency tasks is likely due to the smaller vocabulary of bilinguals and the fact that bilinguals have fewer opportunities to use the words of a given language. It therefore does not contradict an enhanced bilingual competition resolution ability.

In line with this hypothesis, we have found evidence that bilinguals are better at resolving lexical competition within a language (Chapter 4), using the cyclic semantic blocking paradigm (Damian, Vigliocco, & Levelt, 2001; Kroll & Stewart, 1994). In this paradigm participants name pictures that are blocked into either homogenous category blocks (pictures of words from *the same* semantic category) or heterogeneous category blocks (pictures of words from *different* semantic categories). Naming speed is typically longer in homogeneous blocks compared to heterogeneous blocks (Belke, Meyer, & Damian, 2005; Damian et al., 2001; Kroll & Stewart, 1994). This can be explained by the spreading of activation within the semantic network (Collins & Loftus, 1975). Because items in a homogeneous block are from the same semantic category, they strongly activate each other. This is not the case in heterogeneous blocks. Assuming that the speed of selecting a target word depends on the relative activation level of the target and its competitors (Levelt, Roelofs & Meyer, 1999; Roelofs, 2003; Starreveld & La Heij, 1996; Roelofs, 1992), one can understand the semantic blocking effect as the result of stronger lexical competition between semantically related items than unrelated items (but see Oppenheim, Dell, & Schwartz, 2010 for an alternative account). Thus, the more efficiently one resolves the competition, the smaller the semantic blocking effect should be. Accordingly, we found that the blocking of semantically related pictures affected the responses of monolinguals and bilinguals differently. When investigating effects on response distributions, we found that, unlike monolinguals, bilinguals did not show a semantic blocking effect in the main body of the response distribution, but in the tail, i.e. the part that reflects very slow responses. We have argued that these results suggest that bilinguals have enhanced executive control that facilitates the resolution of intra-language lexical conflict.

The aim of the present study was to gather electro-physiological evidence for the enhanced executive control ability in bilinguals when resolving lexical conflict within a language. We used the same cyclic semantic blocking paradigm as Chapter 4, and recorded electroencephalograms (EEG) while participants were performing the task. We investigated executive control by examining event-related potentials elicited by the pictures during the preparation of responses.

Electrophysiological Effects of Interest

The EEG effect of semantic interference has been studied to date using a picture word interference paradigm (e.g. Dell'Acqua et al., 2010; Hoshino & Thierry, 2011; Piai, Roelofs, & van der Meij, 2012), a semantic blocking paradigm (Aristei, Melinger, & Rahman, 2011; Ganushchak & Schiller, 2008b; Janssen, Carreiras, & Barber, 2011; Maess, Friederici, Damian, Meyer, & Levelt, 2002), and a cumulative semantic interference paradigm (Costa, Strijkers, Martin, & Thierry 2009). We will focus our discussion here on the semantic blocking and the cumulative semantic interference literature because they are the most relevant for the current study, even though results using the picture word interference paradigm are also compatible with our predictions. Within the semantic blocking literature, studies have found effects in two time-windows: a) an early effect (from about 150-200 ms onwards), which has usually been interpreted to reflect lexical selection, and b) an effect closer to speech onset, which has been related to self-monitoring.

Electrophysiological effect of lexical selection.

In an MEG study using the semantic blocking paradigm in German, Maess et al., (2002) reported an effect between 150 and 225 ms after stimulus onset. More precisely, the homogeneous and heterogeneous conditions differed significantly with regards to the activation of the left middle temporal region, with the homogeneous condition showing more negative activation. Indefrey and Levelt’s meta-analysis of word production (Indefrey, 2011; Indefrey & Levelt, 2004) estimates that conceptually driven lexical selection takes place between 150 – 275 ms. Therefore, the authors interpreted their effect as being related to lexical selection.

Similarly, in a German EEG study using a combination of a semantic blocking and a picture-word interference paradigm, Aristei et al. (2011) found increased negative activity in the homogeneous condition compared with the heterogeneous condition over frontal and temporal regions, even though the effect occurred later, i.e. between 250 and 400 ms. The distracter effect started to emerge earlier, around 200 ms. The authors concluded that both effects reflect the same mechanism and that this mechanism is related to lexical selection. However, due to the fact that pictures were presented concurrently with a word distracter, the comparability of their results to results of other studies remains debatable. This is the case for both the timing of the effect (which could have been delayed due to the parallel processing of the picture and the distractor word) and the nature of the effect, which might reflect not only the processing of the picture, but also that of the distractor.

Maybe the most comparable study to the current one is the EEG study by Janssen et al. (2011), who tested Spanish speakers in a cyclic semantic blocking task. They found a less negative wave for the homogeneous condition 220 - 450 ms after picture onset, with a frontal distribution. They also conducted a blocked naming experiment with word stimuli instead of pictures. They found that in the word naming experiment, the semantic context effect emerged around 300 ms, which was later than the picture naming condition. They argued that these results meant that the semantic context effect reflects processing at the semantic processing stage rather than the word selection stage. Though interpreted differently, the timing of the effect is consistent with other studies and corresponds to the lexical selection stage specified by Indefrey and Levelt's word production model (Indefrey, 2011; Indefrey & Levelt, 2004). Therefore this result might also be accounted for by the lexical selection interpretation.

Another evidence for the timing of lexical selection stems from a slightly different but related paradigm namely the cumulative semantic interference paradigm (Howard, Nickels, Coltheart, & Cole-Virtue, 2006). In this paradigm, words from different categories are shuffled and named only once with no repetition. The naming speed of a picture increases as the serial position of the picture in a certain category increases. Costa, Strijkers, Martin, and Thierry (2009) found that the ERP amplitude for pictures at different serial positions started to diverge around 200 ms after picture onset and lasted till 380 ms, with posterior P2 and P3 peak amplitude increasing with increasing serial order. Again, when assuming that the semantic interference effect reflects difficulty during lexical retrieval, this result can be taken as evidence that the brain started to engage in lexical selection around 200 ms. Furthermore, Costa et al. (2009) found a point by point correlation between the ERP amplitudes and response times. Specifically, ERP amplitudes started to correlate reliably with behavioural RT from 208 ms after picture onset for 180 ms. The authors interpreted these results as reflecting lexical competition to be accessed among closely related lexical items.

Electrophysiological effect of self-monitoring

Maess et al. (2002) reported an additional effect in a later time window, namely 450 - 475 ms, which they interpreted as increased monitoring because it fitted with the time estimate of self-monitoring processes. According to Indefrey and Levelt's (2004) meta-analysis, self-monitoring during simple picture naming starts after syllabification, i.e. around 355 ms. One goal of self-monitoring is to ensure the semantic appropriateness of the selected lexical item (Maess et al., 2002). Because competitors in the homogeneous condition in the cyclic semantic blocking paradigm are from the same semantic category as the target, the homogenous condition poses a higher monitoring demand. Interestingly, Figure 3A in Janssen et al. (2011) also showed a narrowly distributed transient semantic blocking effect between 475 and 500 ms. This effect was not discussed in the article, but its timing corresponds to that reported in Maess et al. (2002) and might therefore reflect the same increase of self-monitoring demand.

Predictions

To summarize, we have identified two time windows in which semantic blocking effect have occurred: an early window that starts around 150 - 200 ms, which has been attributed to lexical access, and a later window around 400 – 450 ms, which has been related to verbal self-monitoring.

Results from Chapter 4 have suggested that bilingual speakers do engage stronger executive control (e.g. inhibition) during lexical selection. Therefore, bilingual speakers were expected to show activities for enhanced executive control processes around the time of lexical selection. The enhanced control processes should have consequences for the lexical competition effect, namely a reduced semantic blocking effect in the bilingual group. This could manifest itself as a reduced difference in amplitude and/or duration.

We also expected to observe a reflection of higher self-monitoring demand in the homogeneous condition compared to the heterogeneous condition, as we have reviewed above. If bilinguals are more efficient at resolving lexical competition, that is if they suppress competitors more strongly than monolinguals, then they should show reduced response conflict after lexical selection and thus reduced ERP effects related to self-monitoring. Again, this could show up as a reduced difference in amplitude and/or duration.

Last but not least, a correlation between ERP activities and behavioural response times was expected. Particularly, if ERP effect reflects increased on-line processing load due to lexical competition, it should correlate with the RT effect at behavioural level, which reflects the consequences of such competition. It is expected for both monolingual and bilingual speakers. If bilingual speakers are better at resolving lexical competition, which means the lexical selection process is curtailed, then the correlation for bilinguals is expected to last for a shorter time window than for monolinguals.

Examining ERPs, especially close to response onset, in an overt speech production task is controversial due to the contamination of the EEG by motor artefacts of speaking (Laganaro & Perret, 2011; Porcaro, Medaglia, & Krott, 2015; Riès, Janssen, Burle, & Alario, 2013), even though it has been suggested that speech production studies using EEG can lead to meaningful results (Ganushchak, Christoffels, & Schiller, 2011). Therefore, in order to ensure that we are investigating cognitive processes instead of motor activities, we cleaned our data with the Speech Artefact Removal ICA procedure (SAR-ICA) recently proposed by Porcaro et al. (2015).

Methods

Participants

48 participants took part in the experiment: 24 native English speakers and 24 bilingual speakers. They were undergraduate or postgraduate students at the University of Birmingham who participated either for course credits or £20 cash. All participants were right-handed as determined by the Edinburgh handedness inventory (Oldfield, 1971) and had normal or corrected-to-normal vision.

To be classified as bilingual, the following criteria had to be met: the participant (a) learnt English and another language before age 10, (b) was fluent in both languages (i.e. had rated the overall proficiency as more than 5 on a scale of 7 in both languages), and (c) used both languages on a daily basis at the time of the experiment, either in the same setting or in different settings. 20 participants met these criteria. They spoke various other languages (Bulgarian (5), French (2), Romanian (2), Indonesian (2), Armenian (1), Danish (1), Greek (1), Hungarian (1), Italian (1), Latvian (1), Polish (1), Swedish (1) and Vietnamese (1)). Monolingual English speakers were defined as follows: the participant (a) did not speak another language fluently (i.e. proficiency level of another language, if any, was 3 or below on a scale of 7), (b) did not speak another language on a daily basis, and (c) did not learn another language before age 10. Twenty monolingual speakers met these criteria.

Table 5.1 shows a summary of participants' demographic information and language background. The two groups were matched for age, t (38) = .23, p > .05, education, X^2 (2, N = 40) = 1.5, p > .05, and IQ measured by the Raven's Standard Progressive Matrices, t (36) = .95, p > .05. They were not matched in terms of their English proficiency, t (38) = 2.79, p = .008, with monolingual speakers' self-ratings being slightly, but significantly higher.

Table 5.1

Demographic information and language background for monolingual and bilingual speakers.

	Monolinguals	Bilinguals
N (Female)	20 (13)	20 (16)
Age (SD)	20.6 (2.7)	20.8 (2.9)
Median Education (Min: Max)	2 (2:4)	2 (2:4)
Mean Score of Raven's Standard Progressive Matrices	52.6 (4.8)	50.9 (5.9)
Mean Proficiency Rating for English (SD)	6.8 (.41)	6.2 (.73)
Mean Proficiency Rating for Other Language (SD)	-	6.8 (.52)
Mean Age of English Onset (SD)	0	5.0 (2.8)
Mean Age of Onset for Other Language (SD)	-	0.4 (1.6)

General Design and Procedure

This experiment had two sessions. In a first session, participants completed the cyclic semantic blocking paradigm while EEG was recorded. They filled in questionnaires (Appendix A-4) concerning their language use during the EEG preparation stage (i.e. while electro-gel was applied and electrodes were mounted to the cap). In a second session on a different day, participants completed the Standard Raven's Progressive Matrices test.

Semantic Blocking Task

The experiment materials and procedures were exactly the same as that described in Chapter 4 (See page 80).

EEG, EMG and EOG Recording

Electroencephalograms (EEG) were acquired using a 128 channel BioSemi Active Two EEG system. Active electrodes have Ag/AgCl electrode tips. Electrode positions were arranged according to the 10–5 system (Oostenveld & Praamstra, 2001). As active electrodes were used, the level of DC offset is used as an alternative indicator for the quality of the electrode contact instead of impedance. Electrode offsets were made sure to be below 20 mV before starting a recording session. Horizontal and vertical electrooculograms (EOG) were recorded, as well as upper and lower lip electromyograms (EMG). The latter was done as part of the SAR-ICA artefact attenuation procedure (Porcaro et al., 2015). In line with this procedure, we placed surface electrodes at the left orbicularis oris superior (OOS) and left orbicularis oris inferior (OOI), half way between the centre and the corner of the mouth. EEG, EOG and EMG were sampled at 512 Hz. The Active Two AD-box has a built in online low-pass digital filter, which has a 5th order sinc response with a -3 dB point at 102 Hz. The signals were recorded with online reference to the Common Mode Sense (CMS) and were then offline re-referenced in the data analysis (see EEG data pre-processing section below).

Raven's Standard Progressive Matrices

The Raven's Standard Progressive Matrices were used in a timed fashion (Raven, Raven, & Court, 2000). Each participant was allowed 25 minutes to attempt the test. They were instructed that only correct responses count and they were also encouraged to attempt as many items as they could.

Data Analysis

We compared behavioural outcomes as well as stimulus-locked ERPs in heterogeneous blocks with those in homogeneous blocks. Because the semantic blocking effect starts to emerge from the second cycle in a presentation block (Belke et al., 2005), both behavioural and ERP analyses were based on responses from the second presentation cycle onwards. For all analyses, we included only trials with correct responses (see Accuracy section below), and trials with RTs faster than 200 ms were excluded. In addition, trials with RTs beyond 2 *SD* of the mean of each participant were excluded for the mean response time analyses and ERP analyses, while all responses were included for the ex-Gaussian distribution analysis.

For the behavioural analyses, arcsine-transformed percentages of correct responses, mean response times and the ex-Gaussian distribution parameters μ and τ were submitted to 2 (Condition) by 2 (Group) mixed design *ANOVAs*, with Group being a between-group variable and Condition a within-group variable. Greenhouse-Geisser corrections were adopted when appropriate. Bonferroni corrections were applied for any follow-up comparisons.

EEG data pre-processing.

EEG signals were first off-line re-referenced to the average of the left and right mastoids. A band-pass filter of 0.1-30 Hz was applied to the continuous signal. Data segments during experiment breaks, which contained large movement artefacts, were manually identified and rejected.

The data were then cleaned from artefacts by using the SAR-ICA procedure (Porcaro et al., 2015). We adopted this procedure because it has been shown to be the most effective method to attenuate speech motor artefacts to date (Porcaro et al., 2015). The procedure decomposes the EEG signal into independent components on the basis of statistical properties of the signal. Using information from averaged ERPs, power spectrums, topographical distributions, correlations with lip EMG and localization, each component is classified into different clusters: articulatory speech artefacts, environmental and channel noise, ocular artefacts, and cleaned data (i.e. components that do not fall into any of the other categories).

Finally, the signal is reconstructed from the cleaned data cluster by retro-projecting the selected independent components (see Porcaro et al., 2015, for technical details).

For the EEG analysis, the cleaned EEG was segmented into epochs of 700 ms with 200 ms before picture onset. Each segment was baseline corrected with a baseline period of 200 ms prior to picture onset. ERPs were obtained by averaging segments for each condition for each participant. The mean number of segments used for computing the ERP was 105.6 for the homogeneous and 108.6 for the heterogeneous condition for monolinguals; 105.2 for the homogeneous and 107.3 for the heterogeneous condition for bilinguals.

Statistical analyses of ERPs.

Both the timing and the distribution of the stimulus-locked semantic context effects have varied in previous studies, even though the effects tended to be broadly distributed and long-lasting and occurred between 150 and 450 ms. We therefore took a rather explorative approach for all ERP analyses, carefully controlling for multiple comparisons. We chose the cluster-based permutation method (Maris & Oostenveld, 2007), because the cluster-based permutation analysis has been shown to be remarkably good at capturing broad effects (Groppe, Urbach, & Kutas, 2011), and as a permutation test, this method is non-parametric and therefore does not rely on any assumptions of the distribution of the data.

We briefly explain how the cluster-based permutation test works. In this test, a *t*-test is first carried out at a single time point and a single site. A significant point is classified into a cluster; a neighboring point that is temporarily or spatially adjacent to it would enter the same cluster. After all clusters are identified, the mass of a cluster is calculated by adding all the *t*-values in that cluster. The maximum mass with maximum *t*-values is then identified. Random permutation is carried out many times, each time identifying one maximum mass. Therefore a null distribution of the cluster mass is derived. The relative position of the experiment cluster *t*-mass in the null distribution would have a corresponding *p*-value. If the *p*-value is below alpha, the identified cluster is considered to show a significant condition difference.

The cluster-based permutation tests were implemented using the Mass Univariate ERP Toolbox (2015). An electrode's spatial neighbourhood was defined as all electrodes within approximately 3.56 cm, resulting in 7 neighbours for each electrode on average (number of neighbours ranging between 2 and 10). *P*-value threshold for cluster inclusion was .05 and it was not required to have a minimum number of below-threshold neighbours to be included in a cluster. Each test was carried out with 2500 permutations.

We performed the cluster-based permutation tests to statistically test the semantic blocking effect in monolingual and bilingual speakers separately. Two cluster-based permutation tests were carried out for each group, so that the first one covered the time window for the early effect, 150 – 350 ms, whereas the second covered the time window for the late effect 350 – 500 ms.

Results

Behavioural Results

Accuracy.

Speech accuracy in the semantic blocking experiment was checked via audio recordings. Hesitations, repairs, stutters and incorrect object names (e.g. using 'jacket' instead of 'coat') were counted as errors. For illustration purposes, Figure 5.1 shows the average percentage of correct responses for each participant group per condition (instead of arcsine transformed accuracy used in the statistical analysis). We found a significant main effect of Condition, with responses in the heterogeneous condition being more accurate, F (1, 38) = 11.46, p = .002, η_p^2 = .23. There was no main effect of Group, F (1, 38) = 0.72, $p > 0.5$,

η_p^2 = .016, nor an interaction between group and condition, F (1, 38) = 0.13, p > .05, η_p^2 = .004.

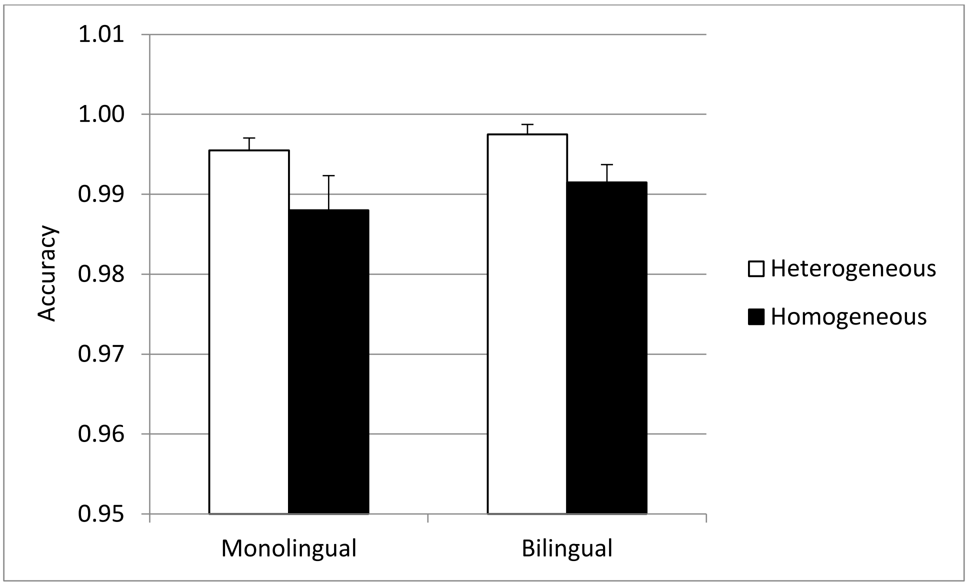


Figure 5.1. Mean response accuracies for monolingual speakers and bilingual speakers in the heterogeneous condition and homogeneous condition. Error bars represent 95% confidence interval.

Mean response time analyses.

Mean response times were calculated for each condition after excluding responses that were faster than 200 ms and responses that were 2 *SD* beyond the mean of each participant. Figure 5.2 illustrates the mean RT for the two participant groups in each condition. Mixed design *ANOVA* showed a significant main effect of Condition, F (1, 38) = 46.56, p <.001, η_p^2 = .55, with responses being slower in the homogeneous condition. There was no main effect of Group, F (1, 38) = 1.73, p > .05, η_p^2 = .04, nor a significant two-way interaction between Condition and Group, F (1, 38) = 0.54, p > .05, η_p^2 = .014, suggesting that the two participant groups performed the task very similarly.

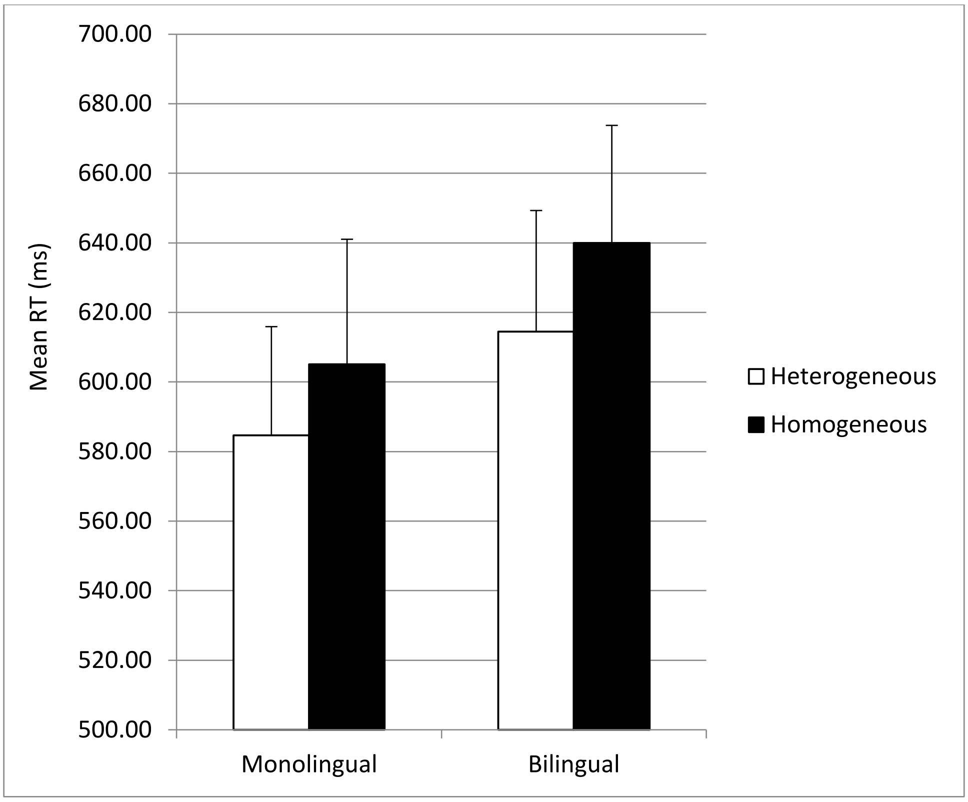


Figure 5.2. Mean RTs for monolingual and bilingual speakers for heterogeneous and homogeneous condition. Error bars represent 95% confidence interval.

Ex-Gaussian analyses.

We analysed the distribution of RTs by fitting an ex-Gaussian distribution to response times on correct trials that were above 200 ms for each condition and each participant. We estimated the ex-Gaussian distribution parameters μ and τ using the QMPE software, which uses the quantile maximum likelihood estimation method (Brown & Heathcote, 2003). Parameters were estimated for each participant under each condition using 10 bins. All ex-Gaussian parameters were successfully yielded with an average iteration of 12.9. Parameter estimations were all trustworthy since the exit codes were all below 128 (see technical manual).

Decile analyses.

Decile analyses (see Chapter 3 page 54 and Chapter 4 page 84) were carried out to make sure that the parameter estimations were reliable. Ten Decile bins were estimated because the ex-Gaussian analyses were based on ten bins. Table 5.2 summarizes the average empirical and theoretical Decile bin values for each condition and each group. 2 (Group) by 2 (Estimation Method) mixed design *ANOVAs* were conducted, for reach bin, to test the effect of fitting method, with Group being a between group factor, and Estimation Method being a within group factor. Results were also presented in Table 5.2. For many bins, there was a significant main effect of Method, suggesting differences between an empirical and a theoretical bin value. However, such discrepancies were mostly small in values and all of them were within 1 *SE* of the empirical values, except for bin 10 for bilinguals in the homogeneous condition. Most importantly, none of the *ANOVA*s showed a significant interaction between Group and Estimation Method. Absence of the interaction effects suggests that although sometimes theoretical values predicted faster or slower responses, this was the same for both groups, and to the same degree, ruling out the possibility that any group differences in the parameters were artificially caused by the ex-Gaussian model fitting.

Table 5.2

Empirical and theoretical Decile bin values for each condition, each participant group and each bin. F-statistics and p-values for the Main effect of Estimation Method (Method), Group and interaction between Group and Method.

		Heterogeneous condition									
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin6	Bin7	Bin8	Bin9	Bin10
Monolingual	Empirical	429.37	497.64	526.75	550.09	571.35	592.54	615.65	642.83	688.52	812.79
	Theoretical	443.20	496.84	524.49	547.50	569.18	591.39	616.01	646.06	688.90	796.01
	Discrepancy	13.82	-0.80	-2.26	-2.59	-2.17	-1.16	0.36	3.22	0.37	-16.78
	SE	17.10	18.04	17.31	16.33	16.62	17.13	17.73	18.78	21.29	29.41
Bilingual	Empirical	448.61	521.69	556.71	580.41	600.68	622.34	650.62	679.27	730.52	873.90
	Theoretical	462.66	520.24	550.03	575.00	598.71	623.19	650.46	683.78	731.26	847.80
	Discrepancy	14.05	-1.46	-6.68	-5.41	-1.97	0.85	-0.16	4.52	0.74	-26.10
	SE	18.67	15.69	15.18	16.71	17.13	17.57	19.72	21.57	25.36	36.84
Method	F	14.87	3.10	15.65	15.38	6.43	0.03	0.01	6.04	0.20	6.45
	p	<.001**	0.09	<.001**	<.001**	0.015*	0.87	0.94	0.019*	0.66	0.015*
Group	F	0.60	1.00	1.43	1.52	1.51	1.56	1.75	1.72	1.68	1.52
	p	0.45	0.33	0.24	0.23	0.23	0.22	0.19	0.20	0.20	0.23
Group*Method	F	0.00	0.26	3.81	1.92	0.02	1.06	0.04	0.17	0.02	0.30
	p	0.98	0.61	0.06	0.17	0.90	0.31	0.85	0.68	0.88	0.58

		Homogeneous condition									
		Bin1	Bin2	Bin3	Bin4	Bin5	Bin6	Bin7	Bin8	Bin9	Bin10
Monolingual	Empirical	437.87	509.65	546.97	571.32	595.50	621.02	645.84	676.10	732.67	896.40
	Theoretical	450.24	510.88	542.31	568.65	593.62	619.39	648.12	683.30	733.59	858.68
	Discrepancy	12.37	1.23	-4.65	-2.67	-1.87	-1.64	2.27	7.21	0.92	-37.71
	SE	15.04	16.30	16.59	16.90	17.49	17.55	17.33	17.60	22.19	37.92
Bilingual	Empirical	455.79	539.96	575.10	602.03	629.00	657.09	684.12	720.81	770.66	941.29
	Theoretical	472.42	538.28	572.18	600.40	626.98	654.13	684.08	720.31	771.39	896.80
	Discrepancy	16.63	-1.69	-2.92	-1.62	-2.02	-2.96	-0.04	-0.50	0.72	-44.49
	SE	17.82	16.98	16.48	16.18	16.56	17.63	18.39	18.61	19.59	27.73
Method	F	13.85	0.07	9.56	4.10	3.76	8.04	1.15	5.50	0.26	16.30
	p	0.001**	0.80	0.004**	0.05*	0.06	0.007**	0.29	0.024*	0.61	<.001**
Group	F	0.71	1.50	1.54	1.79	1.97	2.09	2.20	2.49	1.66	0.92
	p	0.41	0.23	0.22	0.19	0.17	0.16	0.15	0.12	0.21	0.34
Group*Method	F	0.30	2.66	0.50	0.24	0.01	0.67	1.23	7.26	0.00	0.11
	p	0.59	0.11	0.48	0.63	0.94	0.42	0.27	0.01	0.95	0.74

Note. Discrepancy = Theoretical - Empirical Decile means. *SE* = standard error of empirical response times. * signifies that the effect was significant at .05 significance level.

** signifies that the effect was significant at .01 significance level.

The μ parameter.

Figure 5.3 shows the mean μ parameter estimation of the two participant groups for the two conditions. We found a significant main effect of Condition, F (1, 38) = 5.79, p = .02, η_p^2 = .132, with responses in the homogeneous condition being slower. The main effect of Group was not significant, F (1, 38) = 2.03, p > .05, η_p^2 = .051, neither was the interaction between Group and Condition, F (1, 38) = .08, p > .05, η_p^2 = .002. Based on the results in Chapter 4, planned paired-sample t-tests were conducted to investigate how the semantic context affects the distribution parameters in each group. Monolingual speaker showed a marginal semantic context effect in the μ parameter t (19) = 1.98, p = .06, while bilingual speakers did not show a significant difference in the μ parameter across conditions, t (19) = 1.59, p > .05.

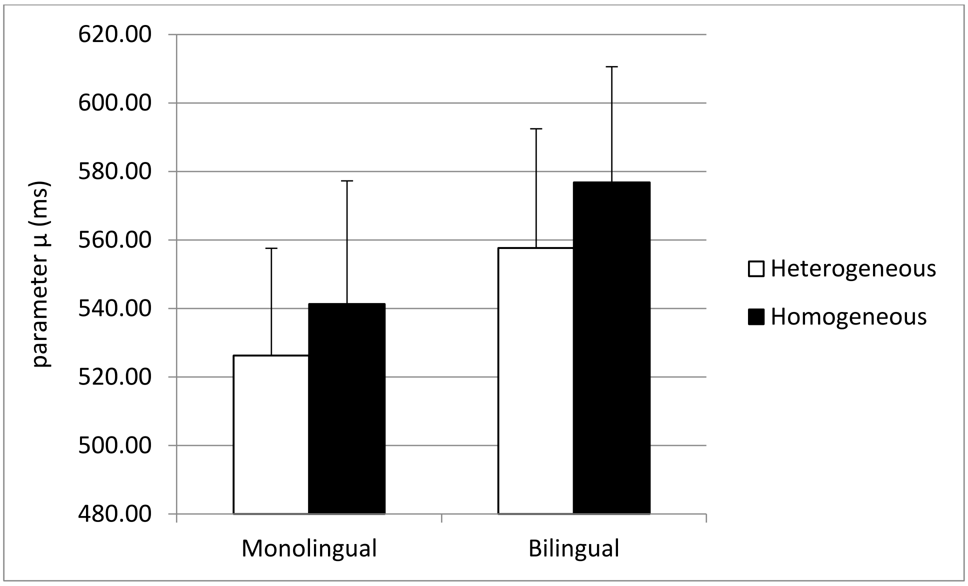


Figure 5.3. Means of the estimated μ parameters for monolingual and bilingual speakers for heterogeneous and homogeneous condition. Error bars represent 95% confidence interval.

The τ parameter.

Figure 5.4 shows the mean τ parameter estimation of the two groups under the two conditions. We did not find any significant main effect of Condition, F (1, 38) = 2.81, p > .05, η_p^2 = .069, or Group F (1, 38) = 0.004, p > .05, η_p^2 = .000. The interaction between

Group and Condition was not significant either F (1, 38) = .06, p > .05, η_p^2 = .002. Planned comparisons revealed that there was no significant difference between the τ parameter across conditions in either the monolingual group, t (19) = 1.42, p > .05 or the bilingual group, t (19) =.97, p > .05.

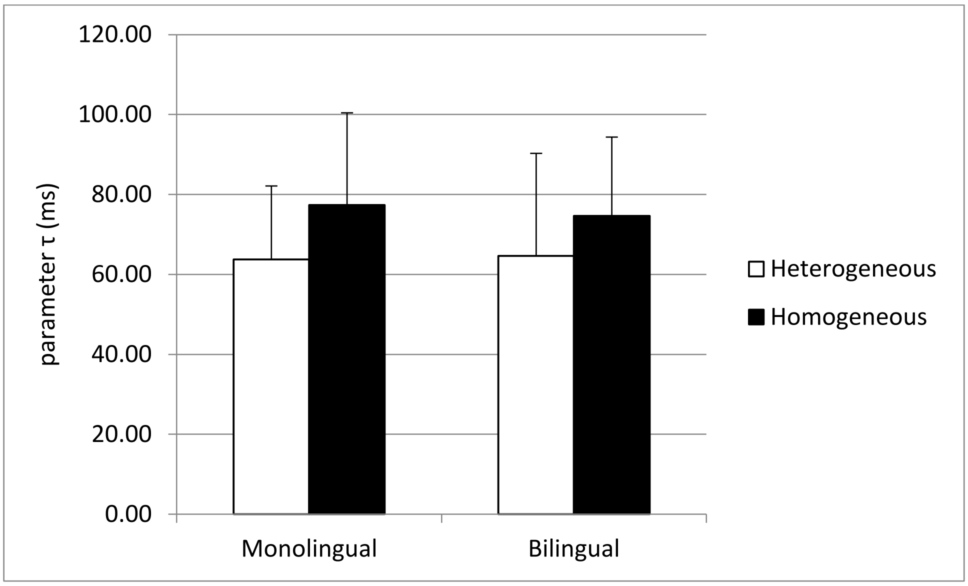


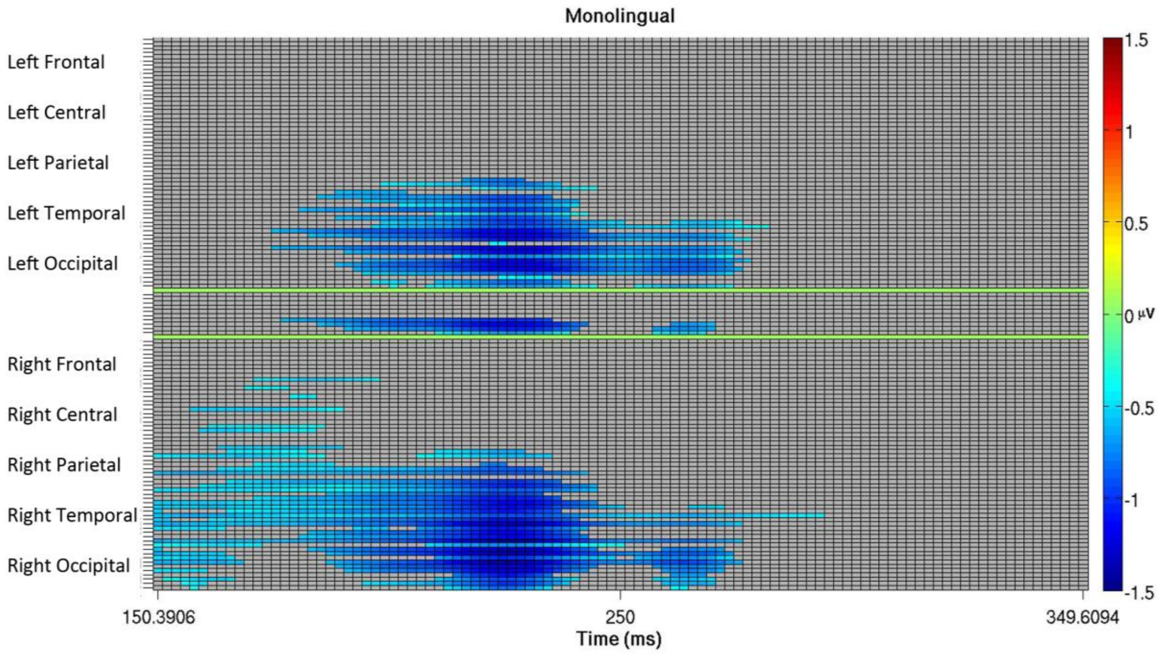
Figure 5.4. Means of the estimated τ parameters for monolingual and bilingual speakers in the heterogeneous and homogeneous condition. Error bars represent 95% confidence interval.

Event Related Potentials

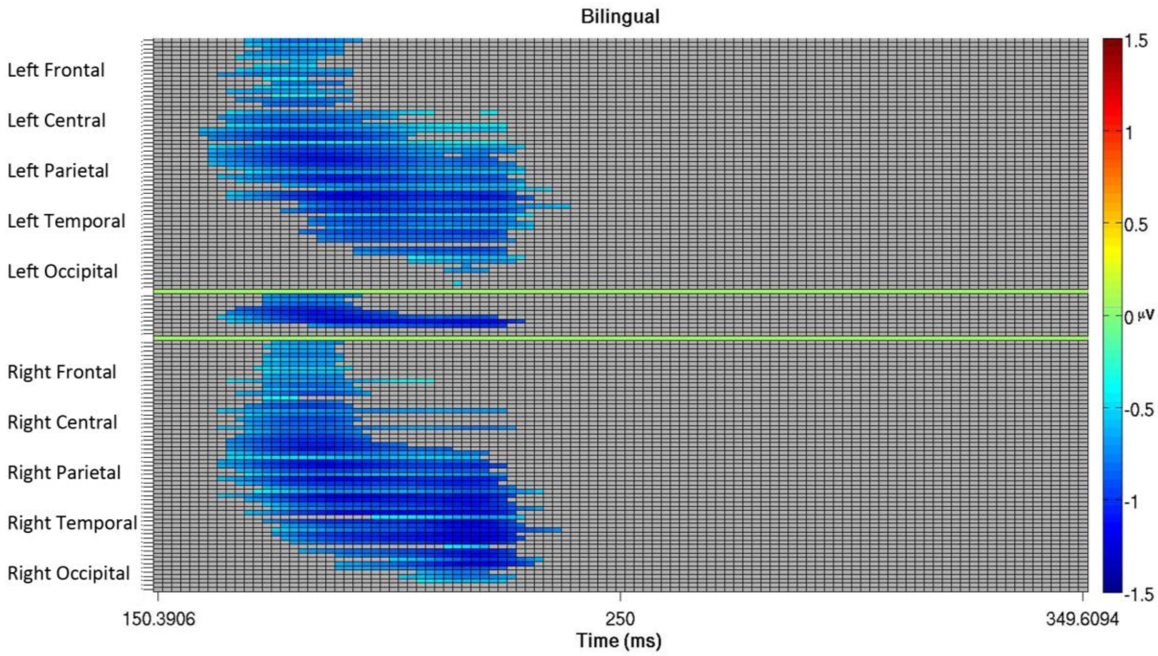
The results for the cluster-based permutation tests for the time window 150 - 350 ms are visualized as a raster diagram in Figure 5.5. For monolingual speakers (Panel A), the procedure returned one significant cluster. The cluster started at 150 ms with a diffuse right distribution and ended at approximately 280 ms, with a parietal-posterior distribution. The effect was most pronounced between 180 ms and 250 ms, with the homogeneous condition leading to reduced positive activation at parietal-posterior sites compared with the heterogeneous condition.

For bilingual speakers (Panel B), again one significant cluster was identified. The cluster started at 160 ms at central and frontal electrodes and shifted backwards to end at 230 ms with a parietal-posterior distribution. Just like monolinguals, bilinguals showed a reduced positivity in the homogeneous condition compared with the heterogeneous condition.

Through visual inspection, we identified time points where these effects were strongest and most widely distributed. For both monolingual and bilingual speakers, the posterior effect peaked around 220 ms. The additional central-frontal effect for bilinguals peaked around 180 ms. Figures 5.6 and 5.7 show the topographic distributions of the effects when they reached their peaks as well as ERPs and difference waves at representative electrodes.



Panel A



Panel B

Figure 5.5. Raster diagram for the cluster-based permutation test results for monolinguals (Panel A) and bilinguals (Panel B) between 150 – 350 post stimuli onset. Coloured square represent electrodes that contributed to the significant cluster at a given time step. Differences between the homogeneous and heterogeneous condition are shown.

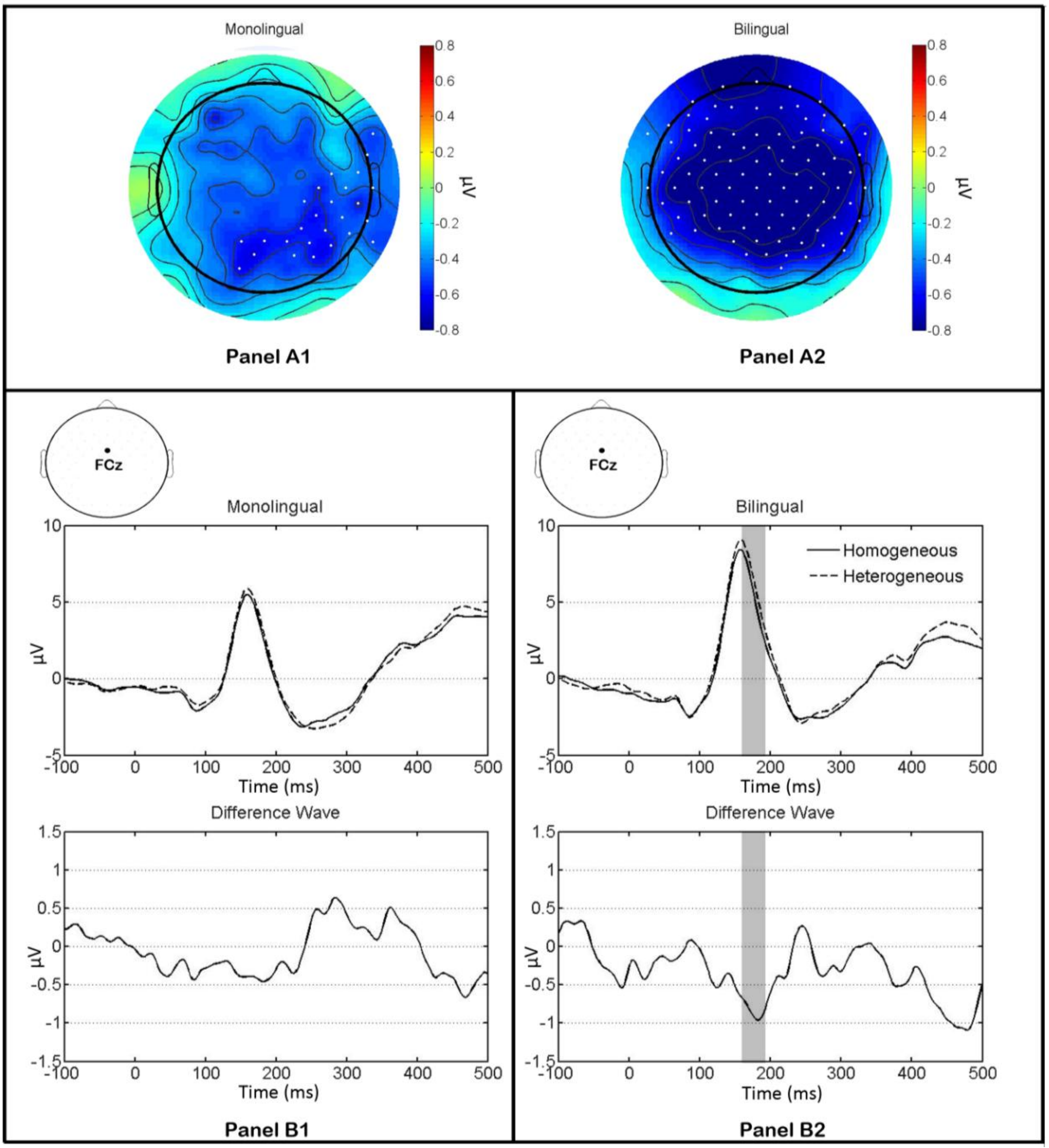


Figure 5.6. Topographic distribution of semantic blocking effect for monolinguals (A1) and bilinguals (A2) at 180 ms. Electrodes that contributed to the significant cluster are marked as white dot. Panel B1 and B2 show the grand average ERP waves (upper panels) and difference waves (lower panels) between homogeneous condition and heterogeneous condition at a frontal-central electrode (FCz) for monolinguals (B1) and bilinguals (B2). Grey shades mark the time windows during which the electrode contributed to the significant cluster.

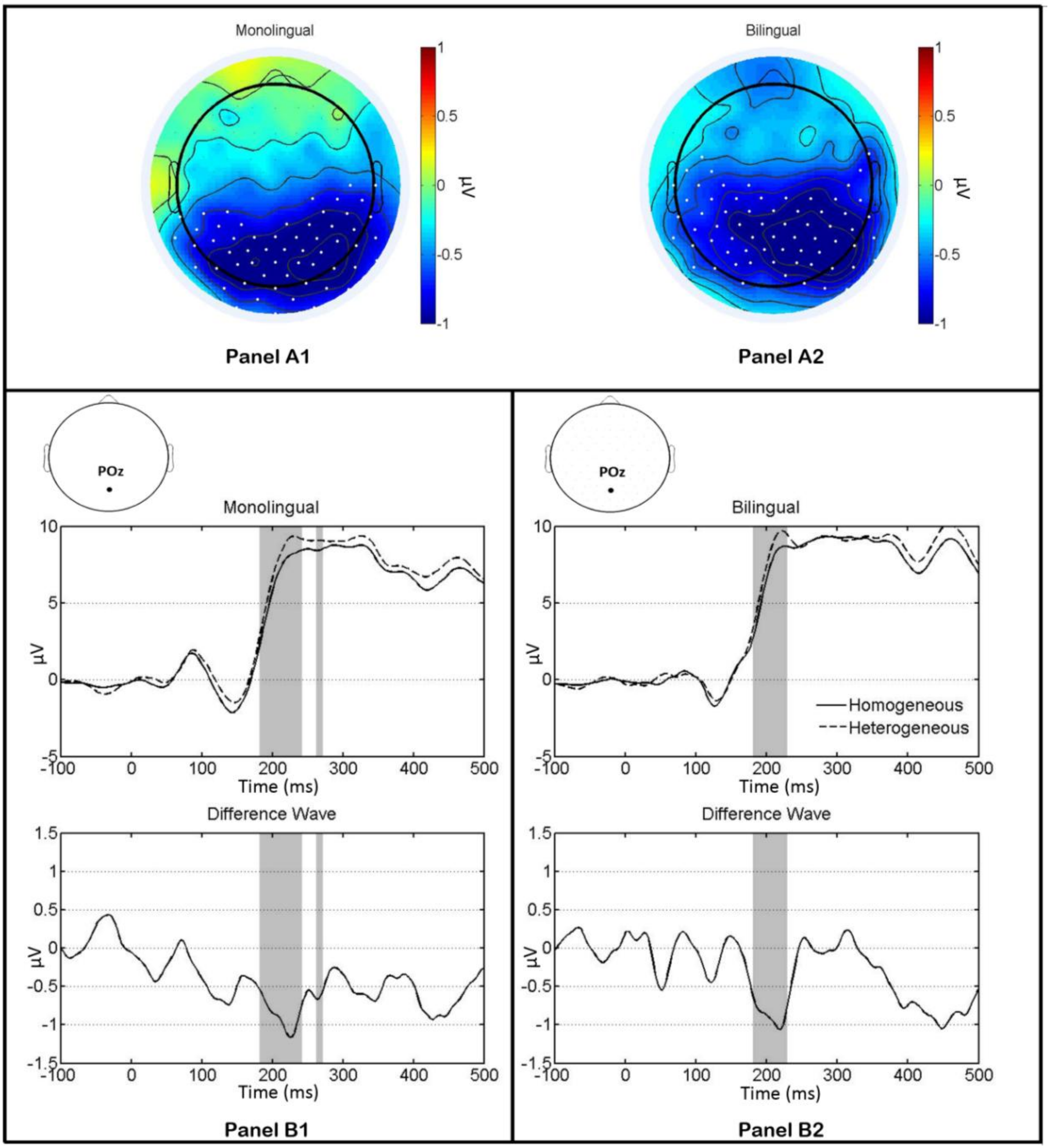
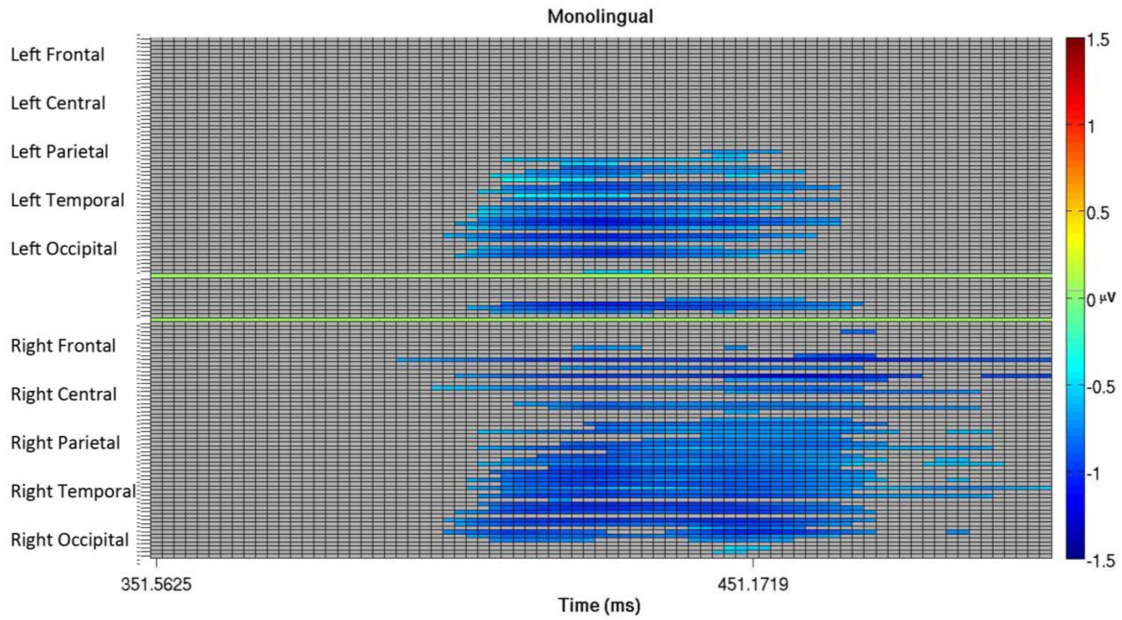
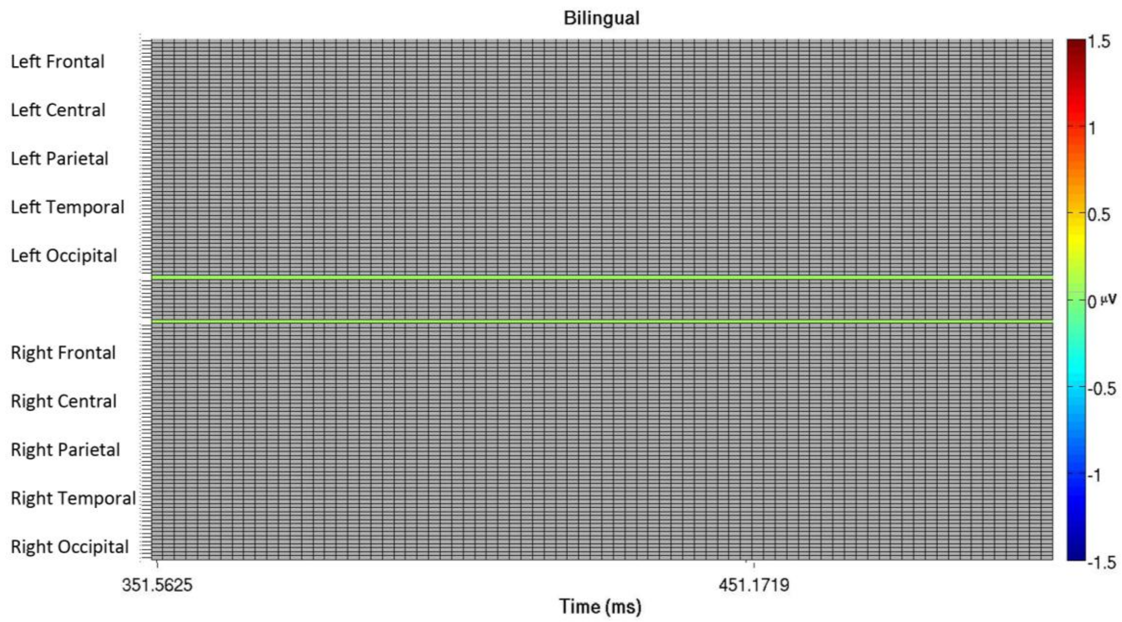


Figure 5.7. Topographic distribution of semantic blocking effect for monolinguals (A1) and bilinguals (A2) at 220 ms. Electrodes that contributed to the significant cluster are marked as white dot. Panel B1 and B2 show the grand average ERP waves (upper panels) and difference waves (lower panels) between homogeneous condition and heterogeneous condition at a representative electrode (POz) for monolinguals (B1) and bilinguals (B2). Grey shades mark the time windows during which the representative electrode contributed to the significant cluster.

The results for the time window between 350 ms and 500 ms are visualized with a raster diagram in Figure 5.8. For monolinguals, one significant cluster was identified starting at parietal-posterior electrodes at 400 ms and lasting until 470 ms (Panel A). The homogeneous condition led to more negative activity than the heterogeneous condition. See Figure 5.9 Panel B1 and B2 for the ERPs and difference wave at a representative electrode. Similar to the early parietal-posterior effect, the topographic distribution plot shows that this effect was present at parietal-posterior sites. Even though descriptively, a similar effect appeared for the bilingual speakers (Figure 5.9 Panel B2), no significant cluster was identified (Figure 5.8 Panel B).



Panel A



Panel B

Figure 5.8. Raster diagram for the cluster-based permutation test results for monolinguals (Panel A) and bilinguals (Panel B) in the stimulus locked analysis between 350 – 500 post stimuli onset. A coloured square represents an electrode that contributes to the significant cluster at a certain time step.

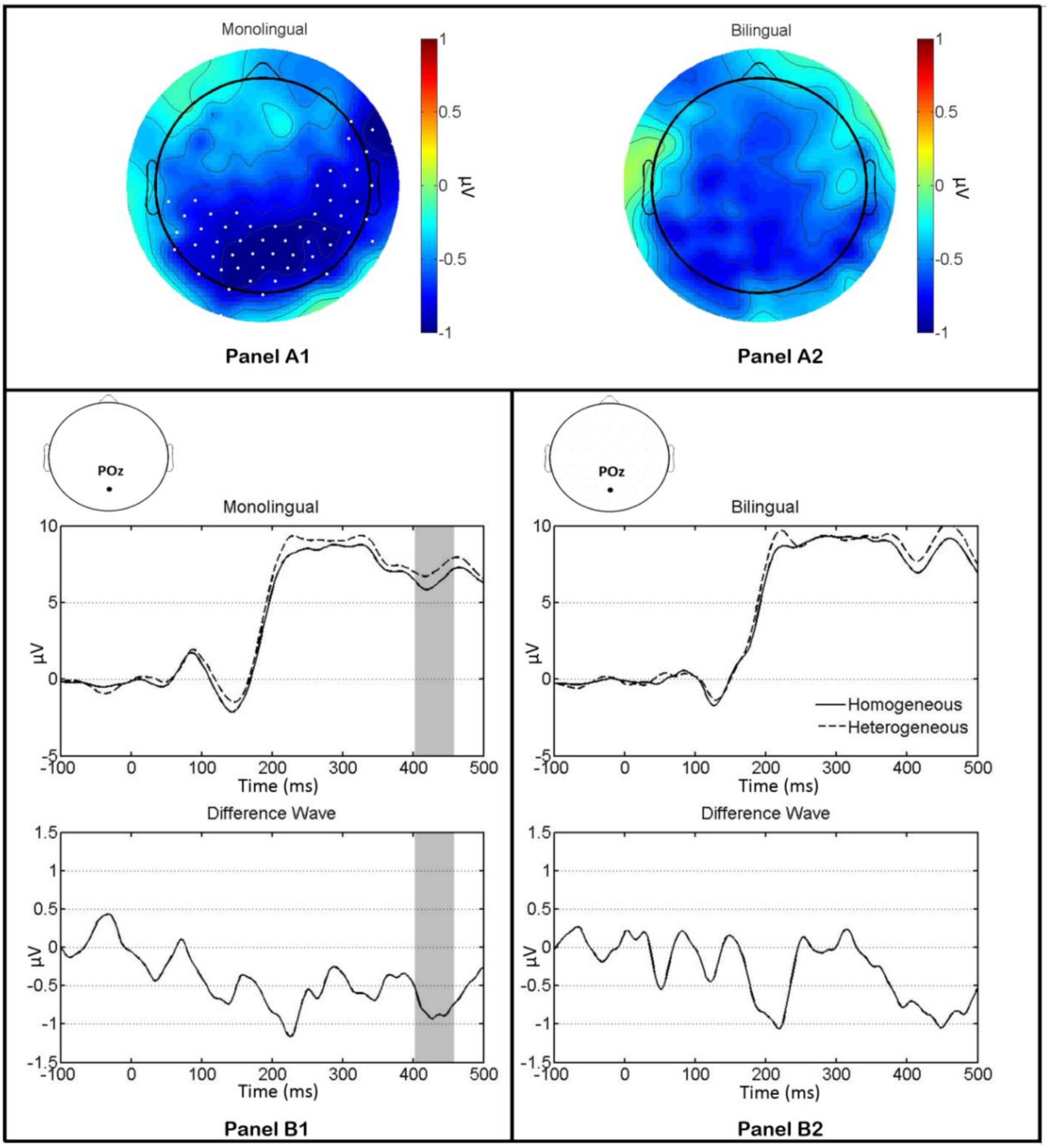


Figure 5.9. Topographic distribution of semantic blocking effect for monolinguals (A1) and bilinguals (A2) at 425 ms. Electrodes that contributed to the significant cluster are marked as white dot. Panel B1 and B2 show the grand average ERP waves (upper panels) and difference waves (lower panels) between homogeneous condition and heterogeneous condition at a representative electrode (POz) for monolinguals (B1) and bilinguals (B2). Grey shades mark the time windows during which the representative electrode contributed to the significant cluster.

Correction Analyses

Point-by-point correlation analyses between behavioural results and ERP results were conducted to investigate whether ERP effects were related to behavioural outcomes. We therefore focussed on the three time windows where a significant ERP effect was found: central-frontal effect between 160 – 200 ms, parietal-posterior effect between 200 – 280 ms and parietal-posterior effect between 400 – 470 ms. Nine representative electrodes were selected from the central-frontal cluster (Fz, FCz, Cz, FFC1h, FFC2h, FCC1h, FCC2h, FC1 and FC2) and the parietal-posterior cluster (Pz, POz, Oz, PPO1h, PPO2h, PO3h, PO4h, POO1, POO2). ERP activities were averaged over these clusters, and correlations between the RT effect and the averaged ERP effect were calculated for every 2 ms. A correlation effect was considered significant when at least 15 consecutive points showed significant correlations, which corresponded to 30 ms.

No significant correlations were found between the RT effect and the ERP effect for 160 – 200 ms at the central-frontal site, neither for monolinguals, nor for bilinguals. Figure 5.10 shows the results for the correlation analyses for 200 – 280 ms at the parietal-posterior site. RT effect correlated significantly with the ERP effect between 200 ms and 270 ms for monolingual speakers, while no significant correlation was found for bilingual speakers. Results for the correlation analyses for 400 – 470 ms at parietal-posterior site also revealed no significant correlations, neither for monolinguals nor for bilinguals.

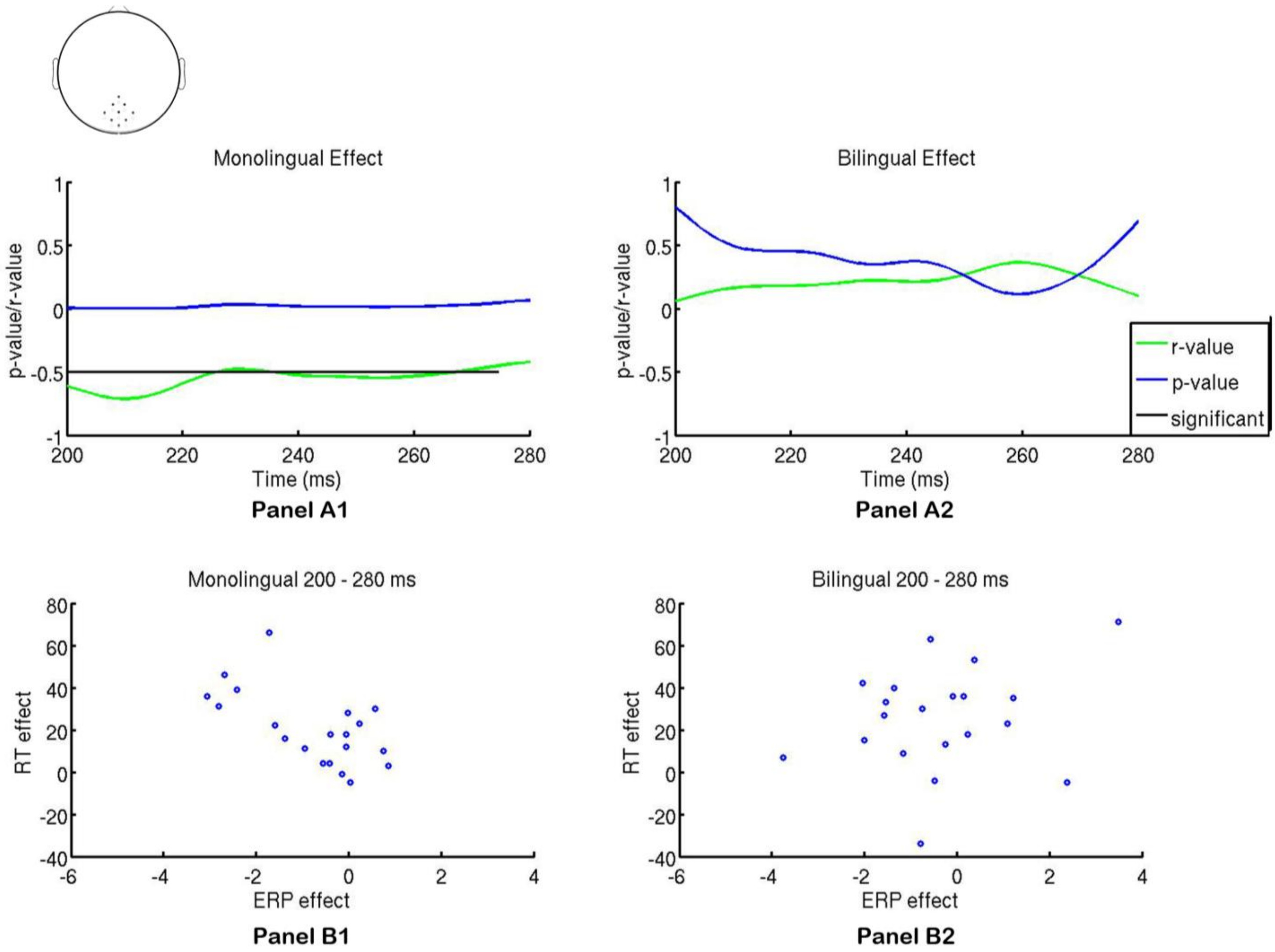


Figure 5.10. Correlation coefficients (green line) and p-values (blue line) for correlations between the semantic blocking ERP effect and the RT effect for monolinguals (Panel A1) and bilinguals (Panel A2). Black bars mark time points at which the correlation was significant at significance level of .05. Scatter plots for the average ERP effect (Homogeneous ERP–Heterogeneous ERP) between 200 – 280 ms against the behavioural RT effect (Homogeneous RT – Heterogeneous RT) for monolinguals (Panel B1) and for bilinguals (Panel B2). Representative electrodes: Pz, POz, Oz, PPO1h, PPO2h, PO3h, PO4h, POO1, POO2.

Discussion

The aim of the present study was to gather electro-physiological evidence for the enhanced executive control ability in bilinguals when resolving lexical conflict within one language. We recorded participants' EEG while they performed a cyclic semantic blocking task. We will first discuss behavioural results and then EEG results.

Behavioural Results

In terms of behavioural results, we found some evidence that monolingual and bilinguals speakers resolve intra-lexical semantic competition differently, partially replicating our previous results (Chapter 4). As for the pervious experiment, accuracy analyses showed significant higher accuracy in the heterogeneous condition than in the homogeneous condition. However, in the current study, we did not find any evidence for a reduced blocking effect in accuracy for bilinguals. An analysis of mean response times revealed the predicted blocking effect but no indication of differences between the two groups. Ex-Gaussian analyses of response time distributions revealed, however, a small difference between the groups. Monolingual speakers showed a marginal semantic blocking effect in the body of the distribution, but not the tail, whereas bilingual speakers did not show any effect. These results are partially consistent with our previous findings (Chapter 4). More precisely, they are consistent for monolinguals, who had shown the blocking effect in the main part of the response distributions but not in the tail. Results for bilinguals are less consistent. In neither study did bilinguals show a semantic blocking effect in the body of the distributions. However, while a significant effect was present in the tails in the previous study for bilinguals, such an effect was absent in the current one. That is to say, bilingual speakers did not show a significant effect, neither in the main body nor the tail of the response distributions. This does not mean, however, that they did not suffer any semantic context effect, as traditional analysis of mean response times showed that they did suffer. It rather means that some participants showed the effect in the body, while others in the tail of the distribution. This suggests that the behavioural response patterns of the bilingual group are less homogeneous than those of the monolingual group in the present experiment and less homogeneous than those of the bilingual group in the previous study.

We have previously interpreted the effect in the main part of the distribution as an indication of more effortful lexical selection, because responses in the homogenous condition are slowed down due to higher activation levels of alternative lexical items. And we have interpreted the effect in the tail of the distribution as capturing recoveries from covert errors. Therefore, results in the current study suggest that the blocking effect of monolinguals was rather caused by lexical competition due to high activation levels of alternative lexical items, while the blocking effect of bilinguals was a mixture of effects of lexical competition and recoveries from covert errors.

The question arises why we did not fully replicate the result pattern reported in the previous behavioural experiment. We speculate that two factors might have led to these discrepancies. First, the present study was carried out in an experimental setting for EEG. Participants were asked to sit still and monitor their eye blinks while performing the task. This must have posed extra demand for control. As we have proposed, the behavioural differences might be rooted in different levels of engagement of executive control. If the experimental setting takes up more cognitive resources of bilingual speakers who are already under higher control demands, some of them might have been left with reduced resources available for executive control to perform the task. Therefore, they might have shown a result pattern rather similar to that of the monolingual group. Furthermore, a comparison of the monolingual data across the two studies shows a higher variance in RTs in the present study, especially for the body of the response distributions. This suggests a more heterogeneous group of monolingual participants as well, and it might explain the weaker blocking effect in the current study.

Second, the bilingual population differed in the two studies. The English/Chinese bilingual group of the previous experiment was more proficient in their L2 and had an earlier onset of L2 than the group of the present study. In addition, instead of being a homogeneous bilingual group as in the previous experiment, bilingual speakers spoke various other languages in the present study. It has been argued that the similarity of the languages of a bilingual can affect their executive control abilities (Coderre & van Heuven, 2014). All these differences mean that some bilinguals might have performed more like monolinguals, showing the blocking effect in the main body, while some bilinguals might have performed more like bilinguals in the previous study, showing the effect in the tail of the response distributions. This might have eliminated any group differences.

EEG Results

Despite weak evidence for a difference between bilingual and monolingual participants in the behavioural data, the EEG data showed striking differences between the groups. The EEG analyses revealed four main findings. First, both monolingual and bilingual speakers showed a parietal-posterior semantic blocking effect with a peak at 220 ms after stimulus onset. The effect was similar for the two participant groups, but it lasted longer for monolingual speakers. Second, only for bilingual speakers, this effect co-occurred with a central-frontal effect that started earlier and peaked at 180 ms. Third, only monolingual speakers showed a significant parietal-posterior effect between 400 ms and 450 ms after stimulus onset, with the homogeneous condition being less positive. Fourth, we found a significant correlation between RT blocking effects and the average ERP effects (difference between homogeneous and heterogeneous ERP) between 200-270 ms in monolinguals but not in bilingual speakers. RT effects were not correlated with the early central-frontal ERP effects or the late parietal-posterior effects. We next discuss each result and its implication.

Early semantic context effect at parietal-posterior region.

We found an early semantic context effect at parietal-posterior regions in both groups. This effect fell roughly between 150 – 280 ms with a peak at 220 ms, which is largely consistent with other studies that have investigated the electrophysiological markers for the semantic blocking effect (Maess et al., 2002; Janssen et al., 2011; Aristei et al., 2011; Costa et al., 2009; the latter for the cumulative semantic interference paradigm). This timing also corresponds to the proposed timing of lexical selection, which has been estimated to occur roughly from 200 to 275 ms (Indefrey, 2011; Indefrey & Levelt, 2004). Following the interpretation of previous studies, this effect is therefore in line with the hypothesis that lexical selection processes are complicated in the homogeneous condition in comparison to the heterogeneous condition due to highly activated lexical competitors.

Despite showing the same parietal-posterior effects in terms of the nature of the effect and its peak, the two participant groups differed in terms of the onset of the effect and the duration of the effect. The effect of monolinguals initiated slightly earlier, around 150 ms. If the effect reflects a lexical selection process, the earlier onset would mean that monolinguals initiated this process earlier. This might have been reflected in the behavioural outcomes. Indeed bilingual speakers were slower than monolingual speakers for about 30 ms, at least descriptively. This might be due to English being the second language for most bilingual speakers. We next consider the duration of the effects. Assuming the effect reflects lexical selection, the duration of the effect then reflects how long it takes to complete the lexical selection process. In case of the semantic blocking effect, this effect reflects extra processing cost posed on participants to resolve the increased lexical selection demand in the homogeneous condition. For monolinguals the effect lasted up to 280 ms, while for bilinguals it lasted only up to 230 ms. Therefore, it appears that it takes longer for monolinguals to resolve the lexical competition. This result is consistent with the idea that bilingual speakers have better executive control, as a result of which bilinguals resolve the lexical conflicts faster. Given that bilingual speakers have to constantly deal with the dual activation of two languages on a daily basis, it is not surprising that their lexical conflict resolution mechanism is well-tuned and more controlled.

Early semantic context effect at central-frontal regions.

Bilingual speakers showed a central-frontal effect between 160 – 200 ms post picture onset. This effect was distinct from the parietal-posterior effect in terms of its topographic distribution, suggesting that it reflected processes other than lexical selection. However, this effect occurred around the time when lexical selection takes place, which means it might be related to lexical selection.

Could this effect be an anterior N2, which is a marker for response inhibition (for a review of the N2, see Folstein & Van Petten, 2008)? An anterior N2 has been observed for speech production tasks. For example when ERP responses were compared for pictures with varying degrees of name agreement, pictures with lower name agreement showed a more pronounced N2 (Cheng, Schafer, & Akyurek, 2010; Shao, Roelofs, Acheson, & Meyer, 2014). The authors interpreted the N2 effect as being caused by the recruitment of selective inhibition to suppress competitors for items with low name agreement (Shao et al., 2014). However, anterior N2 typically has a later peak, for example around 250 ms in simple picture naming tasks (Shao et al., 2014) or around 310 ms in bilingual switch naming paradigms (Jackson, Swainson, Cunnington, & Jackson, 2001; Verhoef, Roelofs, & Chwilla, 2009). Therefore the central-frontal effect in the present study, which was observed around 160 – 200 ms, seems to be too early to be an N2.

We nevertheless suggest that this central-frontal effect reflects executive control processes. Similar to the arguments put forward by Aristei et al. (2010) who found semantic context effects at frontal regions, our effect might reflect increased activation of a neural substrate of executive control in the homogeneous condition. To be more specific, it might reflect some kind of conflict detection and/or control. The effect was only observed in the bilingual group, which suggests that compared to monolingual speakers, bilingual speakers might have more strongly exerted this executive control to resolve conflicts.

As bilingual picture naming is by nature different from monolingual picture naming in that the non-target language is always active, even if only to a small degree, this early frontal-central effect might, alternatively, be caused by the additional task of managing competing lexical activation in the non-target language. This effect cannot reflect the suppression of the other language as a whole, because that should be the same for the two experimental conditions. However, it might reflect the control over the activated competitors in the other language. If competitors in the target language activated their translation equivalents in the non-target language, then the homogeneous condition would mean more competition than the heterogeneous condition because competitors and therefore their translation equivalents would be more strongly activated. What speaks against this interpretation is that the competition seems to be resolved more efficiently in bilingual than monolingual speakers, as reflected in their reduced early semantic context effect as well as non-significant monitoring effect. Also, the central-frontal effect is a reduced positivity in the homogeneous condition, meaning that increased control demands led to *reduced* positivity. However, bilinguals showed generally increased, not reduced positive activity in central-frontal sites during this time window (see Figure 5.6), which does not fit the idea that bilinguals have generally increased control demands. This leaves us with the conclusion that the central-frontal activity in bilinguals most likely reflects recruitment of executive control. More specifically, the control system of bilinguals might have detected the increased conflict in the homogeneous block and facilitated the resolution of such conflicts.

Late effect in the posterior region.

The present data revealed a further effect in monolingual speakers during 400 – 450 ms in parietal-posterior regions. This effect is largely similar to the effect between 450- 475 ms in Maess et al.'s (2002) MEG study, especially when considering that our average responses are faster than those in the MEG study (by about 100 ms). Maess et al. (2002) suggested that their effect most likely reflects internal self-monitoring processes. In the homogeneous condition, where alternatives are more active, it might be more effortful to confirm that the correct response has been chosen. Bilingual speakers showed the same effect numerically, but it was not significant, suggesting high variability across the bilingual sample. This fits with our conclusion about the response distribution results for bilingual participants.

Correlation.

Correlation analyses informed us how our effects in brain activities translated into behavioural outcomes. Correlations for the parietal-posterior site was significant for monolinguals but not for bilinguals, roughly between 200 ms and 270 ms, with participants with a stronger behavioural effect showing a bigger ERP effect. This suggests that the processes during this time-window, which we have interpreted as being related to lexical access and competition, strongly affected response times. The question arises why correlations were only found for monolingual, but not bilingual speakers. In general, if the processing of a task is fairly sequential and linear, one would expect direct mapping between the magnitude of ERP activities and response times. However, if the processing of a task involves multiple operations, it becomes more difficult to observe a direct relation between the neural activity and behavioural outcomes. Responses in the cyclic semantic blocking task are not only affected by lexical selection efficiency but also by top-down executive control (see Belke, 2008). If bilingual speakers engaged increased top-down control, especially in the homogeneous condition, as suggested by early central-frontal effect, then this could have obscured the mapping between the neural activity and response times. We found no correlations for central-frontal sites between 160 – 200 ms or for parietal-posterior sites between 400 – 470 ms, for neither of the two language groups. This suggests that the engagement of executive control processes or monitoring processes might not have contributed to RTs in our study, at least not visibly.

Behavioural results showed that the two participant groups performed quite similarly. However, their brain responses suggested that their internal processes differed considerably. The discrepancy between behavioural and neural effects is not that unusual. Other electrophysiological studies have found substantial neural differences between monolingual and bilingual speakers without behavioural differences in non-verbal conflict tasks, such as the Simon, Colour Stroop and Flanker task (Bialystok et al., 2005; Kousaie & Phillips, 2012). This suggests that the bilingual experience has changed the neural networks that are used for these tasks, but these changes do not necessarily lead do distinct behavioural responses. This depends on the exact task demands and the sample of participants tested.

Conclusion

We tested monolingual English speakers and age matched bilingual speakers in a cyclic blocked picture naming task. Our goal was to gather electro-physiological evidence for the enhanced executive control ability in bilinguals when resolving lexical conflict within a language. Although the semantic blocking effect had been investigated using EEG before, our study differed in two ways. First, it is the first study that compared the responses of monolinguals with bilinguals. Second, we used an artefact cleaning procedure to remove speech artefacts.

Our results support the view that bilingual speakers might engage executive control networks more strongly or in a different way when resolving lexical competition. As a result, it takes bilingual speakers shorter to resolve lexical competition, and the demand for self-monitoring is lower for bilinguals than for monolinguals, even though this is not reflected in their response times. In conclusion, bilingual speakers appear to have a competition resolution system that not only suppresses competitors from their non-target language very efficiently, but also suppresses competitors within the target language.

CHAPTER 6

EXECUTIVE CONTROL NETWORK IN MONOLINGUALS AND BILINGUALS: A STRUCTURAL EQUATION MODELLING APPROACH

Abstract

The study in this chapter compared the organization of the executive control network in monolingual and bilingual speakers. 100 monolingual English speakers and 100 bilingual speakers of English and another language were tested on nine executive control tasks. These tasks were chosen because they tap into each of the sub-functions of the executive control network: inhibition, shifting and updating. Confirmatory Factor Analysis (CFA) led to a two-factor model, consisting of a shifting factor and a common executive control factor (common EC). The fact that this model fits both participant groups suggests that the general executive control network of the two groups is organized similarly. Model invariance analyses suggested that the two groups differed with regard to the factor mean for the shifting factor, with bilinguals having a higher score and therefore a better shifting ability. There was no difference with regard to the common executive control factor, which combines the functions of attentional control and inhibition. Factor covariance was descriptively higher for bilinguals than for monolinguals, implying a more correlated executive control network for bilingual speakers.

Introduction

Most of our cognitive abilities or cognitive processes are complex and not directly measurable. One way to understand such complex process is to understand its components, which can be evaluated by various carefully devised tasks or paradigms. In this thesis so far we have taken this approach and investigated the bilingualism effect on particular cognitive abilities. We have focused on attentional control and inhibitory control in both non-linguistic (Chapter 3) and linguistic tasks (Chapter 4 and Chapter 5). This chapter will take our understanding of the bilingualism effect one step further by investigating its impact at a different level, namely at the level of the whole executive control network rather than an individual aspect of this network. In what follows, we first introduce the risks and limitations of an approach that only investigates the bilingualism effect on a single cognitive construct. The structural equation modelling approach, which has successfully been used to investigate the network of executive control in monolingual speakers is proposed to address this issue. We will also review studies that have investigated the impact of bilingualism in each sub-division of the executive control network before moving on to detail the current investigation and its findings.

Limitation of the Componential Approach

For the purpose of this study, executive control is defined according to the framework put forward by Miyake et al. (2000). This framework distinguishes three sub-functions of executive control: inhibition, updating and shifting, as well as a common attentional control function that actively maintains working memory representations. Therefore, in order to understand the bilingualism effects on executive control, one might take a componential analysis approach and investigate how bilingualism affects each of the sub-functions. This has been the dominant approach in this field of research. One limitation of this approach to executive control is the task impurity problem (Rabbitt, 1997). In other words, an individual task does not provide a pure measure of a single executive control ability, but draws on many aspects of executive control as well as task-specific requirements. This problem is reflected in the lack of correlations among tasks that were thought to tap a similar executive construct, and it has been suggested that this lack of correlations might be due to non-executive processing requirements that mask the similarities between tasks (Miyake et al., 2000). In other words, if non-executive processing plays an important role in performing a task, what a task is measuring then deviates from the original intention. This also means that any group differences found with a task cannot be solely attributed to the targeted cognitive construct.

In addition to the task impurity problem, the componential approach also limits the scope of investigation. Kroll and Bialystok (2013) pointed out that the effect of bilingualism might essentially entail the reorganization of the cognitive network. Efforts trying to capture the bilingualism effect on the constructional unit of executive control might not be fruitful because it is possible that bilingualism shapes the way the executive control network is organized, not necessarily the individual functional potential of each subprocess. Therefore, investigating the network as a whole can provide us with evidence of the bilingualism effect at a different level. This chapter addresses this issue by conducting a comprehensive test of different aspects of executive control and by using confirmatory factor analysis to compare the organisation of the executive control networks in monolingual and bilingual speakers.

Organisation of Executive Control in Monolingual Speakers

As introduced, Miyake et al. (2000)'s proposed a framework for executive control, which consists of three sub-functions: shifting, updating and inhibition. Shifting is the ability to shift back and forth between multiple tasks, operations or mental sets. Updating and monitoring of working memory representations is responsible for coding incoming information that is relevant to the task at hand and revising items when appropriate by replacing old, irrelevant information with new, relevant information. Inhibition is involved when one deliberately inhibits dominant, automatic or prepotent responses. Miyake et al. (2000) used nine executive control tasks to tap into the three sub-functions. They utilized confirmatory factor analysis to model the three sub-functions at the latent factor level and found that their empirical data were fitted best by a model that allowed all three latent factors to correlate with each other (see Figure 6.1). The moderate correlation between the latent factors led them to conclude that the executive control network has two important features: unity and diversity. The unity property suggests that the three sub-functions are not completely independent of each other, and they are related. According to Miyake et al. (2000), the unity aspect of executive control points to common task requirements that are involved in all the tasks used. Such a common process could be active maintenance of task goals and context information in working memory, controlled attention, or some kind of inhibition. The diversity property of the executive control network suggests that although these functions are related, this relation is not perfect, which means each sub-function has its unique feature that is not included or cannot be explained by other sub-function(s).

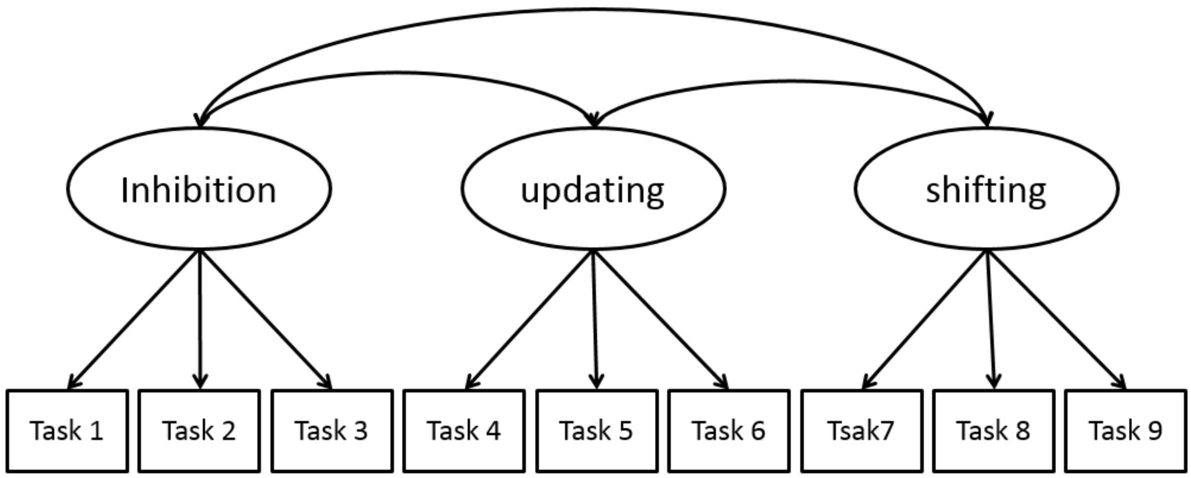


Figure 6.1. Model for executive control proposed by Miyake et al. (2000). Executive control consists of the three sub-functions inhibition, updating and shifting. Each sub-function was tapped by three tasks. The three sub-functions were allowed to correlate with each other.

Friedman et al. (2008) replicated this finding in a larger scale study. In addition, they fitted the data with an alternative hierarchical model, with a second order latent factor to capture the unity property of executive control. They found that the variance in the inhibition function could be fully explained by the common EC component. This motivated them to propose a framework that a) features a domain general function that is common across all ECs and b) captures the uniqueness of particular abilities (see Figure 6.2, see also Miyake & Friedman, 2012). In this framework, both updating and shifting abilities have unique factors, but not inhibition, because it was found all its variance could be accounted for by the common EC component. Similar to the interpretation in Miyake et al. (2000), Friedman et al. (2008) related the common EC to one's ability to actively maintain task goals and use such information to direct lower-level processing.

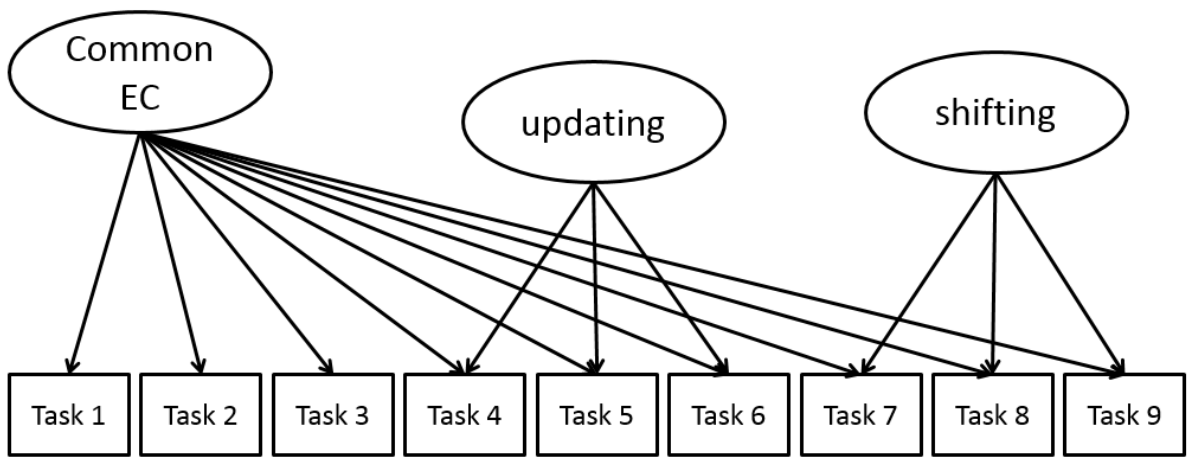


Figure 6.2. Model for executive control proposed by Friedman et al. (2008). According to the bi-factor model, executive control consists of a domain general, common EC component and

two domain specific components (updating and shifting). Common EC was tapped by all nine tasks, and domain specific components were tapped by three tasks each.

These examples show that the structural equation modelling approach can successfully be used to outline the executive control network. It has two advantages. First, it retrieves what is common across tasks, allowing the identification of a cognitive ability despite the task impurity problem. Second, it targets the whole executive network rather than an individual unit, therefore making it possible to observe the relation and interaction between different sub-functions.

It needs to be pointed out that these comprehensive studies assume that the frontal lobe functions are the same for all healthy human beings. They therefore did not compare the properties of the executive control network between participants with different language backgrounds. But because it has been suggested that the bilingual experience might have modified speakers' executive control networks (Grady, Luk, Craik, & Bialystok, 2015; Kroll & Bialystok, 2013), the aim of the present study was to compare the executive control networks of monolingual and bilingual speakers.

The current study addressed three related aims using confirmatory factor analyses (CFA). First, we compared the organisation of the executive control network for monolingual and bilingual speakers. This was done by testing measurement invariance with a multi-group CFA analysis. Second, we compared each sub-function between two participant groups at a latent factor level. This was made possible through factor mean invariance analysis with a multi-group CFA. Third, we investigated whether different subcomponents of the model relate to each other in the same way for monolingual and bilingual speakers.

Monolingual and Bilingual Performance in Inhibition, Shifting and Updating Tasks

Executive control abilities of monolingual and bilingual speakers have been compared in all the three sub-functions proposed by Miyake et al. (2000): inhibition, shifting, and updating. In what follows, we review literature that compared monolingual and bilingual performance in these sub-functions, with a special focus on the tasks used by Friedman et al. (2008) because these are the ones we utilized in the present study.

Inhibition.

Inhibition is the most studied component within the executive control network in the bilingualism literature. In Chapter 2 and 3, we addressed a big portion of the literature that probed speakers' inhibitory control ability with the means of several non-verbal interference tasks. While the Simon task, the Spatial Stroop task and the Flanker task all involve inhibition of prepotent responses, they also rely heavily on the control of attention. By contrast, the inhibition function defined in the Miyake et al. (2000) model focuses on the inhibition of prepotent responses at lower levels. In other words this model addressed inhibitory control at a lower processing level, which was estimated by a Colour Stroop task, a Stop Signal task and an Anti-saccade task. These tasks will be described in detail in the Method section. We next review studies that have investigated the bilingualism effect on performances in these tasks.

The Colour Stroop task and different variations of this task have often been utilized to compare monolingual and bilingual inhibitory control. Evidence for or against the bilingualism effect in this task is far from being conclusive. Some studies have reported a reduced bilingual Colour Stroop effect (Bialystok, Craik, & Luk, 2008), whilst others failed to find any group differences, both in children (Dunabeitia et al., 2014) and adults (Kousaie & Phillips, 2012; Rosselli et al., 2002). In addition, some studies found a bilingualism effect but only under particular conditions. For example, in a Colour Stroop task that manipulated SOA (stimulus onset asynchrony), Coderre and van Heuven (2014) reported a reduced bilingual Stroop interference effect in the zero SOA condition but not in the negative SOA condition (distractor appears earlier than colour stimulus). The bilingualism effect also appears to interact with other factors such as proficiency level and language similarities (Coderre, van Heuven, & Conklin, 2013). Individual difference studies, however, have shown consistently that language proficiency modulates the degree of Stroop interference effect within bilingual speakers (Singh & Mishra, 2012; Tse & Altarriba, 2012).

In comparison to the Colour Stroop task, performance on the Anti-saccade task has not been compared between monolingual and bilingual speakers very often. But similar to the Colour Stroop task, the evidence is inconclusive. Bialystok, Craik, and Ryan (2006) reported no group differences when measuring participants' eye-movements. However they reported a bilingual advantage for manual responses. The authors argued that an eye-gaze is a quite automatic process, while manual responses are more responsive to higher-level executive control. Therefore, the bilingual inhibitory control advantage appears to be related to higher-level executive control. However, this claim was not confirmed by Paap and Greenberg (2013) who reported no group differences with an Anti-saccade task with manual responses, neither in terms of response accuracy nor in terms of response speed.

Concerning the third inhibition task used by Friedman et al. (2008), the Stop Signal task, a difference between monolingual and bilingual speakers was only found with regard to neural measures, not behavioural measures. Morales, Gomez-Ariza, and Bajo (2013) reported no group differences on several measures, including the rate of unsuccessful inhibition, hit-rate, no-signal RT and Stop Signal RT. While Colzato et al. (2008) and Rodriguez-Pujadas et al. (2014) confirmed these results, they did report stronger activations of the anterior cingulated cortex (ACC) for monolinguals compared with bilinguals.

Thus, overall monolingual and bilingual speakers were most often reported to have similar performance in these three inhibition tasks, with exceptional cases where a bilingual advantage had been reported. No study has found evidence for a monolingual advantage. We therefore predicted to find either equal means for the latent inhibition factor or a bilingual advantage, but not likely the other way around.

Shifting.

The shifting function was estimated by Friedman et al. (2008) by means of three switching tasks: the Category Switch task, the Number-Letter task and the Colour-Shape task. Overall, the evidence for a bilingual advantage in these tasks is again mixed. In an initial attempt to address the bilingualism effect on shifting or switching ability, Prior and MacWhinney (2010) utilized a mixed block design Colour-Shape task. In this task, each stimulus is preceeded by a cue, which signals the participant either to make a colour decision or a shape decision. Cues on consecutive trials can either be the same (a repeat trial) or different (a switch trial). The average RT difference between repeat trials and switch trials is the switch cost. Prior and MacWhinney (2010) reported a bilingual advantage in that bilingual speakers had smaller switching cost. However such bilingual advantage was not observed using similar Colour-Shape tasks in later studies (Paap & Greenberg, 2013; Prior & Gollan, 2013). Interestingly, though, Rodriguez-Pujadas et al. (2013) found significant differences in brain involvement despite no differences at behavioural level.

Different switching paradigms have found much more consistent evidence for a bilingual advantage. With a paradigm that involves switching of attention, Christoffels, de Haan, Steenbergen, van den Wildenberg, and Colzato (2015) reported smaller switching cost for bilingual speakers than for monolingual speakers in a Global-Local task. The authors concluded that this suggested a greater mental flexibility for bilinguals. An individual difference study using a Colour-Word Stroop switch task has revealed a relationship between language proficiency and switching ability, with more proficient speakers having smaller switch cost (Tse & Altarriba, 2015). Furthermore, the bilingualism effect is sometimes reflected in response accuracy instead of response latencies. Marzecova et al. (2013) reported a tendency for bilinguals to be more accurate on a social category switching task, especially in situations where one needed to flexibly switch between stimulus-task bindings.

To sum up, evidence is again inconclusive regarding the bilingualism effect on the switching function. Nevertheless, whenever there was a group difference, it pointed to a bilingual advantage. This led to the prediction that we either would find no difference on the latent factor mean for the shifting ability or a higher latent factor mean for the bilinguals.

Updating.

Updating is the least tested sub-function within the executive control network. Friedman et al. (2008) estimated it by the Letter Track task, the Keep Track task and the Space Two-Back task. To our knowledge, there is no study that compared monolinguals' and bilinguals' performance on these tasks. Nevertheless, there are two studies that investigated effects of bilingual experience, such as frequency of language switch, age of second language acquisition, and language proficiency, on updating ability within the bilingual population (Soveri, Rodriguez-Fornells, & Laine, 2011; Yow & Li, 2015). But bilingual experience did not predict updating performance, neither in a space N-back task (Soveri et al. 2011), nor an N-back task (Yow & Li, 2015). Thus, bilingual experience does not seem to impact speakers' updating ability.

Despite the scarce literature on bilingualism and updating, a related construct, namely working memory, has been investigated quite often, and the picture is similarly mixed to the one of inhibition and shifting. Adesope, Lavin, Thompson, and Ungerleider (2010) presented two contrasting hypotheses as to how bilingualism might affect speaker's working memory abilities. The need to keep two languages in mind might pose extra demands on working memory and might thus result in reduced proficiency of processing information in working memory. Alternatively, bilinguals might have increased working memory efficiency because they are able to manage their working memory resources through inhibitory processes. In line with the second hypothesis, some studies reported enhanced bilingual working memory ability (Blom, Kuntay, Messer, Verhagen, & Leseman, 2014; Luo, Craik, Moreno, & Bialystok, 2013, in a verbal working memory but not spatial working memory). But others reported no significant difference in a digit-span task, standard operation span tasks, a non-verbal symmetry span task and a digit span task (Kousaie, Sheppard, Lemieux, Monetta, & Taler, 2014; Ratiu & Azuma, 2015). Similarly to the evidence for advanced inhibition and shifting ability, whenever a group difference was observed, it tended to hint at a bilingual advantage rather than the other way around. Thus, there is no support for Adesope et al.'s (2010) first hypothesis, namely that bilinguals have reduced working memory efficiency.

The executive control network.

In addition to or in place of exerting effects on individual executive control components, it has been suggested that bilingualism might lead to a reconfiguration of the executive network (Bialystok et al., 2006; Kroll & Bialystok, 2013). In other words, instead of quantitatively modifying the functional efficiency of a certain aspect of executive control processing, the bilingual experience might qualitatively alter the way those functions are interrelated. A functional connectivity study has indeed shown that bilingual speakers had higher connectivity in the fronto-parietal control network (FPC) and the default mode network (DMN), which are important networks for executive control (Grady et al. 2015). In addition, the same study also reported stronger correlations among the between-network connectivity. The authors suggested that this might mean bilingualism enhances the links between different network activities. With regard to the current study, if bilinguals have a more interconnected executive network, it is expected that co-variances among the latent factors of the network will be stronger in the bilingual group than in the monolingual group.

To the best of our knowledge, no study to date has tested all three executive sub-functions with more than one test per sub-function and compared monolingual and bilingual speakers' performances. However there are two studies that have investigated different types of executive control within the bilingual population (Soveri et al. 2011, Yow & Li, 2015). Using a within group approach, Soveri et al. (2011) found that bilingualism only affected switching function but not the updating and inhibition function of executive control. Frequent switch of languages in daily life predicted a reduced mixing cost in a non-verbal switching task. Using a similar approach, Yow and Li (2015) found balanced use of two languages had an impact on switching task performance and Stroop task performance, but not performance in an updating task. Despite being intriguing, in contrast to the present study, these studies used only one index for the ability, therefore were not able to account for the task impurity problem outlined earlier. In addition they did not compare monolingual and bilingual speakers directly and therefore were not able to pinpoint the bilingualism effect on executive control.

Based on what we have seen in the literature, we predicted to potentially find a bilingual advantage in the latent factor means of shifting and inhibition, with bilinguals having larger latent factor means of shifting and inhibition, but not updating. In addition, we also predicted higher covariance amongst latent factors in the bilingual group than in the monolingual group.

Method

Participants

228 participants took part in this experiment. All were undergraduate and postgraduate students at the University of Birmingham and participated either for course

credits or cash. For the analysis, participants were selected on the basis of their responses to a questionnaire examining their language use history (see Appendix A-4), adapted from Silverberg and Samuel (2004). Participants were asked to rate their self-perceived proficiency in English. They were asked to list all the languages that they learnt or were able to speak and the age at which they started to learn the language. In addition, bilingual speakers were asked to rate their proficiency in their other language. They also indicated their current language use pattern (e.g. using mainly one language, or using both languages on a daily basis).

In order to be classified as monolingual speakers for the analysis, participants needed to fit the following criteria: 1) participants did not speak another language fluently; 2) did not speak another language on a daily basis; 3) did not learn another language before age 10 (although 9 of them reported being exposed to a second language before age 10, but it was within school setting and they never became proficient or fluent in that language); 4) rated their proficiency in another language 3 or lower on a scale of 7. These criteria led to 100 monolingual speakers being included in the analysis. Bilingual speakers were defined as follows: 1) participants who were fluent in both languages and had rated their proficiency at least 5 on a scale of 7; 2) participants learnt both languages before age 10; 3) participants who used both languages on a daily basis. These criteria led to 100 bilinguals being included in the analysis, see Table 6.1 for various languages spoken by the mixed group of bilingual speakers (six of which were multilingual speakers).

Table 6.1

Languages spoken by the bilingual speakers.

Language	Count	Bilingual/ Multilingual	Language	Count	Bilingual/ Multilingual
Cantonese	14	bilingual	Finnish	1	bilingual
Romanian	8	bilingual	Hebrew	1	bilingual
Chinese	7	bilingual	Hindi	1	bilingual

Punjabi	6	bilingual	Kadazan	1	bilingual
Urdu	6	bilingual	Latvian	1	bilingual
French	5	bilingual	Maltese	1	bilingual
German	4	bilingual	Mauritian creole	1	bilingual
Italian	4	bilingual	Polish	1	bilingual
Mandarin	4	bilingual	Pothwari	1	bilingual
Bulgarian	3	bilingual	Shona	1	bilingual
Danish	3	bilingual	Swedish	1	bilingual
Albanian	2	bilingual	Tamil	1	bilingual
Arabic	2	bilingual	Woilot	1	bilingual
Bengali	2	bilingual	Yoruba	1	bilingual
Farsi	2	bilingual	German/Italian	1	multilingual
Greek	2	bilingual	German/Russian	1	multilingual
Portuguese	2	bilingual	Gujarita/Hindi	1	multilingual
Somali	2	bilingual	Moldovan/Romanian	1	multilingual
Welsh	2	bilingual	Spanish/Basque	1	multilingual
		bilingual	Spanish/French	1	multilingual

Monolingual and bilingual speakers were matched for their age, *t* (198) = 1.33, *p* > .05, education background, χ^2 (2) =3.99, *p* > .05, and social economic status estimated by family income, χ^2 (9) = 11.04, *p* > .05, and self-reported social classes, χ^2 (2) = .33, *p* > .05. See Table 6.2 for a summary of the demographic and language background information of both participant groups.

Table 6.2

Demographic and language background of monolingual and bilingual speakers.

	Monolingual	Bilingual
N (Female)	100 (84)	100 (84)
Age (SD)	20.0 (2.7)	20.5 (3.1)
Median Education level (Min: Max)	2 (2:4)	2 (2:4)
Mean Proficiency Rating for English (SD)	6.51 (.73)	6.36 (.80)

Mean Proficiency Rating for Other Language (SD)	-	5.92 (1.21)
Mean Age of English onset (SD)	-	4.26 (3.95)
Mean Age of the Other Language (SD)	-	0.85 (2.11)
Speak L2 fluently	No	Yes

Tasks and Procedures

Task overview.

The same nine executive control tasks as in Friedman et al. (2008) were administered. Each three tasks tested one of the three sub-functions of executive control: updating, shifting and inhibition. All tasks were programmed using E-prime. In manual response tasks, response time was measured using a Cedrus RB-834 response pad. In the speech production task, responses were audio-recorded and response time was measured with a Cedrus SV-1 Voice Key. In addition to the nine executive control tasks, a computerized Corsi Block task, i.e. test for visio-spatial working memory, was administered to control for non-executive cognitive ability between the two groups. For that task, a mouse was used to make responses. The procedures of the nine executive control tasks followed closely those detailed in Friedman (2008). We nevertheless report the details below.

General procedure.

Each participant completed the experiments in a testing cubicle in two sessions, which were one week apart. Tasks that were believed to tap into the same executive control were not administered in direct succession. The Corsi Block task was administered at the end of the second session. Session 1 included the Letter Track task, the Category Switch task, the Colour Stroop task, the Keep Track task and the Number-Letter task. Session 2 Included the Space Two-Back task, the Stop Signal task, the Colour-Shape task, the Anti-saccade task and the Corsi Block task. Participants then completed the language history questionnaire and were debriefed before they left. The order of tasks was the same for each participant.

Updating tasks.

Letter Track task.

The Letter Track task was adapted from Morris and Jones (1990). In each trial, 5, 7 or 9 letters appeared one at a time without breaks in the centre of a computer screen; each letter was presented for 1500 ms. Participants monitored the letter strings and were asked to report the last three letters that they saw, in the order they appeared. To ensure constant updating, they were instructed to always rehearse aloud the last three letters that they saw after the presentation of each letter. This requested mentally adding the most recently presented letter and drop the first representation in working memory. For example, if the string was ‘C,H,E,N,F’, then the participants should have said, ‘C’, ‘C, H’, ‘C, H, E’, ‘H, E, N’, ‘E, N, F’. And at the end of the trial, the participant should have recalled ‘E, N, F’. Letters were presented in font 18. There were twelve trials, four trials each with five, seven or nine letters. The sequence of trials was randomized for each participant, but the trials themselves had fixed orders of letters. Responses were recorded by the experimenter on an answering sheet. Responses were scored as correct when the three letters were recalled correctly and in the correct order. Each participant completed three practice trials before the actual experiment.

Keep Track task.

The Keep Track task was adapted from Yntema (1963). At the beginning of each trial, participants were shown two to four category names from six possible categories (animals, colours, countries, distances, metals, and relatives). Then 15 words were presented serially at the centre of the screen. Each word was presented for 1500 ms. The 15 words consisted of words from the categories shown at the beginning of the trial. Participants were instructed to remember and report the last word from each category. At the end of each trial, i.e. when the participants had to report the items, the category names appeared again. Words were presented in font 18. There were twelve trials, four trials each with two, three or four categories, leading to a total of 36 words to be recalled. The sequence of trials was randomized for each participant, while the order of words in each trial was fixed. Responses were recorded by the experimenter on an answering sheet. The proportion of correctly recalled words was entered into the analyses. Each participant completed three practice trials (one trial at each difficulty level) before the actual experiment started.

Space Two-Back Task.

The space two-back task was taken from Friedman et al. (2008). The experiment consisted of four blocks. In each block, there were 10 open squares (2 cm) scattered on the screen. During a block, one square at a time became solid black for 500 ms, appearing to be flashing. There were 24 flashes in each block with a SOA of 1500 ms. Each flash was accompanied by a ‘ding’ tone. Participants were instructed to indicate whether the currently flashing square was the same as the one that flashed two trials earlier (i.e. the trial before the immediate previous trial). They were instructed to respond ‘no’ for the first two flashes since there were no two previous squares for those two trials. The order of flashes was predetermined so that there were 6 yes responses in each block. In addition, there was no case where the currently flashing block was the same as one-back or three-back. Responses were made on a keyboard (“0” key for yes and “1” key for no responses). The proportion of correct responses was entered into the analyses. Each participant completed one practice block before the actual experiment.

Shifting tasks.

In the shifting tasks, each stimulus could be classified according to two dimensions (e.g. colour or shape) with two values for each dimension (e.g. red or green on the colour dimension; triangle or circle on the shape dimension). Participants were required to make a forced choice on one of the dimension according to a given cue. Participants were instructed to use the index fingers of each hand to make a response using the Cedrus RB-834 response pad. Responses were mapped onto two buttons.

The general procedure of the three shifting tasks was the same, which was taken from Friedman et al. (2008). In each trial, a cue was shown before the onset of the stimulus. The stimulus then appeared in the centre of the computer screen (but see Number-Letter Task). The cue and the stimulus then stayed until a response was made or until 5000 ms, whichever was earlier. A blank screen was presented for 350 ms before the onset of the next cue. There were four blocks of 48 trials each. In the first and third block, the cue was presented for 150 ms before the stimulus onset. In the second and the forth block, the cue was presented for 1500 ms before the stimulus onset. Participants also completed two blocks of 24 trials for practice. At the beginning of each block, there were 4 additional warm-up trials that were not included in the analysis.

The order of the trials was predetermined and was pseudo-randomised. The trial order fulfilled the following constraints: first, there were 24 switch trials and 24 non-switch trials in each block. Second, no more than 4 switch/non-switch trials would occur in a row. Third, there was no item-specific negative priming trial. That is, the stimulus on a switch trial was not the same as the stimulus on the previous trial (but not for the colour-shape task, which had only four possible stimulus items). Fourth, each stimulus and cue combination occurred with equal likelihood. Response accuracies and the differences between average RTs for switch trials and non-switch trials were entered into the analyses.

Category Switch task.

For this task, stimulus words were object names. Each word could be classified on two dimensions: 1) whether the object that the word refers to was living or non-living and 2) whether the size of the object that the word refers to was larger or smaller than a soccer ball. The cue indicated which of the rules to follow. When the cue was a heart, participants needed to decide between living or non-living; when the cue was a cross with arrow heads, participants needed to decide on size. The 16 words (see Table 6.3) were drawn from Mayr and Kliegl (2000). Texts were presented in font 18.

Table 6.3

Word stimuli used in the Category Switch task.

	Larger than soccer ball	Smaller than soccer ball
Living	shark, lion, oak, alligator	mushroom, sparrow, goldfish, lizard
Non-living	table, bicycle, coat, cloud	pebble, knob, marble, snowflake

Number-Letter task (Rogers & Monsell, 1995).

Stimuli were number-letter or letter-number pairs (e.g. 3U or U3). Numbers used included the even numbers 2, 4, 6 and 8, and the odd numbers 3, 5, 7 and 9. Letters used included the consonants G, K, M and R, and the vowels A, E, I and U. The stimuli appeared inside squares that acted as the cues for the response. The squares appeared either above or below the middle line of the screen, with either edge being approximately 0.3 inch from the middle line. They were centred on the central vertical axis. When the cue was above the middle line of the screen, participants needed to decide whether the number within the stimulus was odd or even. When the cue was below the middle line of the screen, they needed to decide whether the letter within the stimulus was a consonant or a vowel. Texts were presented in font 18.

Colour-Shape task.

The Colour-Shape Task was based on a version by Miyake, Emerson, Padilla, & Ahn (2004). Stimuli were a triangle (about 1.25 inch on each side) and a circle (about 1.1 inch in diameter), either red or green, which were presented in the centre of the screen following cues. The cue was a capital letter that appeared about 0.33 inch above the top edge of the stimuli. When the cue was “C” for colour, one needed to decide whether the colour of the shape was red or green. When the cue was “S” for shape, one needed to decide whether the shape was a triangle or a circle. Cue letters were presented in font 18.

Inhibition tasks.

Colour Stroop task (Stroop, 1935).

Just as Friedman et al. (2008) we used two types of stimuli: neutral stimuli and incongruent stimuli. Incongruent stimuli were colour words (RED, GREEN, BLUE, ORANGE, YELLOW and PURPLE) printed in a different colour (e.g. the word RED printed in green). Neutral stimuli were strings of six Xs printed in one of the six colours. Participants were asked to name the ink colour of the stimuli. Each trial started with a fixation cross presented for 500 ms in the centre of the screen. The stimulus was then presented at the same location for 3000 ms or until a response was made, whichever was earlier. This was followed by a blank screen of 1000 ms. There were 18 practice trials, and the experiment consisted of 60 neutral trials and 60 incongruent trials. The stimuli were in pseudorandom order so that a colour that appeared in one trial would not appear in the subsequent trial. Participants took a short break after every 30 trials. Response time was measured by a voice key. The measure entered in the analyses was the Stroop interference calculated as the difference between the incongruent condition RT and the neutral condition RT.

Stop Signal task (Logan, 1994).

In this task, participant presses a button to indicate whether the object that the word referred to is an animal (e.g. hare) or a non-animal (e.g. ball). 24 words were selected, half of them being animals. Frequencies of animal and non-animal words were matched. Each trial started with a fixation cross for 500 ms. A stimulus was then presented for 1500 ms or until response was made. The screen went blank for 500 ms until the next trial started. Participants practiced for 24 trials (each stimuli word appeared once in the practice). The experiment had 5 blocks. The first block (48 trials) was a simple categorization task (animal versus non-animal). It was used to build up prepotent responses and to calculate the average RT for each participant. The subsequent 4 blocks had 96 trials each and there were two types of trials. 75% of the trials were response trials where participants performed the categorization task normally. The remaining 25% trials were stop trials where participants were required to withhold their responses. Stop trials were signified by a computer-generalized ‘ding’ tone which occurred at three different times: 50 ms or 225 ms before a participant’s average RT, or 50 ms after the onset of the trial. Each type of stop trial occurred equally often across the four blocks. Following Friedman (2008), we used stop signal RT, i.e. the time when the stopping process finished, as the dependent measurement for the analyses. The stop signal RT was calculated as in Friedman (2008). First, the probability of responding at a certain delay was calculated. This was then multiplied by the number of total non-stop signal trials to get a constant n. All non-stop signal responses were then rank-ordered from fastest to slowest. The stop signal RT for that delay is calculated as the RT of the n^{th} response minus the stop signal delay. The stop signal RT was obtained by averaging the stop signal RT for the three delay conditions.

Anti-saccade task (Roberts, Hager, & Heron, 1994).

In this task participants were asked to indicate the direction of a target arrow. Each trial started with a fixation cross at the centre of the screen which lasted for various durations (fixation intervals), that is between 1500 ms and 3500 ms with intervals of 250 ms. A black cue square (4 mm) then appeared about 3 inches away, on one side of the fixation cross for 150 ms. A target arrow (7 mm, pointing left, up or right) in an open square (10 mm) then appeared about 3.6 inches away from the centre, on the opposite side of the screen for 175 ms. The target was backward masked with grey cross-hatching which remained until response or 5000 ms, whichever was earlier. Participants were seated so that their eyes were 45 cm from the computer screen. Participants practiced for 22 trials and completed two blocks of 45 trials each. Each fixation interval appeared equally likely and was randomized. The measure entered into the analyses was the proportion of correct responses.

Computer-based Corsi Block task.

In the computer-based Corsi Block task, 9 white squares (3 cm) were scattered in a pseudo-random way on a blue background. A ‘ding’ tone signified the start of a trial. Then squares turned black one at a time for 1 second each, appearing to be flashing. Trials consisted of 2 to 9 flashing squares. Participants were instructed to click on the squares that had turned black in order of flashing. Each participant started with a 2-unit sequence. If a sequence was correctly identified, the length of the next sequence increased by one. The experiment stopped when participants made 5 consecutive errors. The measure entered into the analyses was the maximum length of the sequence that a participant successfully recalled.

Results

Results will be presented in two steps. We will first present the results for each task (single task analyses) and then results from the confirmatory factor analysis (CFA).

Single Task Analyses

For tasks that used response accuracy as index (the three updating tasks and the Anti-saccade task), participants with performances below chance were excluded (4 in the Space Two-Back task, 1 monolingual and 3 bilinguals; 13 in the Anti-saccade task, 6 monolinguals and 7 bilinguals). Accuracies were arcsine-transformed before submitted to independent sample t-tests, comparing performance between the two participant groups. Stop Signal RT was also submitted to an independent sample *t*-test. For tasks that used RT as index (Colour Stroop task and the three switching tasks), response times beyond 2 *SD* of the mean were excluded before calculating the final average RT for each participant. Mixed design *ANOVAs* were performed on RT data, with Condition (e.g. neutral versus incongruent in the Colour Stroop task) being a within group variable and Language Group (monolingual versus bilingual) a between group variable. Response accuracies for these tasks were also analysed using mixed design *ANOVA*s.

Corsi Block task.

Results of the Corsi Block task (see Figure 6.3) showed that monolinguals and bilinguals did not differ in terms of their spatial memory span, t (198) = .48, p > .05.

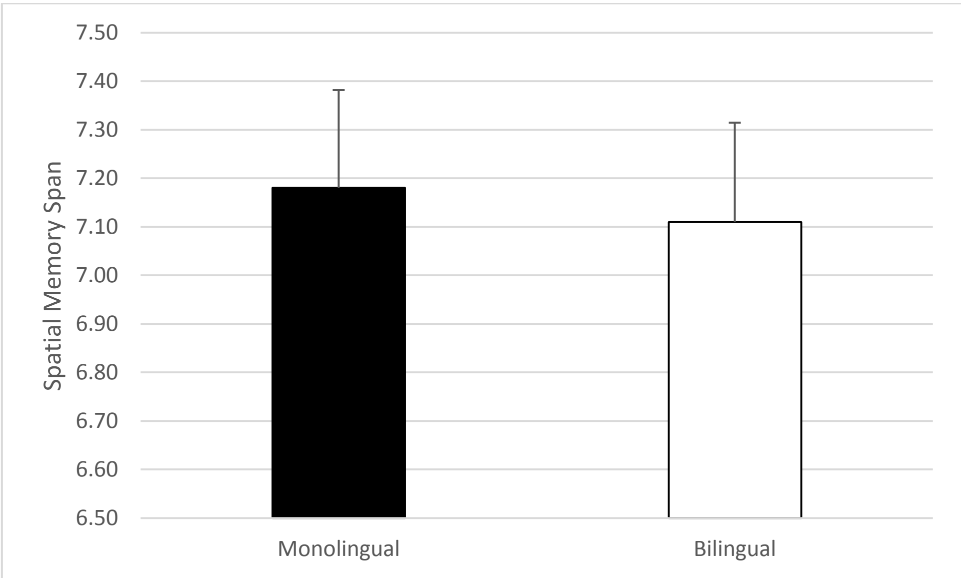


Figure 6.3. Average Spatial Memory Span for monolingual and bilingual speakers in the Corsi Block task. Error bars represent 95% confidence interval.

Updating tasks.

There was no significant difference in accuracy between the two participant groups in the Letter Track task, t (198) = 1.03, p > .05, and the Keep Track task, t (198) = .13, p > .05. Monolinguals outperformed bilingual speakers in the Space Two-Back task, t (194) = 2.14, p = .034 (see Figure 6.4 for task performance summaries).

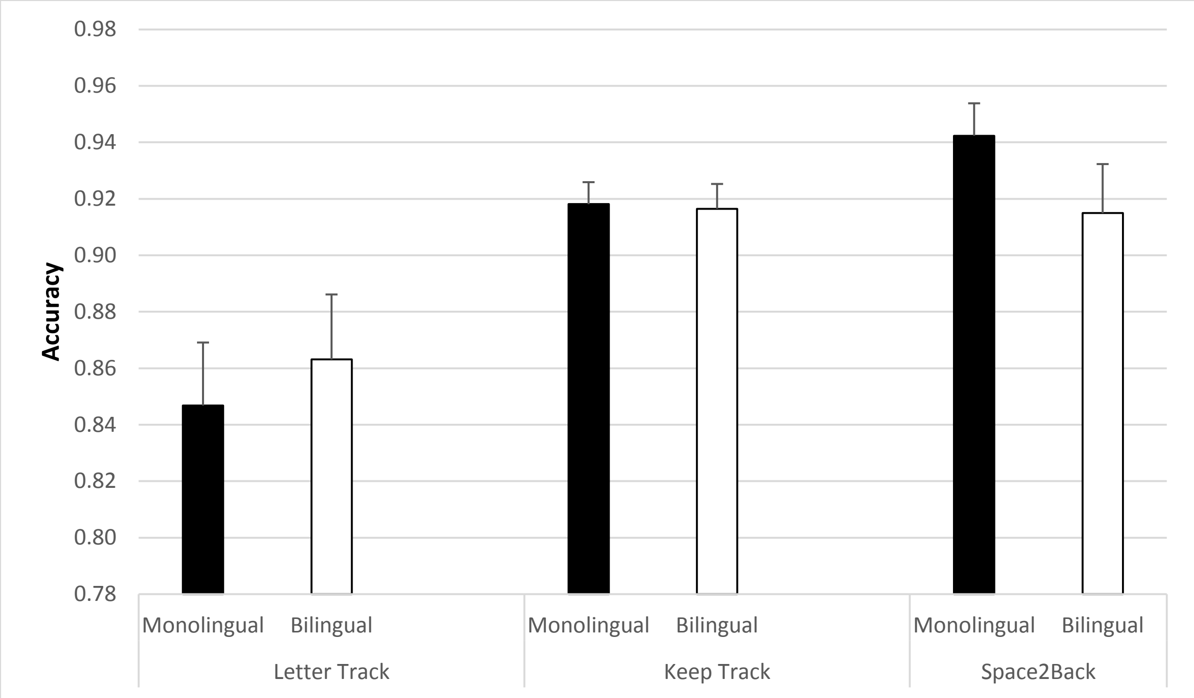


Figure 6.4. Response accuracies in updating tasks for monolingual and bilingual speakers. Error bars represent 95% confidence interval.

Inhibition tasks.

Colour Stroop task.

Results are presented in Figure 6.5. There was a significant main effect of Condition, F (1,198) = 877.88, $p < .001$, $\eta_p^2 = .82$. In line with the Stroop effect, the incongruent condition led to longer responses than the neutral condition. There was also a main effect of Group, with monolingual speakers responding faster than bilingual speakers overall, F (1,198) = 22.39, $p < .001$, $\eta_p^2 = .10$. The interaction between Condition and Group was also significant, F (1,198) = 7.63, $p = .006$, $\eta_p^2 = .04$. Follow up independent sample *t*-tests revealed that monolinguals were faster than bilinguals speakers in the neutral condition, t (198) = -4.49, $p < .001$, and in the incongruent condition, t (198) = -4.80, $p < .001$. Follow up paired sample *t*-tests showed that both groups showed a significant Stroop effect, t (99) = -23.02, $p < .001$ for monolingual, and t (99) = -19.94, $p < .001$ for bilinguals. This suggested that the significant interaction was driven by bilingual speakers suffering from larger Colour

Stroop effects than monolingual speakers. This was confirmed by an independent sample *t*-test, t (198) = -2.76, p = .006.

In terms of response accuracies, there was a main effect of Condition, F (1,197) = 91.68, $p < .001$, $\eta_p^2 = .32$, with the incongruent condition leading to more error responses than the neutral condition. The main effect of Group was not significant, F (1,197) = .40, $p > .05$, $\eta_p^2 = .002$. The interaction between Condition and Group was significant, F (1,197) = 7.12, $p = .008$, $\eta_p^2 = .04$. This interaction was followed up with *t*-tests. Paired sample *t*-tests revealed that both groups showed significant worse performance in the incongruent condition than in the neutral condition, t (98) = 4.60, $p < .001$ for monolinguals, and t (99) = 9.24, $p < .001$ for bilinguals. Independent sample *t* tests revealed that two groups did not differ in the neutral condition, t (197) = .83, $p > .05$, whereas monolinguals were marginally more accurate in the incongruent condition, t (197) = 1.78, $p = .08$.

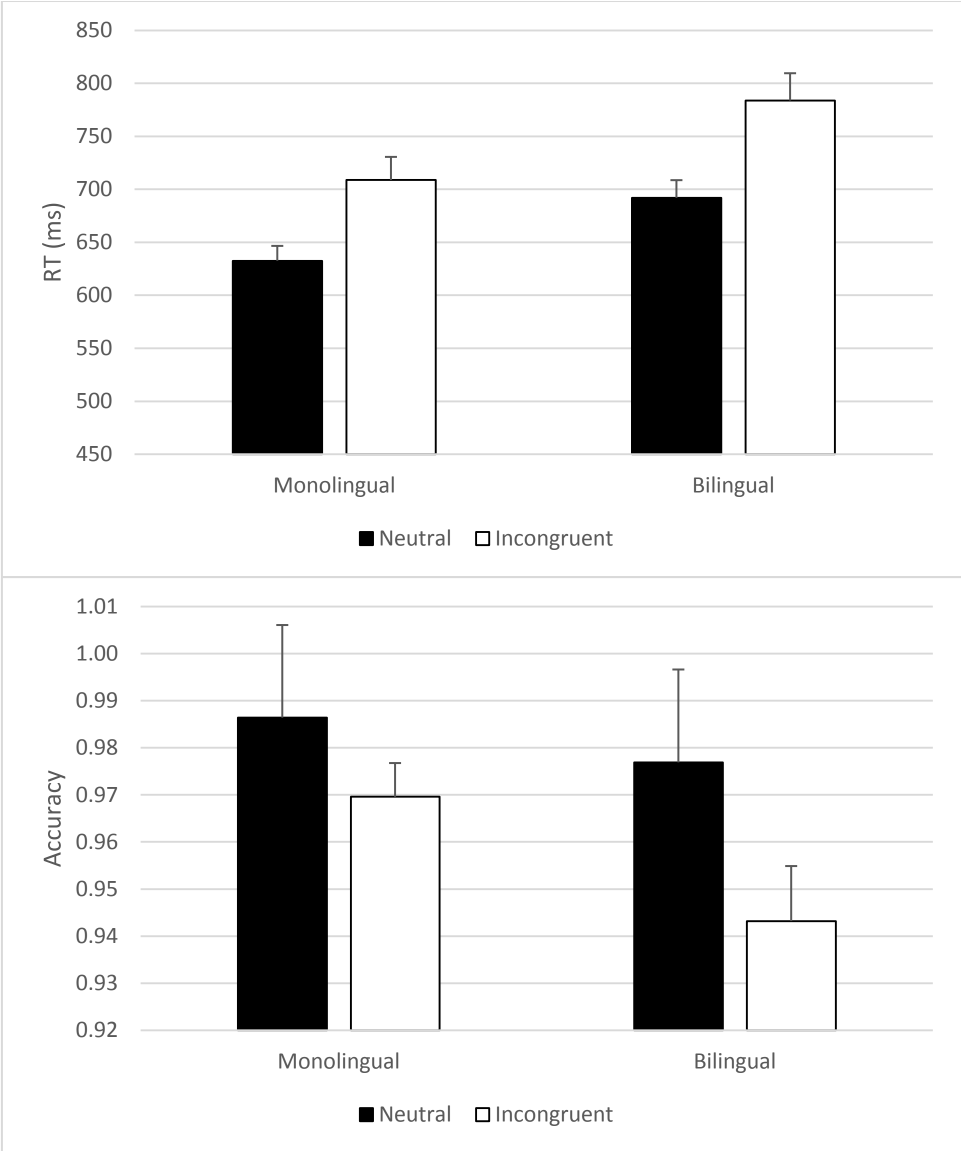


Figure 6.5. *RT and accuracy in the Colour Stroop task for monolingual and bilingual speakers. Error bars represent 95% confidence interval.*

Stop Signal task.

Results showed that monolingual speakers had shorter Stop Signal RTs than bilingual speakers, *t* (147.58) = 2.49, *p* = .014 (see Figure 6.6).

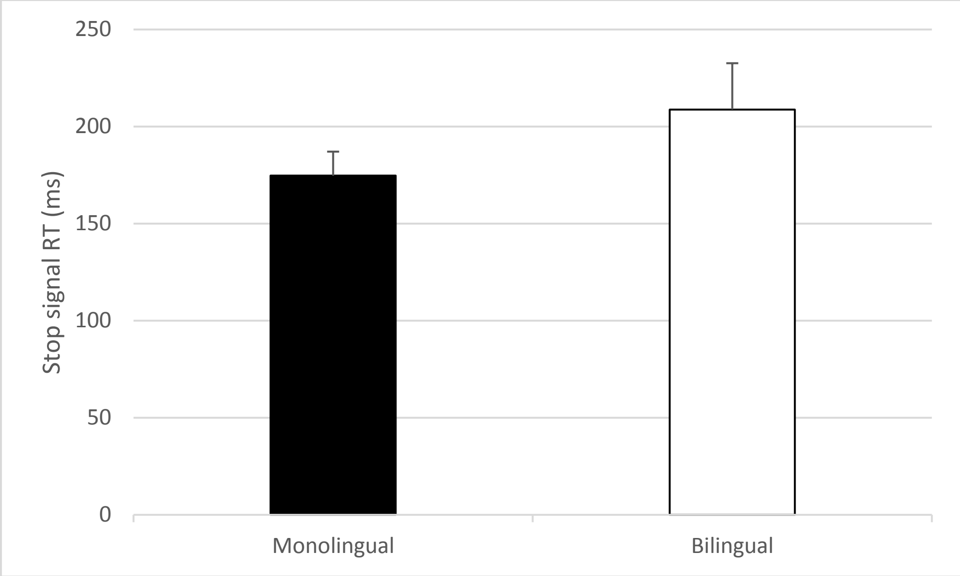


Figure 6.6. Stop Signal RT for monolingual and bilingual speakers. Error bars represent 95% confidence interval.

Anti-saccade task.

There was no significant difference between the two participant groups in the Anti-saccade task, t (185) = .96, p > .05 (see Figure 6.7).

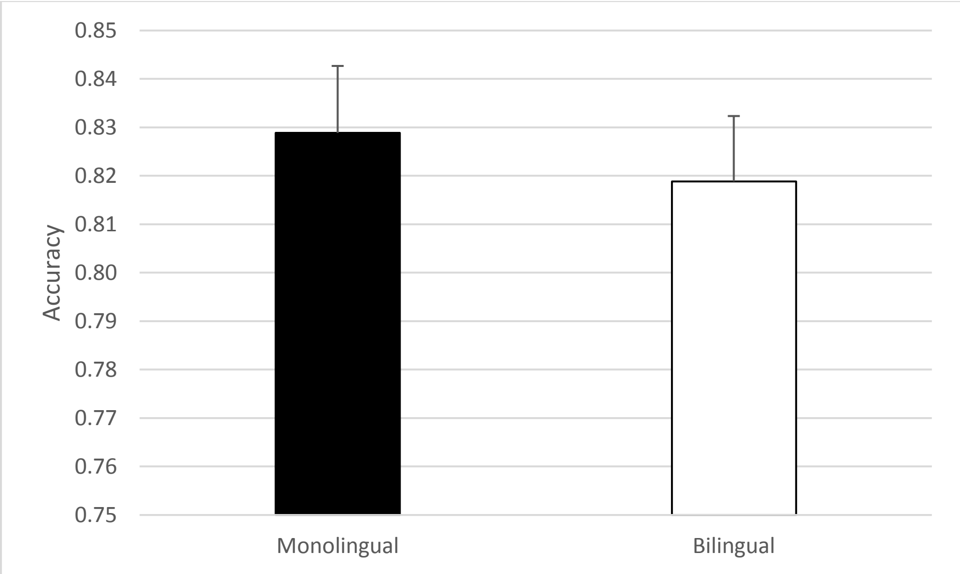


Figure 6.7. Accuracy for the Anti-saccade task for monolingual and bilingual speakers. Error bars represent 95% confidence interval.

Shifting Tasks.

Category Switch task.

Results of the Category Switch task are presented in Figure 6.8. Analyses revealed a significant main effect of Condition, F (1,198) = 364.26, $p < .001$, $\eta_p^2 = .65$. Responses were slower in the switch than non-switch condition. There was also a main effect of Group, F (1,198) = 16.75, $p < .001$, $\eta_p^2 = .08$, with monolingual speakers responding faster than bilingual speakers overall. There was no significant interaction, F (1,198) = 1.05, $p > .05$, $\eta_p^2 = .005$, suggesting that the two groups had similar switch costs. For response accuracy, there was a significant main effect of Condition, F (1,198) = 6.36, $p = .012$, $\eta_p^2 = .03$, with switch trials leading to more erroneous responses than non-switch ones. The main effect of Group was not significant, F (1,198) = 1.7, $p > .05$, $\eta_p^2 = .009$. Neither was there a significant interaction between Condition and Group, F (1,198) = 1.35, $p > .05$, $\eta_p^2 = .007$.

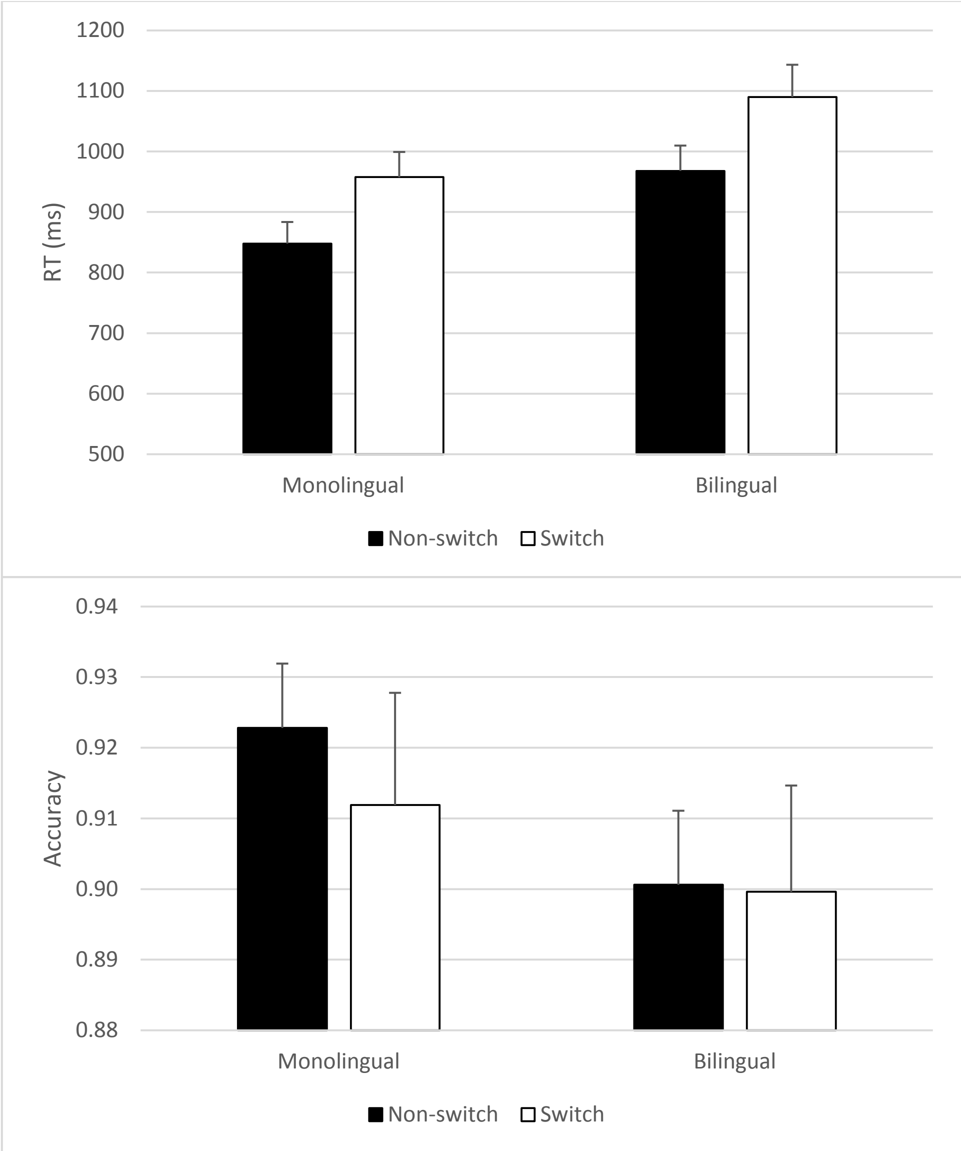


Figure 6.8. RT and accuracy for monolingual and bilingual speakers in the Category Switch task. Error bars represent 95% confidence interval.

Number-Letter task.

Results (see Figure 6.9) showed a significant main effect of Condition, F (1,198) = 518.54, p < .001, η_p^2 = .72, with responses on switch trials being slower than on non-switch trials. There was no main effect of Group, F (1,198) = .03, p > .05, η_p^2 < .001, and no Condition by Group interaction, F (1,198) = .03, p = .862, η_p^2 < .001, suggesting that the two groups did not differ in terms of response speed. Regarding response accuracies, results revealed a significant main effect of Condition, F (1,198) = 37.06, p < .001, η_p^2 = .16, with non-switch trials leading to more accurate responses than switch trials. There was no main effect of Group, F (1,198) = 2.3, p > .05, η_p^2 = .012, but a significant interaction between Condition and Group, F (1,198) = 5.4, p = .021, η_p^2 = .03. This was due to bilingual speakers showing a smaller switch cost (accuracy on switch trials minus accuracy on non-switch trials), t (198) = 2.33, p = .021.

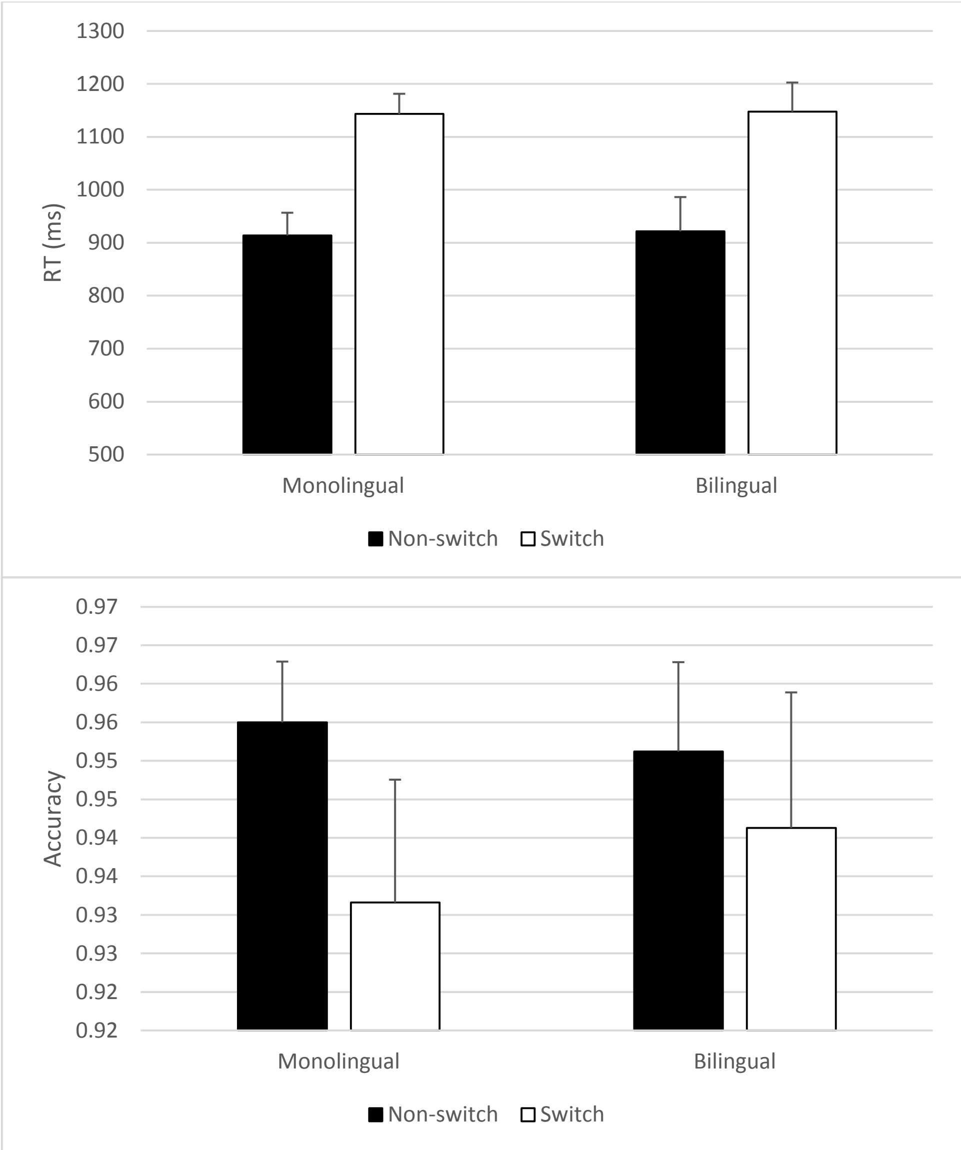


Figure 6.9. RT and accuracy for monolingual and bilingual speakers in the Number-Letter task. Error bars represent 95% confidence interval.

Colour-Shape task.

Results (see Figure 6.10) showed a significant main effect of Condition, F (1,198) = 374.65, $p < .001$, $\eta_p^2 = .65$, with responses on switch trials being slower than on non-switch trials. There was no main effect of Group, F (1,198) = 2.64, $p > .05$, $\eta_p^2 = .01$, and no Condition by Group interaction, F (1,198) = .56, $p > .05$, $\eta_p^2 = .003$, suggesting that the two groups did not differ in terms of response speed. Regarding response accuracies, results revealed a significant main effect of Condition, F (1,198) = 95.13, $p < .001$, $\eta_p^2 = .33$, with non-switch trials leading to more accurate responses than switch trials. There was no main effect of Group, F (1,198) = 1.36, $p > .05$, $\eta_p^2 = .007$, but a significant interaction between Condition and Group, F (1,198) = 5.5, $p = .02$, $\eta_p^2 = .03$, with bilingual speakers showing a smaller switch cost (accuracy on switch trials minus accuracy on non-switch trials), t (198) = 2.35, $p = .02$.

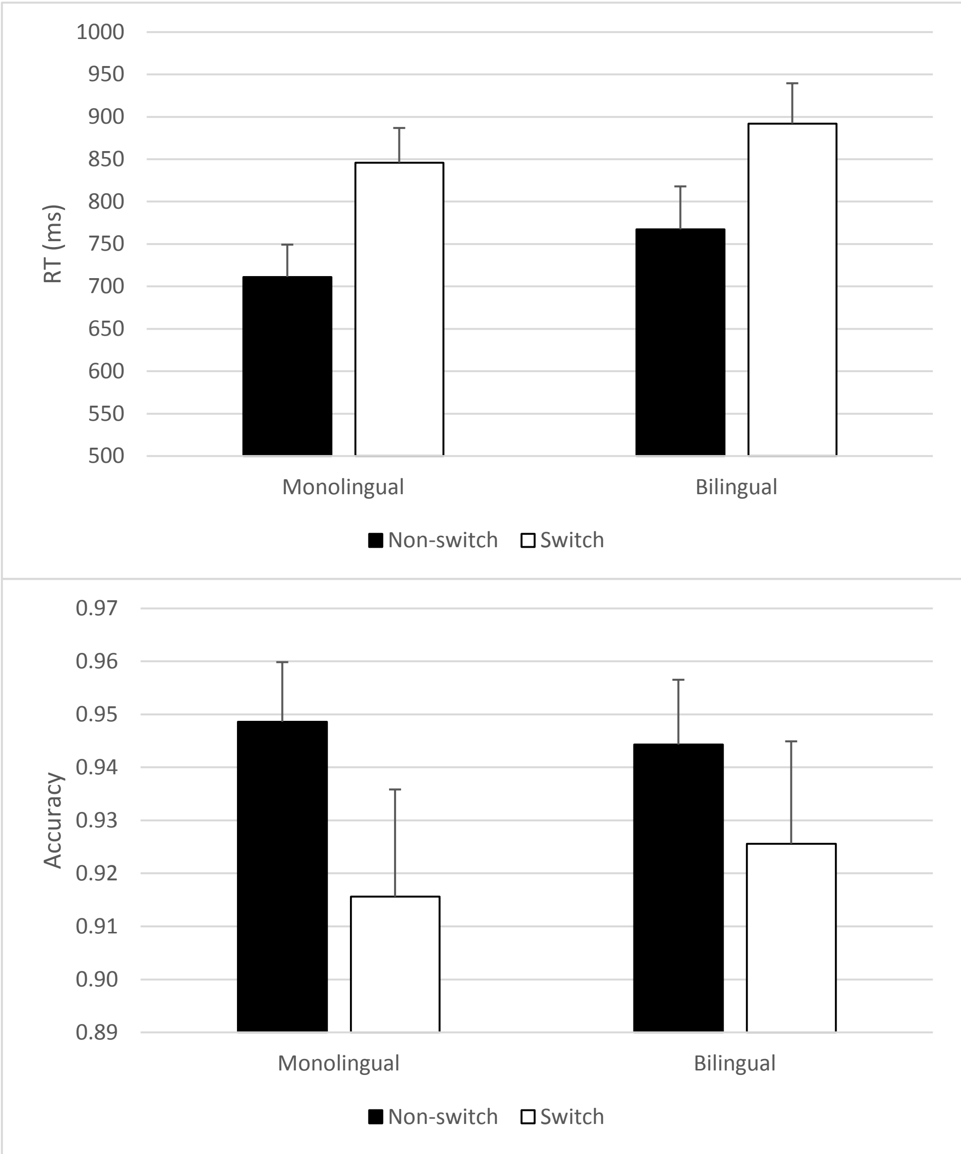


Figure 6.10. RT and accuracy for monolingual and bilingual speakers in the Colour-Shape task. Error bars represent 95% confidence interval.

Summary of Single Task Analyses

Results showed that general effects that have been reported for these tasks were replicated in both groups, including the Colour Stroop effect and the switch effects in all switching tasks, in terms of both response speed and accuracy.

Importantly, no difference in performance was found in the Corsi Block task, which was used as a control measurement of non-executive aspects of cognitive processing. This allows further comparisons between the performances of two groups on executive control tasks. Overall, group comparison showed some processing advantages of monolingual speakers. Bilinguals only showed superior performance in switching accuracy. With regard to the updating performances, the two groups did not differ in terms of accuracy in any task, except for the Space Two-Back task, which showed a monolingual advantage. With regard to inhibition performance, monolinguals generally outperformed bilinguals. Monolinguals were overall faster in the Colour Stroop task and showed smaller Stroop effects in RTs and accuracies. Monolinguals had also shorter Stop Signal RTs, which suggested stronger inhibition ability. No difference was found in the Anti-saccade task. With regard to the shifting performance, both groups showed similar RT switching costs in all three switching tasks. But monolinguals were overall faster in the Category Switch task. By contrast, bilinguals showed a smaller switch cost in terms of response accuracy in the Number-Letter task and the Colour-Shape task.

Confirmatory Factor Analysis

Data preparation.

Following the procedure by Friedman et al. (2008), several transformation steps were carried out to the raw data to enhance the normality of the data in order to perform the confirmatory factor analysis. For accuracy measures, participants with accuracies below chance were excluded (see Single Task Analyses). As in single task analysis, accuracy data were arcsine transformed. For RT measures, RT was based on accurate responses that were above 200 ms. Within-participant trimming procedures were applied that RT beyond 3SD of a participants mean was excluded. To improve normality, observations beyond three *SD*s from the group mean were replaced with three *SD*. Last but not least, the directionality of the RT measures were adjusted so that higher scores always meant better performance.

Single task analyses showed that sometimes result patterns for response accuracy and response times differed, e.g. in the Colour Stroop task and in the switching tasks. In order to account for these differences and to combine RT and accuracy into a single measure, inverse efficiency scores were computed (see Townsend & Ashby 1978). This was done by dividing the mean RTs of the correct trials by the proportion of accurate responses. The inverse efficiency score was used in the CFA analysis instead of the raw switching cost or the raw Stroop effect.

Data analysis was carried out in two stages, namely 1) model identification and 2) invariance analyses. In the first stage, a confirmatory factor analysis (CFA) was conducted to identify a model that best fit the data of both participant groups. This allowed us to examine the structure of the cognitive control network in the two groups. In the second stage, invariance analysis was carried out with the best-fitting model to test whether the measurements are invariant across the two groups. Invariances of latent factor means and latent factor covariances across the two groups were also examined. This allowed us to compare the two groups on individual sub-functions of the executive control network as well as the relationship between those functions.

CFA analysis was carried out in R (version 3.1.1) with the lavaan package 0.5-18 (Rosseel, 2012). The robust variance of maximum likelihood estimator "mlr" was utilized because some of the indicators did not meet the normality assumption. This estimator provides maximum likelihood estimation with robust (Huber-White) standard errors and a scaled test statistic that is (asymptotically) equal to the Yuan-Bentler test statistic. Missing data were handled using the list-wise deletion method. This led to 12% of the monolingual and 14% of the bilingual data being excluded.

Model identification.

Model evaluation criteria.

The following statistics were considered jointly to determine the fitness of a model, Chi-square, CFI (Comparative Fit Index), RMSEA (Root Mean Square Error of Approximation) and SRMR (Standardized Root Mean Square Residual). For Chi-square goodness of fit, .05 was chosen as the significance level. A chi-square smaller than .05 suggests the predicted outcome differed significantly from the observed outcome, hence the model was considered to be a bad fit. According to Hu & Bentler (1999), CFI statistics greater than .95 indicates a good fit, while CFI greater than .90 but smaller than .95 indicates adequate fit. Last but not least, cut-offs for RMSEA and SRMR were set to be .06 and .08 respectively for a good fit, and .08 and .10 for a mediocre fit (West, Taylor, & Wu, 2012).

The primary goal of this study was (a) to test whether the factor structure for executive control is the same for monolingual and bilingual speakers and (b) if yes, further compare the factor means and factor covariance of the two groups. Therefore, we started with testing the executive control model proposed by Miyake et al. (2000). As introduced, this model specified three latent factors, updating, inhibition and shifting. Each of the latent factors has three indicator variables (see Figure 6.1 in the introduction). Next, we present the results of the confirmatory factor analysis to verify the factor structure as proposed by Miyake (2000) in each group separately.

Table 6.4 shows the descriptive statistics for the nine executive control tasks for the two groups and Table 6.5 lists the correlations among the nine executive control tasks. The full three-factor model did not fit the monolingual data well, χ^2 (24) = 45.55, p = .005, CFI = .725, RMSEA = .101 and SRMR = .065. The examination of modification indices (MI) showed that the MI for the residual covariance of the Colour Stroop task and the Number-Letter task was 8.19 (higher than the critical value 3.84 for chi-square with df = 1). The full three-factor model did not fit the bilingual data well either, χ^2 (24) = 43.30, *p* = .009, CFI = .742, RMSEA = .097 and SRMR = .079. The examination of modification indices (MI) showed that the MI for the residual covariance of the Category Switch task and the Stop Signal task was 15.70 (higher than the critical value 3.84 for chi-square with *df* = 1). In order to accommodate both groups, we next tested a model where residual covariances between the Colour Stroop task and the Number-Letter task as well as those between the Category Switch and the Stop Signal task were freely estimated (see Figure 6.11). This model provided a good fit for the monolinguals, χ^2 (22) = 26.46, *p* = .23, CFI = .943, RMSEA = .048 and SRMR = .053. This model also fit well for the bilingual data, χ^2 (22) = 26.84, *p* = .22, CFI = .935, RMSEA = .051 and SRMR = .069.

Table 6.4

Descriptive statistics for the nine executive control tasks for monolingual and bilingual speakers.

	Monolingual						Bilingual					
Task	N	Mean	SD	Min	Max	Skewness	N	Mean	SD	Min	Max	Skewness
Letter Track	88	0.9	0.4	0.3	1.6	0.6	86.0	0.9	0.4	0.2	1.6	0.3
Keep Track	88	1.0	0.1	0.6	1.3	-0.1	86.0	1.0	0.2	0.7	1.5	0.5
Spatial 2-Back	88	1.1	0.2	0.5	1.4	-1.1	86.0	1.0	0.0	0.3	1.4	-0.9
Colour Stroop	88	92.0	48.9	-35.4	297.4	0.9	86.0	122.4	62.0	24.8	299.5	0.8
Stop Signal	88	171.8	49.5	101.3	484.7	3.2	86.0	192.1	72.7	53.0	484.7	2.1
Anti-saccade	88	0.7	0.2	0.5	1.2	0.5	86.0	0.7	0.2	0.5	1.2	0.8
Category Switch	88	135.2	85.8	-67.2	385.7	0.4	86.0	132.7	112.4	-75.0	481.7	0.8
Number-Letter	88	270.9	162.6	28.0	792.1	1.3	86.0	245.6	140.2	-8.1	792.1	1.0
Colour-Shape	88	178.9	116.5	-48.7	544.2	1.1	86.0	146.2	106.1	-128.2	474.8	0.7

Table 6.5

Correlation coefficients between executive control tasks for monolingual and bilingual speakers.

Monolingual	Keep Track	Spatial 2-Back	Colour Stroop	Stop Signal	Anti-saccade	Category Switch	Number-Letter	Colour-Shape
Letter Track	0.17	.394**	0.01	-.204*	.299**	0.09	-0.06	-0.10
Keep Track		0.16	-0.09	-0.17	0.12	0.00	0.01	0.01
Spatial 2-Back			-0.02	-.222*	.209*	0.15	-0.19	0.01
Colour Stroop				0.03	0.01	0.07	0.20	0.07
Stop Signal					-.262*	0.14	0.06	.198*
Anti-saccade						0.02	-0.04	-0.11
Category Switch							.379**	.492**
Number-Letter								.475**

Bilingual	Keep Track	Spatial 2-Back	Colour Stroop	Stop Signal	Anti-saccade	Category Switch	Number-Letter	Colour-Shape
Letter Track	.288**	0.17	-0.16	-.244*	.341**	-0.01	-0.04	0.08
Keep Track		-0.03	-0.16	-0.15	0.05	0.16	0.09	0.15
Spatial 2-Back			-0.01	-0.10	.364**	-0.14	-.221*	-0.12
Colour Stroop				0.06	-0.20	-0.13	0.04	-0.08
Stop Signal					-.229*	.239*	-0.04	0.00
Anti-saccade						-.256*	-.332**	-.260*
Category Switch							.294**	.429**
Number-Letter								.516**

* correlation was significant at .05 significant level. ** correlation was significant at .01 significant level.

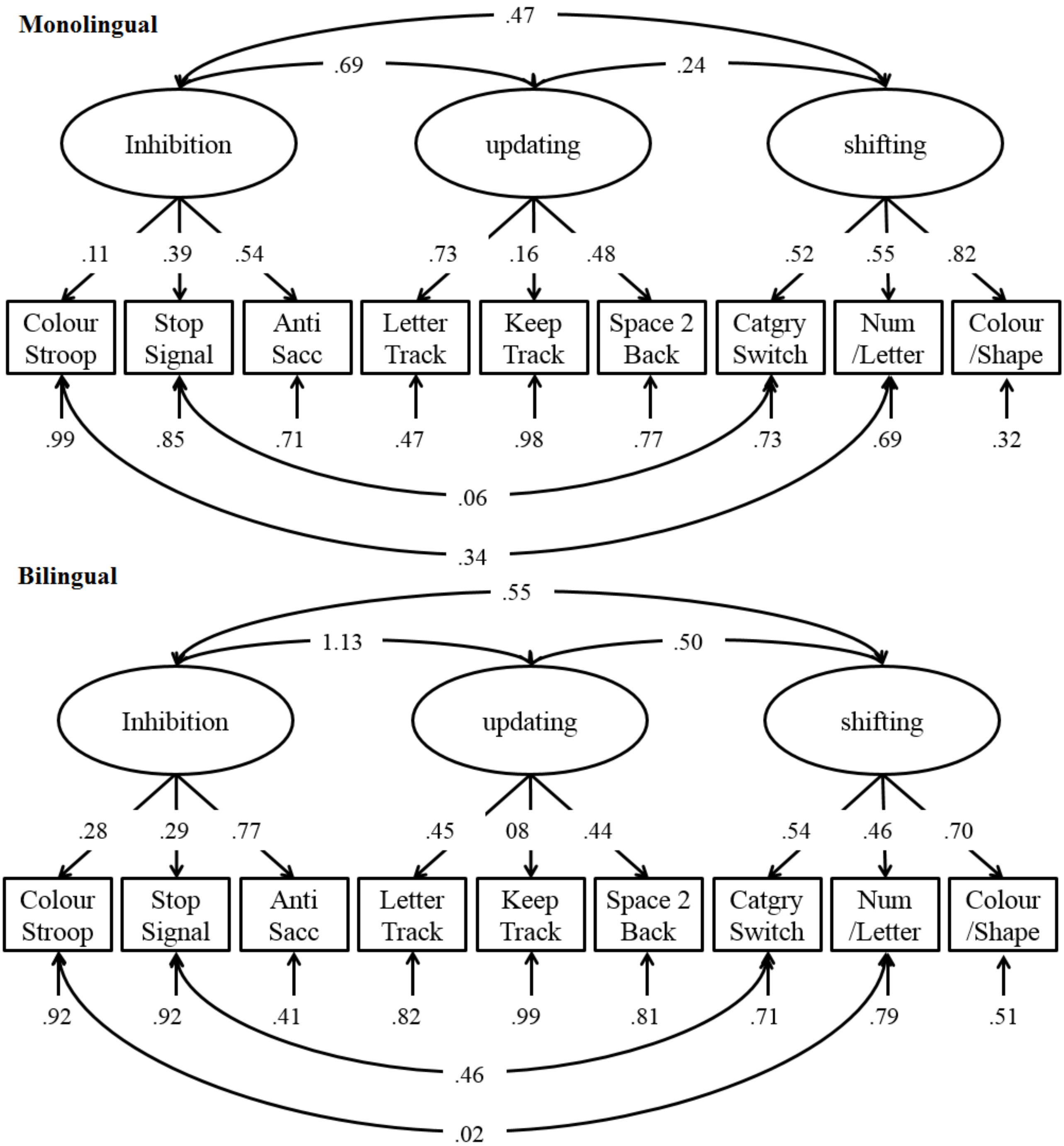


Figure 6.11. Confirmatory factor analysis model of the three executive functions inhibition, updating, and shifting. Numbers on arrows are standardized factor loadings, those next to the vertical arrows (pointing up) are residual variances, and those on curved double-headed arrows are inter-factor correlations.

Figure 6.11 shows the structural model with fitting indices for monolinguals and bilinguals. Inspection of the inter-factor correlations showed that inhibition and updating had a correlation of more than 1 for the bilingual group. This means that these two latent constructs were highly correlated and there might be no need to separate them at the latent factor level. In order to test this, we next fitted the data with a two-factor model, with inhibition and updating collapsed (common executive control, see Figure 6.12). This model provided a good fit for the monolingual data, $\chi^2\,(24) = 28.82$, $p = .28$, CFI = .939, RMSEA = .048 and SRMR = .059, and the bilingual data, $\chi^2\,(24) = 25.93$, $p = .36$, CFI = .974, RMSEA = .031 and SRMR = .068. Comparing this model to the full three-factor model showed that it was not significantly worse than the full model, neither for the monolingual data, $\Delta\chi^2\,(2) = 2.37$, $p = .30$, nor for the bilingual data, $\Delta\chi^2\,(2) = 0.18$, $p = .91$. Furthermore, this two-factor model was better than a one-factor model (collapsing all latent three factors into one common EC factor), which did not fit the data well for monolinguals, $\chi^2\,(27) = 59.96$, $p < .001$, CFI = .580, RMSEA = .118 and SRMR = .087, nor bilinguals, $\chi^2\,(27) = 49.57$, $p = .005$, CFI = .699, RMSEA = .099 and SRMR = .085.

Therefore the two-factor model in Figure 6.12 was identified as the best-fitting model, with Letter Track, Keep Track, Space Two-Back, Colour Stroop, Stop Signal and Anti-saccade task loading onto the first latent factor (common EC), with Category Switch, Number-Letter task and Colour-Shape task loading on to the second latent factor (shifting). Factors were allowed to co-vary, and the residuals between Colour Stroop and Number-Letter as well as between Category Switch and Stop Signal were allowed to co-vary with each other as well.

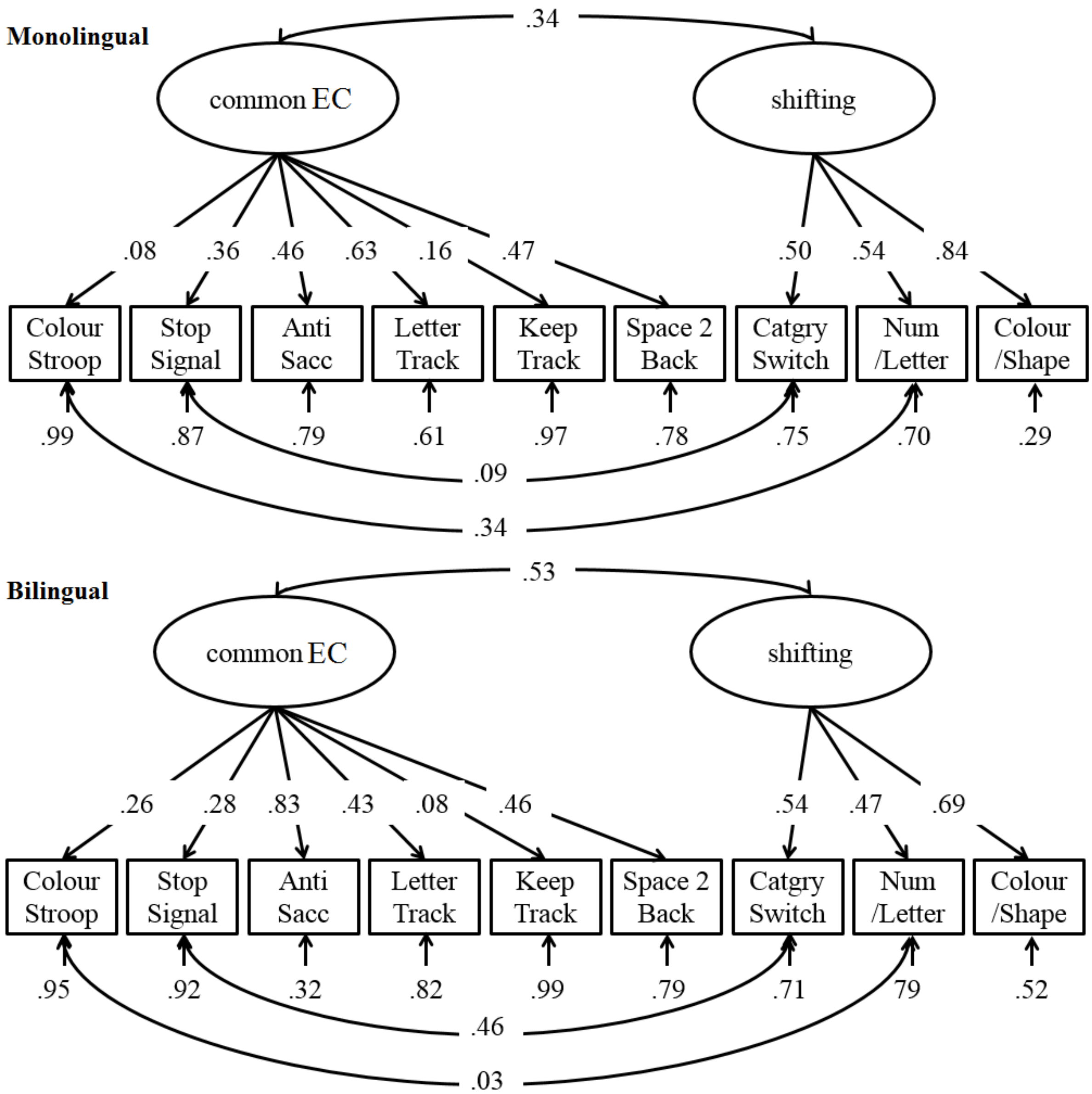


Figure 6.12. Two factor model for executive functions after collapsing inhibition and updating into a single factor (common EC). Numbers on arrows are standardized factor loadings, those next to the vertical arrows (pointing up) are residual variances, and those on curved double-headed arrows are inter-factor correlations.

Model invariance.

To test for factorial invariance, a forward fitting approach was adopted where a model with no constrains is fitted first. This model is then modified by adding invariance constraints, and the new and old model are compared using a chi-square fitting index and other fitting indices (see below for details).

In order to compare group latent factor means, measurement invariance has to be established before meaningful comparisons between latent factor means are possible. Once strong invariance is established, or invariance constraints are imposed on all measurement intercepts, the entire mean structure for the factor model across groups can be identified, allowing a latent mean invariance evaluation (Green & Thompson, 2012). Therefore, the order of the test was as follows: individual groups were fitted first and then configural invariance was tested, followed by metric invariance and scalar invariance. Partial invariance models were tested if a more constrained model did not provide a good fit. This was done by inspecting modification indices and constraining the most non-invariant item. Partial invariance is established when $\Delta\chi^2$ is not statistically significant compared with a less constraint model. In addition to factor mean invariance, factor covariance invariance was also examined, which showed whether the latent factors were related to each other in the same way in the two groups.

Invariance evaluation criteria

Two types of fitting indices were considered jointly. First, the traditional change in chi-square values was considered. Invariance is established or accepted if the chi-square test is not statistically significant at significance level of .05. Second, as chi-square is sensitive to sample sizes, other indices were considered collaboratively to evaluate whether an invariance model should be retained or not. These indices included changes in the fitting indices. According to Chen (2007), delta CFI < .01, delta RMSEA < .015 and delta SRMR < .03 indicates an acceptable model fit for a more restrictive model for metric invariance (i.e. factor loadings are the same across groups), and delta CFI < .01, delta RMSEA < .015 and delta

SRMR < .01 indicates a good fit for scalar invariance (i.e. indicator intercepts are the same across groups).

Invariance analyses

As the two factor model with two sets of correlated residual variances fit both groups reasonably well, invariance analyses were carried out with this model. Table 6.6 summarizes model fits indices for each level of the invariance test. It needs to be pointed out that correlated residual variances should be treated with caution, as they may indicate potential mini-factors.

Table 6.6

Summary of model invariance analyses.

Model ID	Model Name	n	χ2	df	p	CFI*	TFI	RMSEA	SRMR	Δχ2	Δdf	p	ΔCFI	ΔRMSEA	ΔSRMR
a	Monolingual	88	28.82	24	0.227	0.94	0.91	0.048	0.059						
b	Bilingual	86	25.93	24	0.357	0.97	0.96	0.031	0.068						
1	Configural	174	54.59	48	0.239	0.96	0.94	0.040	0.064						
2	Metric	174	63.76	55	0.196	0.94	0.93	0.043	0.068	9.30	7	0.232	-0.014	0.003	0.004
3	Scalar	174	86.99	62	0.020	0.84	0.81	0.068	0.081	22.97	7	0.002	-0.106	0.025	0.013
3a	Scalar(a)	174	75.53	61	0.100	0.91	0.89	0.052	0.075	11.49	6	0.074	-0.038	0.009	0.007
4	Factor Mean	174	84.50	63	0.037	0.86	0.84	0.063	0.082	6.93	2	0.003	-0.045	0.011	0.007
4a	Common EC Equal	174	78.16	62	0.081	0.90	0.88	0.055	0.079	2.12	1	0.146	-0.010	0.003	0.004
4b	Shifting Equal	174	79.13	62	0.070	0.88	0.87	0.056	0.078	3.20	1	0.074	-0.025	0.004	0.003
5	Factor Covariance	174	78.95	62	0.085	0.90	0.88	0.054	0.081	0.07	1	0.392	0.001	-0.001	0.002

* CFI robust

Meade, Johnson, and Braddy (2008) have recommended assessing model fitting for each individual group before carry out invariance analysis. The first two rows in Table 6.6 summarise the model fitting indices for the base model. All indices were within the limit of a good fit for both the monolingual and the bilingual group, except for the CFI for the monolinguals, which was between .90 and .95. This suggested that the model provided not a good, but an acceptable fit to the data for monolingual speakers. Therefore, there is enough evidence to proceed to the invariance analysis with the base model.

Starting with configural invariance, results suggested that configural invariance existed across monolinguals and bilinguals, as all fitting indices met pre-specified good model fit criteria. This configural invariance means that the factor model or factor structure was the same across both groups.

Next, metric invariance was assessed. When metric invariance is established, it means that factor loadings before standardization can be constrained to be the same for both group. Results (model 2 in Table 6.6) suggested that the model with constrained factor loading also fitted the data well. The CFI was below .95, but .94 still indicates an acceptable fit. Other fitting indices all suggested good model fit. In addition, all the change in model fit statistics indicated that the more constrained model was not significantly worse than the less constrained model. Although the delta CFI was -.014, which is marginally smaller than -.01, when all the indices were considered together, there appears to be enough evidence to support a metric invariance.

The next step assessed scalar invariance, which tested whether the indicator intercepts can be constrained to be the same across both groups. Results showed that this model (model 3 in Table 6.6) did not fit the data well. Both CFI (.84) and SRMR (.081) fell below the pre-specified criteria. Change in model fit statistics also suggested that this model was significantly worse than model 2. This indicated that at least one indicator intercept could not be equivalent in two groups. Examining the test score for each group, it was identified that the Colour Stroop task had the largest discrepancy across groups. A partial invariance model was then tested, where the intercepts for the Colour Stroop task were not constrained to be equal in two groups. The revised partial invariance model (3a) provided a good fit to the data, although change in CFI index was -.038, which is lower than -.01. Nevertheless, considering all fitting indices and change in statistics, there was enough evidence to support this partial invariance model, with the Colour Stroop task having different intercepts across the two groups and other indicators having the same factor loadings and intercepts. This measurement invariance is acceptable as only 1 out of 9 indicators (11.1%) failed intercept invariance. In summary, results suggested that measurement invariance existed for monolinguals and bilinguals.

Factor mean invariance was assessed because strong measurement invariance was established. In model 4 (Table 6.6), when the factor means were constrained to be the same in two groups, all fitting indices suggested a bad fit. Change in fit statistics also suggested that this model was significantly worse than the partial scalar invariance model. This suggested that factor means had to differ for monolinguals and bilinguals in order to fit the data. To examine which factor mean differed, a revised model with partial factor mean invariance was tested. Model 4a tested a model where the common EC was constrained to be equal and the shifting factor means were allowed to vary across groups. This model provided a just acceptable model fit, with CFI = .90. In addition, change in fitting statistics all indicated that this model was not significantly worse than the partial scalar invariance model (3a). Thus, it appears that monolinguals and bilinguals had a similar common EC factor mean. Model 4b tested an alternative model where the common EC factor means were allowed to vary, but the shifting factor mean was constrained to be equal across two groups. This model did not fit the data well, with CFI = .88. In addition, change in fit indices showed a marginal effect, $\Delta\chi^2$ (1)

= 3.2, p = .074. When all indices were considered together, it appeared that constraining the shifting factor mean to be equal resulted in a bad fit and a worse model compared with the model 3a. Therefore model 4a was accepted. Interestingly, in this model bilingual speakers scored .21 units higher than monolingual speakers on the shifting factor.

Last but not least, factor covariance invariance was tested (Model 5). Results showed that this model fit the data well apart from SRMR = .081. However all other fitting indices and change in fit statistics showed that this model provided an acceptable fit and was no worse than Model 4a. Thus, the factor covariance invariance model was accepted, with the factors correlating with each other in the same way across two groups.

To conclude, based on the two-factor model for executive control in Figure 6.12, we examined the invariance in the mean and covariance structure between monolinguals and bilinguals. Model 5 was identified which fitted the data of both groups well with equivalent constraints. In this model, the factor means for the common EC as well as the factor covariance were constraint to be equivalent for two groups. Factor means for Shifting could not be constraint, and bilingual speakers showed a larger mean for the Shifting factor.

Discussion

This chapter described a comprehensive study that surveyed the performance of monolingual and bilingual speakers on a series of executive control tasks, tapping into updating, inhibition and shifting abilities. Confirmatory factor analysis was conducted to compare the structure of the executive control network in monolingual and bilingual speakers. We first discuss the results from single task analyses, then move on to CFA model identification and results from the invariance analysis.

Single Task Analyses

Results from single task analyses were largely mixed, similar to what we have seen in the literature. In terms of updating ability, two updating tasks (the Letter Track and the Keep Track task) showed no group difference, whereas the Space Two-Back showed a monolingual advantage. Individual differences studies had revealed that bilingual language experience did not predict performance in N-back tasks (Soveri et al. 2011; Yow & Li, 2015). The absent of a bilingualism advantage is thus unsurprising. However, it is unclear what gave rise to the monolingual advantage. Further studies will need to address this issue.

Results were similarly mixed for tasks that tap into inhibition ability. Monolinguals showed superior performance in the Colour Stroop task and the Stop Signal task, whereas the Anti-saccade task showed no group difference. Contrary to most previous findings of bilingual advantage (Bialystok, Craik, & Luk, 2008) or no group differences (Dunabeitia et al., 2014; Kousaie & Phillips, 2012; Rosselli et al., 2002), we seem to be the first to have found a monolingual advantage in the Colour Stroop effect. The overall speed advantage of monolinguals in the Colour Stroop task and the Category Switch task might be explained by them being native speakers. This task is an overt naming task and there has been evidence that bilinguals are slower in naming pictures than monolinguals (Gollan et al., 2005) due to, for instance, slower lexical access. The larger Colour Stroop effect in bilinguals might be due to the activation of the other language: while the translation equivalent in the non-target language needs to be suppressed in both the incongruent and the neutral condition, the incongruent condition activates an additional competitor word in the non-target language leading to a larger competition for bilingual than monolingual participants. Previous studies have compared an incongruent with a congruent, instead of a neutral, condition. The interference from the second language is arguably similar for congruent and incongruent conditions, explaining why the bilingual disadvantage has not been reported previously.

We seem to be also the first to have found a monolingual advantage in the Stop Signal task. Previous studies have reported no group differences (Colzato et al., 2008; Morales et al., 2013; Rodriguez-Pujadas et al., 2014). The diverging results in this task might be due to the nature of the stimuli used. Previous studies all used non-verbal stimuli, while the current study used English words as stimuli. Consistent with earlier studies where bilinguals were found to have inferior performance on verbal tasks, our result suggests a bilingual handicap in a verbal task (Darcy, 1953).

Results for tasks that tap into shifting were more consistent, especially if one considers the effect of verbal stimuli. Monolinguals were faster and showed less of a RT shifting effect in the Category Switch task, again a task that used English words as stimuli. RT results for the other two tasks showed compatible performances for two groups. This strengthens the argument that the worse bilingual performance in the Category Switch task was, due to the verbal nature of the task. Accuracy results, by contrast, were more consistent. Bilinguals showed reduced switching effect in all three tasks, significantly so in the two non-verbal switching tasks, and descriptively so in the verbal switching task.

Taking all results from the single task analyses together, it is difficult to strongly conclude whether there is a bilingualism effect or not. The situation was further complicated by the verbal tasks, on which monolinguals tended to outperform bilinguals, making it impossible to conclude whether this difference was due to enhanced monolingual executive control abilities or purely due to their better verbal abilities. This is a perfect demonstration of the task impurity problem that we have introduced. We next discuss results from the confirmatory factor analyses, which overcame this problem.

Model Identification

A two-factor model was identified as the model that best fit the empirical data, with all inhibition tasks and updating tasks loading onto one factor and all shifting tasks loading onto another factor. Factors were allowed to co-vary. There are two differences between our best fitting model and what has been reported in the literature. First, there were only two latent factors instead of three as reported in the seminal study by Miyake et al. (2000), with tasks that tap inhibition and updating loading onto a common factor. Second, there were two sets of residual covariances that had to be assumed in order to achieve a good fit.

Our two factor model was not significantly worse than a full three factor model. This suggests that the inhibition function and the updating function shares features that are not separable. In contrast, in the three factor model by Miyake et al. (2000), the confidence interval for the correlation between the updating and inhibition factors was [.30. 96]. Because it did not include 1, the authors argued that the two factors were not perfectly correlated and thus should not be collapsed into a single factor. When comparing a three factor model with models that combined two of the factors (see Table 6.2 in Miyake et al. 2000), fitting indices showed that such models provided good fit to the data, but they were all significantly worse than the full three factor model. The authors therefore argued that the full three factor model should be retained. This three-factor structure has been replicated in different populations, such as children (Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003), young adults (Friedman et al., 2006), and clinical populations (Willcutt et al., 2001). However this has not always been the case. In an elderly population, Fisk and Sharp (2004) identified an additional, fourth factor for long term memory access. St Clair-Thompson and Gathercole (2006) identified a reduced two-factor model that best fit data from a child population. Therefore, depending on the population and the tasks used, the model might contain two, three or even four factors. What is common in all models is that executive control consisted of inter-related sub-functions, showing both unity and diversity. The current findings are therefore not completely at odds with previous findings. First, although the confidence interval did not include 1 in Miyake et al. (2000), it was very close to 1. This means that in their model the relationship between updating and inhibition was stronger than for the other factor combinations (inhibition and shifting, updating and shifting). Second, consistent with our observations, Hull et al. (2008) reported for an elderly population that a two-factor model, where updating and inhibition were collapsed, fitted the data equally well as a full three factor model.

There are two possible causes for the discrepancy between our finding and the original one by Miyake et al. (2000). First, the current study found that inhibition and updating were correlated more strongly in the bilingual group than in the monolingual group. For monolinguals, inhibition and updating showed only moderate covariance, while this was estimated to be above 1 for bilinguals, which is not possible and a symptom for highly correlated factors. Miyake et al. (2000) might have tested rather monolingual participants, which can explain why Miyake et al. (2000) did not find a high correlation between the two factors. A second potential cause for the discrepancy is sample size. In a large-scale study by Friedman et al. (2008) with over 900 participants, unlike in Miyake et al. (2000), the model with inhibition and updating collapsed no longer fitted the data well. Therefore, it could also be that we did not have enough power in the current study to tease the two latent factors apart.

The second difference between our model and the original model by Miyake et al. (2000) were two sets of residual covariances that had to be assumed in order to achieve a good fit. The first pair was the Category Switch task and the Stop Signal task. These two tasks are similar in that stimuli were both words, and participants were required to make categorical judgements. Therefore, similar processes were involved, such as word recognition, concept retrieval and conceptual categorization. In addition, both tasks required animation judgements at a conceptual level. The residual covariance likely accounts for such similarities. The second pair is the Colour Stroop task and the Number-Letter task. It is not immediately clear how these two tasks share common processes, because stimuli were different and one task required manual responses, the other verbal responses. However, it has been suggested that there might be an inhibition component in shifting tasks in that participants need to deactivate an old mental set and switch to a new mental set. Therefore, the residual covariance might account for these shared processing requirements.

The question remains, however, how one can interpret the common EC factor, that is, a factor that captures what is common to all updating and inhibition tasks in our study. One potential common process is maintenance of task goals (Miyake & Friedman, 2012; Miyake et al., 2000). In all tasks, participants had to keep the task goal in mind and use it to inform their response choices. A second candidate process is inhibition. In the updating tasks, participants had to discard items that used to be the target but were no longer relevant. One could argue that this process entails some form of suppression (Miyake et al., 2008). Given this uncertainty, the term ‘common EC’ was used here to mean it entails a variety of functions, potentially goal maintenance and inhibition.

Model Invariance

An invariance analysis was carried out on the two-factor model in order to examine how the indicators measured the latent factors across groups. Strong partial measurement invariance was established between monolingual and bilingual speakers. This suggests that both groups have the same general factor structure (i.e. both can be described using the same two factor model) and that indicators measure the latent factors in the same way across groups. A partial scalar invariance was established after freeing the intercepts for the Colour Stroop task. This means that apart from the Colour Stroop task, all the tasks had the same intercepts across groups. The result regarding the Colour Stroop task is consistent with the single task analysis, which showed that monolingual speakers outperformed bilingual speakers. Therefore, regarding the first aim of the study, our results suggest that the outline of executive control is the same for monolingual and bilingual speakers.

Structural invariance analysis demonstrated that the latent factor mean for shifting was not the same across groups, whereas the mean for the latent common EC factor was the same. This result is consistent with Soveri et al.'s (2011) finding using an individual difference approach, namely that bilingualism affects the shifting task performance but not performance in inhibition and updating. The latent factor mean for shifting was higher for bilinguals than for monolinguals, consistent with findings of enhanced bilingual shifting ability (Christoffels et al., 2015; Marzecova et al., 2013; Prior & MacWhinney, 2010). As introduced, using a factor analysis approach, we extracted what is common across tasks and therefore gained a purer measure of the cognitive constructs in question. Therefore, we have shown for the first time at a latent factor level that bilingual speakers have developed enhanced shifting ability. This observation corresponds well with the bilingual experience. As a daily exercise, bilingualism entails shifting back and forth between languages. And indeed, most of the bilingual participants that took part in our study often switch between languages. 70% of the bilingual speakers reported using both languages in the same conversation, 28% reported this occurred rarely and only 2% reported never to use both languages in a conversation.

We have discussed above that the common EC factor could be described as goal maintenance, attentional control or inhibition. Results of invariance analyses showed that the two groups do not differ in this general executive control factor. However, because the common EC is too broadly and vaguely defined, it is possible that it has absorbed much of the variances caused by different factors and effects might have even cancelled each other out. For instance, if the common EC factor absorbs variances due to language abilities (e.g. bilinguals performed worse in the language based executive control tasks such as the Stroop task and the Stop Signal task), then this disadvantage might potentially mask any advantages in attentional control. In other words, the aggregated outcome being the same across participant groups does not rule out the possibility that group differences exist if we were able to break the common EC down into smaller components. But the high correlation between the latent updating and inhibition factors in bilingual participants did not allow us to compare participant groups at such sub-components. Therefore, there is not adequate evidence in the current study to conclude that bilingualism did not result in a general advantage or disadvantage in inhibition, updating, general attentional control or monitoring.

With regard to the last aim of the study, we did not find a statistically significant difference between groups regarding the covariance among latent factors. This means that the components of the executive control network were related to each other in largely the same way for monolinguals and bilinguals. Nevertheless, descriptively individual components had higher covariance scores for bilinguals than for monolinguals (.53 for bilinguals and .34 for monolinguals). In addition, the correlations between all pairs of the three sub-functions were descriptively higher for bilinguals than for monolinguals (see Figure 6.11). In fact, the three factor model provided good fit to the monolingual data, while the model was problematic for the bilinguals due to the correlation between inhibition and updating being more than 1.These results suggested that the three sub-functions were more independent for monolinguals than for bilinguals. This is consistent with the claim that bilinguals have developed cognitive networks with a stronger connectivity (Grady et al., 2015). Garcia-Penton, Fernandez, Iturria-Medina, Gillon-Dowens, and Carreiras (2014) reported two structural sub-networks that were more connected by white matter (WM) tracts in bilinguals than in monolinguals. One of them included the medial superior frontal gyrus (MSF), which is related to executive functions.

Lack of statistically significant differences for covariance among latent factors between the two participant groups in our study could be due to the effects being too subtle and the study not having enough power.

Limitations

The first limitation of the study is that the sample size was at the lower limit of conducting a confirmatory factor analysis. This might have limited the power of detecting any group difference regarding latent factor means.

The second limitation is that many tasks used in this experiment have a verbal component. A lot of the tasks used verbal stimuli, namely the Category Switch task, the Colour Stroop task, the Stop Signal task and the Keep Track task. Other tasks also used linguistic elements such as the Letter Track task and the Number/Letter task. Bilinguals tend to have worse performance in verbal tasks, such as naming speed (Ivanova & Costa, 2008) or verbal fluency (Sandoval, Gollan, Ferreira, & Salmon, 2010). If the majority of tasks have a verbal component, then variances caused by verbal abilities might be absorbed by the latent factor rather than being left in the unique variance of the tasks. In the present study, four out of six of the tasks that loaded onto the common EC factor had some verbal or linguistic component. This has added noise to the factor analysis in that it might have reduced any chance of observing a group difference in the common EC factor, if the verbal ability had pointed to a bilingual disadvantage and the common EC had pointed to a bilingual advantage. We employed the original tasks without much modification in order to be able to compare our results with those of previous studies. Future studies should carefully select non-linguistic tasks to avoid the same complications.

Conclusion

This is the first study that compared monolingual and bilingual executive control networks. Instead of comparing speakers' performance on a single task, or focusing on a single executive function, this study investigated the network as a whole. A factor analysis approach allowed us to compare groups on purer measures of executive functions. The confirmatory factor analysis identified a two-factor model for executive function: common EC and shifting. Measurement invariance analysis showed that the two participant groups had the same factor structure and the same factor loadings. Most interestingly, the latent factor mean for the shifting ability was higher for bilingual speakers than for monolingual speakers, suggesting enhanced shifting ability. No difference was found for the common EC factor means.

CHAPTER 7

GENERAL DISCUSSION AND FUTURE DIRECTIONS

For researchers striving to understand the cognitive impact of bilingualism, it is a controversial time. It is challenging, because inconsistent findings reveal that the relationship between bilingualism and its cognitive impact is a complex one, making it difficult to agree on how bilingualism affects speakers' general cognitive abilities. If the claims about bilingual cognitive advantage were driven by publication bias, then it calls into question whether this line of research is warranted. However, we are also experiencing a promising period in this field, since discussions and debates are prompting deeper investigations. This thesis contributes to our understanding of the effect of bilingualism on speakers' executive control from three aspects: bilingualism and executive control in non-verbal tasks, bilingualism and executive control in verbal tasks, and bilingualism and the executive control network.

Bilingualism and Executive Control in Non-verbal Tasks

This work began by addressing inconsistent findings in the literature on bilingual cognitive advantages. In Chapter 2, we reviewed studies on the effect of bilingualism on non-verbal interference task performance from a novel perspective. We proposed that one reason for observing or not observing bilingualism advantages is whether data is trimmed or not. The results of the comprehensive review showed that in studies with shorter response time allowances, bilingualism advantages were less likely to be observed. This finding has at least two critical implications. First, it implies that the effect of bilingualism resides at least partially in long responses. Second, it promotes the usage of more fine-grid analytical approaches for revealing subtle differences between monolingual and bilingual performances.

Empirical evidence for the effect of bilingualism in Chapter 3 was gained by comparing a group of monolingual English speakers and a group of bilingual English/Chinese speakers across three non-verbal interference tasks. Using insights gained from Chapter 2, we conducted ex-Gaussian analyses of response time distributions and attempted to distinguish the relative contribution of different executive controls to the bilingualism advantage. Two aspects of executive control were examined: inhibitory control and attentional control. While the main body of the response distributions were similar across the two groups, the extent of the tails differed significantly, with bilingual speakers having shorter tails in all three tasks, regardless of the experimental condition. These results suggest that bilingual speakers have developed enhanced attentional control. Evidence for enhanced inhibitory control, which would have been indicated by a reduction in the magnitude of the congruency effect in the body of the response distribution, was not found. In conclusion, the evidence suggests that bilingualism enhances attentional control but not inhibitory control.

Bilingualism and Executive Control in Verbal Tasks

Chapters 4 and 5 addressed the issue of whether bilingual speakers' enhanced executive control affects their resolution of intra-language lexical competition. Under the assumption that the management of more than one language cultivates some domain-general executive control that benefits speakers' performance in the non-verbal domain, the bilingualism advantage hypothesis has been extensively tested with non-verbal tasks. There is a dearth of literature on whether such an enhanced executive control also facilitates verbal control, such as lexical selection within a single language in a highly competitive context. We tested this hypothesis with a semantic blocking paradigm (Chapter 4). Results revealed that response distribution profiles for monolingual and bilingual speakers were differentially affected by competitive semantic context. Results demonstrated a blocking effect for the main body of the responses for monolinguals only. This means that the majority of responses of monolingual speakers in the homogeneous condition were slowed due to strong intra-language lexical competition, and that this was not the case for bilingual speakers. In contrast, a blocking effect for the response distribution tails was found for bilinguals, but not for monolinguals. In other words, strong lexical competition appears to result in more extreme responses within the bilinguals. We argued that these results reflect the involvement of executive control in the bilingual groups, which facilitated their resolution of lexical competition through inhibiting alternative items. As a result, recovering from an internal error takes longer.

Chapter 5 followed up the findings of Chapter 4, seeking electrophysiological evidences for the engagement of executive control when bilinguals were performing the semantic blocking task. Behavioural results partially replicated the finding in Chapter 4: monolinguals who showed a marginal semantic blocking effect in the main body of the responses demonstrated consistent outcomes, whereas bilinguals did not show statistically significant effects, either in the body or the tail of the distributions. This is plausibly due to the heterogeneity of the participant group. In contrast to the small behavioural difference between monolinguals and bilinguals, EEG results showed divergent outcomes for the two groups. While both groups showed a parietal-posterior semantic blocking effect with a peak at 220 ms, this effect lasted longer for monolinguals. Interestingly, bilingual speakers showed an additional, early central-frontal effect beginning at approximately 160 ms. Thus, both groups seem to suffer from more effortful lexical selection in the homogeneous condition than in the heterogeneous condition, but bilinguals seem to possess a more engaged top-down executive control that interacts with bilingual lexical selection. Furthermore, only monolinguals demonstrated an additional late posterior effect, at approximately 400 – 450 ms. This suggests stronger response-monitoring in the homogeneous condition, in which alternatives were more active than in the heterogeneous condition. This effect was numerically present, but statistically not significant in the bilingual group, possibly because for a large of number of bilinguals the lexical competition had been sufficiently resolved in early processing stages, meaning that later monitoring demand was not significantly affected. In conclusion, there is converging electrophysiological evidence to support the stronger involvement of executive control in the resolution of intra-language lexical competition in the bilingual group, although differences in neural activities might not necessarily lead to strong differences at the behavioural level.

Bilingualism and Executive Control Network

In Chapter 6, the effect of bilingualism on executive control was investigated from a different level. Instead of investigating the impact of bilingualism on an individual facet of executive control, this issue was addressed from a holistic view, by targeting the executive control network as a whole. In order to do that, a confirmatory factor analysis approach was taken. This approach has several advantages. First, by statistically extracting the common components from several similar tasks, one obtains purer measurements of executive control subcomponents. Second, this approach allowed one to examine the relationship amongst different executive sub-functions. Simple task analysis showed mixed results. Bilinguals demonstrated advantages in switching tasks in terms of response accuracies, but they did not show any advantages in inhibition or updating. On the contrary, bilingual speakers performed worse than monolingual speakers on some tasks. Closer inspection revealed that bilingual speakers were disadvantaged in tasks that involved verbal stimuli. Their inferior performance might thus be due to verbal ability rather than executive control. Confirmatory factor analysis (CFA) suggested a slightly different picture. First, a two-factor model was identified as the best fit for the observed data. That is, the executive control network was best described as a general executive control factor and a shifting factor, which correlated with each other. Importantly, this model fitted the data of monolinguals and bilinguals equally well, suggesting that executive control is organised similarly in the two groups. An invariance analysis revealed group differences. Particularly, an invariance analysis of latent factor means revealed that bilingual speakers had better shifting ability, i.e. were better able to shift cognitive sets than monolingual speakers. The groups did not differ regarding the 'general executive control' component of the model. Lastly, descriptive statistics revealed a more correlated network for bilingual speakers than for monolingual speakers, suggesting that bilingualism might have affected the organization of the executive control network, enhancing its coordination or connectivity. Lack of statistical significance of this difference might be due to the limited power of this study.

Limitations

The way that levels of variables are classified necessarily restricts the research questions that may be posed and the possible answers that may be given (Bialystok, 2001). In the Introduction, bilingualism was defined within the scope of this thesis. Throughout this work, a binary distinction was made between monolingualism and bilingualism. This means that individual differences in monolinguals and bilinguals were largely unconsidered. Bilingualism is in fact better described as a continuum than as categorical variable (Luk & Bialystok, 2013). The bilingual population tested in the current series of studies is taken from one end of the continuum, i.e. speakers whose proficiency in the second language is fairly high and who started to acquire the second language relatively early in life. This raises the question to what extent the current findings may be applied to bilingual speakers with various linguistic backgrounds. Do early (proficient) bilinguals differ from other late or non-proficient bilinguals? If they do differ, is this difference quantitative or qualitative? A quantitative difference would predict the bilingual advantage to increase with the accumulation of their language experience. If the difference is qualitative, then the findings from this line of research may not be directly applicable to bilinguals with different language backgrounds.

Lastly, this work focused on young adults, who are at the apex of their cognitive abilities. As discussed in Chapter 2, findings about the cognitive impact of bilingualism in children, young adults and the elderly diverge largely. Therefore, it is unclear to what extent these findings can be generalised to different age groups.

Towards a Better Understanding of the Bilingualism Effect

Work detailed in this thesis provides converging evidence that bilingualism affects cognitive control. However, this effect is sometimes subtle and weak, and it might be well masked by noise. In order to better understand the cognitive consequences of bilingualism, a more fine-grid examination of performances and a more nuanced evaluation of the evidence are required. In the following section, directions for future research are suggested. First, expanding the scope to examine a wider bilingual population might be insightful. Second, while allowing variations on one dimension, other dimensions of language experience should be more strictly controlled in order to clarify results. Third, recruiting the most appropriate population and applying the appropriate level of task demand are prerequisites for accurate observation of group differences. Last but not least, detailed analysis of results might provide richer information.

Expansion of the Scope

Language experience is an extremely diverse variable. I have defined the scope of the examination of bilingualism by focusing on early bilinguals (Age of Acquisition before the age of ten years) who speak a second language on a daily basis and who have high proficiency in the second language. It is immediately evident that many factors potentially interact with bilingualism, thus accounting for its cognitive consequences. These include age of acquisition, active use of both languages, level of proficiency, and so on. Given the diversity of bilingual speakers, it is perhaps unsurprising that findings on bilingual executive function are inconsistent. In order to fully understand the effect of speaking more than one language on one's general cognitive ability, we need to expand the scope of our analysis and sample a wider range of the bilingual population.

It is important that the analytical programs are designed according to the sample population. More precisely, when the population is so diverse that a categorical analysis becomes no longer appropriate, an individual differences approach might be adopted instead. Take for instance a study by Paap and Greenberg (2013), in which participants with various degrees of second language knowledge were dichotomised as monolingual and bilingual speakers according to self-ratings on language proficiency. Participants who rated themselves as advanced, intermediate or higher were classified as bilingual speakers; those who rated themselves as intermediate or lower were classified as monolingual. This means that the groups were very close to each other on the continuum. It could be argued that the lack of group differences in this study was due to the adjacency of the two groups on the language proficiency continuum. In this case, an individual differences approach might have been more beneficial. This approach has recently been successful, revealing that many factors, other than bilingualism *per se,* modulate performance on executive control tasks, including age of second language acquisition (Yow & Li, 2015), language proficiency of both first and second language (Singh & Mishra, 2015; Tse & Altarriba, 2012; 2015), and the relative dominance of languages (Goral, Campanelli, & Spiro, 2015).

Strict Control of Variables

Strict control of participants' language background has proven critical since early on. Peal and Lampart (1962) were the first who strictly controlled the socio-economic background and proficiency level of their bilingual group. The authors reported superior performances for bilingual speakers, in contrast to almost all previous findings.

In this thesis, different types of bilinguals were recruited, such as English/Chinese bilinguals and those speaking English and another language. Using the same experimental paradigm but different bilingual populations as in Chapter 4, Chapter 5 did not fully replicate the behavioural outcomes. Similarly, using the same experimental setting, a group of English/Chinese bilingual speakers has been shown to out-perform a group of English/French bilingual speakers who did not differ from monolinguals (Bialystok et al., 2005). Such evidence suggests the influence of cross-linguistic differences. Indeed, a recent study revealed a language-specific effect of bilingualism on the ability to resolve conflict (Coderre & van Heuven, 2014). The authors suggested that similar languages activate each other to a greater degree, leading to a higher requirement for cognitive control. Thus similar languages promote executive control to a larger degree. The mechanism of cross-linguistic influence is beyond the scope of this discussion, but it is suggested that more insight would be gained by testing more homogeneous bilinguals, and that this line of research is warranted.

It should be acknowledged that, when comparing monolingual and bilingual speakers, it is almost impossible to control for every relevant factor. Therefore, one way to circumvent possible complications caused by the mismatch of populations would be to look for evidence within a group. For example, Wu and Thierry (2013) showed that a simple change of language context affected bilingual executive control performance. Specifically speakers' performance was enhanced after being exposed to a mixed language context than to a single language context. Therefore, study designs like this can provide direct evidence with regards to how engagement of different language activities can modulate executive control abilities.

Careful Design of Experiment

The strength of the bilingualism advantage is not commensurate across the lifespan (Bialystok et al., 2004). It is most prominent in the elderly, and less so in children and young adults (Hilchey & Klein, 2011). One explanation for this is that young adults are at the peak of their cognitive performance, and that tasks that are relatively easy do not reveal group differences due to a ceiling effect. Consistent with this idea, Bialystok (2006) reported no group difference between monolingual and bilingual speakers in a moderate response switch condition, but a bilingual advantage only under a high response switch condition. Costa et al. (2009) manipulated the proportion of incongruent trials in an ANT task and reported bilingual advantage in a medium (50% incongruent) and a high (75% incongruent) monitoring condition, but not in a condition with low (25%) monitoring demand. Participants tested in studies of this thesis were young adults. In order to investigate the effect of bilingualism, one needs to devise a sensitive measure and set the appropriate level of task demand.

Alternative Analytical Approaches

Data always come with noise. As works in this thesis has shown, in order to better understand the signal, alternative analytical approaches can be fruitful. First, detailed inspection of reaction time data, such as analysing response time distributions, might provide further insights into underlying cognitive processes. In this thesis, an ex-Gaussian response distribution analysis had been used to analyse reaction times in non-verbal executive control tasks (Chapter 3), as well as verbal speech production tasks (Chapter 4). In both cases, this analysis provided information that was not available from a traditional central tendency analysis.

Second, Chapter 6 demonstrated the potential usefulness of more sophisticated statistical methods. This is particularly useful when investigating a certain aspect of cognitive processing, for example, inhibitory control or attentional control. The task impurity problem exists because tasks conceivably draw upon many different levels of processing. It is therefore difficult to isolate the contribution of a certain processing of interest using a single task. One way to circumvent this problem is by taking a factor analysis approach, as in Chapter 6. This requires taking a few measures of the same target cognitive processing and statistically extracting the common component.

Beyond Behavioural Data

This study moved beyond the investigation of behavioural responses. In Chapter 5, the two language groups showed no group difference in the picture-naming task, but they revealed divergent brain responses. There are many cases in the literature where a brain response difference was observed despite showing identical behavioural outcomes. For example, in an ERP study using a modified version of a continuous performance task, Morales, Yudes, Gomez-Ariza, and Bajo (2015) showed that bilingual groups showed different N2 and P3a components, reflecting different conflict detection and response suppression processes in bilinguals and monolinguals. But the two groups did not differ in their behavioural measurements. Similarly, in an MEG study investigating performance on a Simon task, Bialystok, Craik, et al. (2005) reported that different neural networks were involved in a group of English/French bilingual speakers, compared with a monolingual English group, although the two groups demonstrated identical average speed and response accuracy. Therefore, operational processing alternatives might not necessarily lead to functional differences. Going beyond the surface level will help in gaining a better understanding of the processing differences between monolingual and bilingual speakers.

Conclusion

The present thesis has accumulated empirical evidence that bilingualism does affect executive control. These effects were visible in the non-verbal domain, verbal domain, as well as when the executive control network was investigated as a whole. Due to the clearly defined scope of the target population of bilinguals, the exact generalisability of findings in the present thesis to other populations remains to be investigated. Due to the various facets of the bilingualism phenomenon, further research should investigate how certain aspects of the bilingual experience affect cognitive control, with more fine-grained analytical procedures.

LIST OF REFERENCES

Aarts, E., Roelofs, A., & van Turennout, M. (2009). Attentional control of task and response in lateral and medial frontal cortex: Brain activity and reaction time distributions. *Neuropsychologia, 47*(10), 2089-2099. doi: 10.1016/j.neuropsychologia.2009.03.019

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. *Language and Cognitive Processes, 24*(5), 713-734.

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex, 49*(3), 905-911. doi: 10.1016/j.cortex.2012.08.018

Acheson, D. J., Ganushchak, L. Y., Christoffels, I. K., & Hagoort, P. (2012). Conflict monitoring in speech production: Physiological evidence from bilingual picture naming. *Brain and Language, 123*(2), 131-136. doi: 10.1016/j.bandl.2012.08.008

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A Systematic Review and Meta-Analysis of the Cognitive Correlates of Bilingualism. *Review of Educational Research, 80*(2), 207-245. doi: 10.3102/0034654310368803

Alladi, S., Bak, T. H., Duggirala, V., Surampudi, B., Shailaja, M., Shukla, A. K., . . . Kaul, S. (2013). Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology, 81*(22), 1938-1944.

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology-Learning Memory and Cognition, 27*(2), 514-540. doi: 10.1037//0278-7393.27.2.514

Aristei, S., Melinger, A., & Rahman, R. A. (2011). Electrophysiological Chronometry of Semantic Context Effects in Language Production. *Journal of Cognitive Neuroscience, 23*(7), 1567-1586. doi: 10.1162/jocn.2010.21474

Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 49*(1), 5-28.

Bak, T. H., Vega-Mendoza, M., & Sorace, A. (2014). Never too late? An advantage on tests of auditory attention extends to late bilinguals. *Frontiers in Psychology, 5*, 6. doi: 10.3389/fpsyg.2014.00485

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology-General, 128*(1), 32-55. doi: 10.1037//0096-3445.128.1.32

Balota, D. A., & Yap, M. J. (2011). Moving Beyond the Mean in Studies of Mental Chronometry: The Power of Response Time Distributional Analyses. *Current Directions in Psychological Science, 20*(3), 160-166. doi: 10.1177/0963721411408885

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59*(4), 495-523. doi: 10.1016/j.jml.2007.10.004

Belke, E. (2008). Effects of working memory load on lexical-semantic encoding in language production. *Psychonomic Bulletin & Review, 15*(2), 357-363. doi: 10.3758/pbr.15.2.357

Belke, E. (2013). Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *Journal of Memory and Language, 69*(3), 228-256. doi: 10.1016/j.jml.2013.05.008

Belke, E., Brysbaert, M., Meyer, A. S., & Ghyselinck, M. (2005). Age of acquisition effects in picture naming: Evidence for a lexical-semantic competition hypothesis. *Cognition, 96*(2), B45-B54.

Belke, E., Meyer, A. S., & Damian, M. F. (2005). Refractory effects in picture naming as assessed in a semantic blocking paradigm. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 58A*(4), 667-692.

Belke, E., & Stielow, A. (2013). Cumulative and non-cumulative semantic interference in object naming: Evidence from blocked and continuous manipulations of semantic context. *The Quarterly Journal of Experimental Psychology*, 1-26. doi: doi: 10.1080/17470218.2013.775318

Bialystok, E. (1992). Attentional Control in Childrens Metalinguistic Performance and Measures of Field Independence. *Developmental Psychology, 28*(4), 654-664. doi: 10.1037/0012-1649.28.4.654

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* : Cambridge, UK: Cambridge University Press.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Canadian Journal of Experimental Psychology-Revue Canadienne de Psychologie Experimentale, 60*(1), 68-79.

Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *NeuroImage, 24*(1), 40-49.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*(3), 89-129.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290-303.

Bialystok, E., Craik, F. I. M., & Luk, G. (2008). Lexical access in bilinguals: Effects of vocabulary size and executive control. *Journal of Neurolinguistics, 21*(6), 522-538.

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(4), 859-873.

Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences, 16*(4), 240-250.

Bialystok, E., Craik, F. I. M., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459-464. doi: 10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I. M., Grady, C., Chau, W., Ishii, R., Gunji, A., & Pantev, C. (2005). Effect of bilingualism on cognitive control in the Simon task: evidence from MEG. *NeuroImage, 24*(1), 40-49.

Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*(3), 89-129.

Bialystok, E., Craik, F. I. M., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. *Psychology and Aging, 19*(2), 290-303.

Bialystok, E., & DePape, A.-M. (2009). Musical Expertise, Bilingualism, and Executive Functioning. *Journal of Experimental Psychology-Human Perception and Performance, 35*(2), 565-574. doi: 10.1037/a0012735

Bialystok, E., Kroll, J. F., Green, D. W., MacWhinney, B., & Craik, F. I. M. (2015). Publication Bias and the Validity of Evidence: What's the Connection? *Psychological Science, 26*(6), 944-946. doi: 10.1177/0956797615573759

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. J. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism-Language and Cognition, 13*(4), 525-531. doi: 10.1017/s1366728909990423

Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism, 9*(1), 103-119. doi: 10.1177/13670069050090010701

Bialystok, E., Peets, K. F., & Moreno, S. (2014). Producing bilinguals through immersion education: Development of metalinguistic awareness. *Applied Psycholinguistics*, 35(1), 177–191. doi:10.1017/S0142716412000288

BijeljacBabic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition, 25*(4), 447-457.

Blanco, M. J., & Alvarez, A. A. (1994). Psychometric Intelligence and Visual focused Attention - Relationships in Nonsearch Tasks. *Intelligence*, *18*(1), 77-106. doi: 10.1016/0160-2896(94)90021-3

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language, 48*(3), 468-488. doi: 10.1016/s0749-596x(02)00503-x

Blom, E., Kuntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish-Dutch children. *Journal of Experimental Child Psychology, 128*, 105-119. doi: 10.1016/j.jecp.2014.06.007

Blumenfeld, H. K., & Marian, V. (2014). Cognitive control in bilinguals: Advantages in Stimulus-Stimulus inhibition. *Bilingualism-Language and Cognition, 17*(3), 610-629. doi: 10.1017/s1366728913000564

Botvinick, M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652. doi: 10.1037//0033-295x.108.3.624

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*(6758), 179-181. doi: 10.1038/46035

Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior Research Methods Instruments & Computers, 35*(4), 485-492. doi: 10.3758/bf03195527

Calabria, M., Hernandez, M., Martin, C. D., & Costa, A. (2011). When the tail counts: the advantage of bilingualism through the ex-Gaussian distribution analysis. *Frontiers in Psychology, 2*, 8. doi: 10.3389/fpsyg.2011.00250

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science, 11*(2), 282-298. doi: 10.1111/j.1467-7687.2008.00675.x

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling-a Multidisciplinary Journal, 14*(3), 464-504.

Cheng, X. R., Schafer, G., & Akyurek, E. G. (2010). Name agreement in picture naming: An ERP study. *International Journal of Psychophysiology, 76*(3), 130-141. doi: 10.1016/j.ijpsycho.2010.03.003

Christoffels, I. K., de Haan, A. M., Steenbergen, L., van den Wildenberg, W. P. M., & Colzato, L. S. (2015). Two is better than one: bilingual education promotes the flexible mind. *Psychological Research-Psychologische Forschung, 79*(3), 371-379. doi: 10.1007/s00426-014-0575-3

Coderre, E. L., & van Heuven, W. J. B. (2014). The effect of script similarity on executive control in bilinguals. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01070

Coderre, E. L., van Heuven, W. J. B., & Conklin, K. (2013). The timing and magnitude of Stroop interference and facilitation in monolinguals and bilinguals. *Bilingualism: Language and Cognition, 16*(Special Issue 02), 420-441.

Collins, A., & Loftus, E. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407-428. doi: citeulike-article-id:2976359

Colome, À. (2001). Lexical Activation in Bilinguals' Speech Production: Language-Specific or Language-Independent? *Journal of Memory and Language, 45*(4), 721-736.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *Journal of Experimental Psychology-Learning Memory and Cognition, 34*(2), 302-312.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*(5), 1283-1296. doi: 10.1037/0278-7393.26.5.1283

Costa, A., Hernandez, M., & Sebastian-Galles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106*(1), 59-86.

Costa, A., Hernandez, M., Costa-Faidella, J., & Sebastian-Galles, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition, 113*(2), 135-149.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences of the United States of America, 106*(50), 21442-21446. doi: 10.1073/pnas.0908921106

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition, 81*(3), B77-B86.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology, 2*, 10. doi: 10.3389/fpsyg.2011.00233

Darcy, N. T. (1953). A Review of the Literature on the Effects of Bilingualism upon the Measurement of Intelligence. *Journal of Genetic Psychology, 82*(1), 21-57.

de Bruin, A., Treccani, B., & Della Sala, S. (2015). Cognitive Advantage in Bilingualism: An Example of Publication Bias? *Psychological Science, 26*(1), 99-107. doi: 10.1177/0956797614557866

Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E., & Grainger, J. (2010). ERP evidence for ultra-fast semantic processing in the picture-word interference paradigm. *Frontiers in Psychology, 1*. doi: 10.3389/fpsyg.2010.00177

de Zubicaray, G., McMahon, K., Eastburn, M., & Pringle, A. (2006). Top-down influences on lexical selection during spoken word production: A 4T fMRI investigation of refractory effects in picture naming. *Human Brain Mapping, 27*(11), 864-873.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283-321. doi: 10.1037//0033-295x.93.3.283

Dunabeitia, J. A., Hernandez, J. A., Anton, E., Macizo, P., Estevez, A., Fuentes, L. J., & Carreiras, M. (2014). The Inhibitory Advantage in Bilingual Children Revisited Myth or Reality? *Experimental Psychology, 61*(3), 234-251. doi: 10.1027/1618-3169/a000243

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). British Picture Vocabulary Scale (2nd ed.).

Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The Source of Enhanced Cognitive Control in Bilinguals: Evidence From Bimodal Bilinguals. *Psychological Science*, *19*(12), 1201-1206. doi: 10.1111/j.1467-9280.2008.02224.x

Engel de Abreu, P., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism Enriches the Poor: Enhanced Cognitive Control in Low-Income Minority Children. *Psychological Science, 23*(11), 1364-1371. doi: 10.1177/0956797612443836

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., . . . Altaye, M. (2011). Evidence for Higher Reaction Time Variability for Children With ADHD on a Range of Cognitive Tasks Including Reward and Event Rate Manipulations. *Neuropsychology*, *25*(4), 427-441. doi: 10.1037/a0022155

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of Noise Letters Upon Identification of A Target Letter in A Nonsearch Task. *Perception & Psychophysics*, *16*(1), 143-149.

Fisk, J. E., & Sharp, C. A. (2004). Age-related impairment in executive functioning: Updating, inhibition, shifting, and access. *Journal of Clinical and Experimental Neuropsychology, 26*(7), 874-890. doi: 10.1080/13803390490510680

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172-179. doi: 10.1111/j.1467-9280.2006.01681.x

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology-General, 137*(2), 201-225.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45*(1), 152-170. doi: 10.1111/j.1469-8986.2007.00602.x

Ganushchak, L. Y., Christoffels, I. K., & Schiller, N. O. (2011). The use of electroencephalography in language production research: a review. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00208

Ganushchak, L. Y., & Schiller, N. O. (2008a). Brain error-monitoring activity is affected by semantic relatedness: An event-related brain potentials study. *Journal of Cognitive Neuroscience, 20*(5), 927-940. doi: 10.1162/jocn.2008.20514

Ganushchak, L. Y., & Schiller, N. O. (2008b). Motivation and semantic context affect brain error-monitoring activity: An event-related brain potentials study. *NeuroImage, 39*(1), 395-405. doi: 10.1016/j.neuroimage.2007.09.001

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Guasch, N. V., . . . Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and rnetalinguistic tasks. *Frontiers in Psychology, 5*, 14. doi: 10.3389/fpsyg.2014.00011

Garcia-Penton, L., Fernandez, A. P., Iturria-Medina, Y., Gillon-Dowens, M., & Carreiras, M. (2014). Anatomical connectivity changes in the bilingual brain. *NeuroImage, 84*, 495-504. doi: 10.1016/j.neuroimage.2013.08.064

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58*(3), 787-814. doi: 10.1016/j.jml.2007.07.001

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220-1234.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology, 16*(4), 562-576.

Gollan, T. H., Sandoval, T., & Salmon, D. P. (2011). Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. *Psychological Science, 22*(9), 1155-1164.

Goral, M., Campanelli, L., & Spiro, A. (2015). Language dominance and inhibition abilities in bilingual older adults. *Bilingualism-Language and Cognition, 18*(1), 79-89. doi: 10.1017/s1366728913000126

Grady, C. L., Luk, G., Craik, F. I. M., & Bialystok, E. (2015). Brain network activity in monolingual and bilingual older adults. *Neuropsychologia, 66,* 170-181. doi: 10.1016/j.neuropsychologia.2014.10.042

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the Use of Information - Strategic Control of Activation of Responses. *Journal of Experimental Psychology-General, 121*(4), 480-506. doi: 10.1037//0096-3445.121.4.480

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1*(02), 67-81.

Green, S., & Thompson, M. (2012). Investigating Measurement Invariance Using Confirmatory Factor Analysis. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 393-416). New York: Guilford Press.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711-1725. doi: 10.1111/j.1469-8986.2011.01273.x

Grosjean, F. (1982). *Life with Two Languages: An Introduction to Bilingualism*. Cambridge, Mass: Harvard University Press.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*(02), 131-149. doi: doi:10.1017/S136672899800025X

Grosjean, F. (2013). Bilingualism: A Short Introduction. In F. Grosjean & P. Li (Eds.), *The Psycholinguistics of Bilingualism* (pp. 1-21). Malden, MA & Oxford: Wiley-Blackwell.

Hakuta, K., & Diaz, R. (1985). The Relationship Between Degree of Bilingualism and Cognitive Ability: A Critical Discussion and Some New Longitudinal Data. In K. E. Nelson (Ed.), *Children's Language* (Vol. 5, pp. 321-344): Lawrence Erlbaum.

Heathcote, A., Popiel, S. J., & Mewhort, D. J. K. (1991). Analysis of Response - Time Distributions - An Example Using the Stroop Task. *Psychological Bulletin, 109*(2), 340-347. doi: 10.1037/0033-2909.109.2.340

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20*(5), 511-528. doi: 10.1037/0894-4105.20.5.511.supp

Hernandez, M., Costa, A., & Humphreys, G. W. (2012). Escaping capture: Bilingualism modulates distraction from working memory. *Cognition, 122*(1), 37-50. doi: 10.1016/j.cognition.2011.08.002

Hervey, A. S., Epstein, J. N., Curry, J. F., Tonev, S., Arnold, L. E., Conners, C. K., . . . Hechtman, L. (2006). Reaction time distribution analysis of neuropsychological performance in an ADHD sample. *Child Neuropsychology, 12*(2), 125-140. doi: 10.1080/09297040500499081

Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review, 18*(4), 625-658.

Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: Electrophysiological evidence for cross-language competition. *Brain Research, 1371*, 100-109. doi: 10.1016/j.brainres.2010.11.053

Hoshino, N., & Thierry, G. (2012). Do Spanish–English bilinguals have their fingers in two pies – or is it their toes? An electrophysiological investigation of semantic access in bilinguals. *Frontiers in Cognition.* **3**(9), 1-6.

Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: experimental and computational studies. *Cognition, 100*(3), 464-482.

Hu, L. T., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling-a Multidisciplinary Journal, 6*(1), 1-55. doi: 10.1080/10705519909540118

Hull, R., Hamilton, A. C., Martin, R. C., Beier, M. E., & Lane, D. (2008). Executive function in older adults: A structural equation modeling approach**.** *Neuropsychology, 22*(4), 508-522. doi: 10.1037/0894-4105.22.4.508

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00255

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1-2), 101-144. doi: 10.1016/j.cognition.2002.06.001

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*(2), 277-288.

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*(02), 169-178. doi: doi:10.1017/S1366728901000268

Janssen, N., Carreiras, M., & Barber, H. A. (2011). Electrophysiological effects of semantic context in picture and word naming. *NeuroImage, 57*(3), 1243-1250.

Kalia, V., Wilbourn, M. P., & Ghio, K. (2014). Better early or late? Examining the influence of age of exposure and language proficiency on executive function in early and late bilinguals. *Journal of Cognitive Psychology, 26*(7), 699-713. doi: 10.1080/20445911.2014.956748

Kan, I. P., & Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive Affective & Behavioral Neuroscience, 4*(1), 43-57. doi: 10.3758/cabn.4.1.43

Kapa, L. L., & Colombo, J. (2013). Attentional control in early and later bilingual children. *Cognitive Development, 28*(3), 233-246. doi: 10.1016/j.cogdev.2013.01.011

Kirk, N. W., Fiala, L., Scott-Brown, K., & Kempe, V. (2014). No evidence for reduced Simon cost in elderly bilinguals and bidialectals. *Journal of Cognitive Science,* 26(6), 640-648. doi: 10.1080/20445911.2014.929580.

Kousaie, S., & Phillips, N. A. (2012). Conflict monitoring and resolution: Are two languages better than one? Evidence from reaction time and event-related brain potentials. *Brain Research, 1446*, 71-90. doi: 10.1016/j.brainres.2012.01.052

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. *Frontiers in Behavioral Neuroscience, 8*. doi: 10.3389/fnbeh.2014.00250

Kramer, A. F., Bherer, L., Colcombe, S. J., Dong, W., & Greenough, W. T. (2004). Environmental influences on cognitive and brain plasticity during aging. *Journals of Gerontology Series a-Biological Sciences and Medical Sciences, 59*(9), 940-957.

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology, 25*(5), 497-514. doi: 10.1080/20445911.2013.799170

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica, 128*(3), 416-430. doi: 10.1016/j.actpsy.2008.02.001

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

Laganaro, M., & Perret, C. (2011). Comparing Electrophysiological Correlates of Word Production in Immediate and Delayed Naming Through the Analysis of Word Age of Acquisition Effects. *Brain Topography, 24*(1), 19-29. doi: 10.1007/s10548-010-0162-x

Laming, D. (1979). Choice reaction performance following an error. *Acta Psychologica, 43*(3), 199-224. doi: 10.1016/0001-6918(79)90026-x

Lapatki, B. G., Oostenveld, R., Van Dijk, J. P., Jonas, I. E., Zwarts, M. J., & Stegeman, D. F. (2010). Optimal placement of bipolar surface EMG electrodes in the face based on single motor unit analysis. *Psychophysiology, 47*(2), 299-314. doi: 10.1111/j.1469-8986.2009.00935.x

Lee, T. M. C., & Chan, C. C. H. (2000). Stroop interference in Chinese and English. *Journal of Clinical and Experimental Neuropsychology, 22*(4), 465-471. doi: 10.1076/1380-3395(200008)22:4;1-0;ft465

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59-80. doi: 10.1348/026151003321164627

Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. *Acta Psychologica, 104*(2), 167-190. doi: 10.1016/s0001-6918(00)00019-6

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

Linck, J. A., Hoshino, N., & Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment Lex, 3*(3), 349-374. doi: 10.1075/ml.3.3.06lin

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. *Dagenbach, Dale [Ed]; Carr, Thomas H [Ed] (1994) Inhibitory processes in attention, memory, and language (pp 189-239) xiv, 461 pp San Diego, CA, US: Academic Press; US*, 189-239.

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., & Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: Response inhibition versus interference suppression. *Brain and Cognition, 74*(3), 347-357. doi: 10.1016/j.bandc.2010.09.004

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605-621. doi: 10.1080/20445911.2013.795574

Luo, L., Craik, F. I. M., Moreno, S., & Bialystok, E. (2013). Bilingualism Interacts With Domain in a Working Memory Task: Evidence From Aging. *Psychology and Aging, 28*(1), 28-34. doi: 10.1037/a0030875

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29-41.

Macleod, C. M. (1991). Half a century of research on the Stroop effect - an integrative review. *Psychological Bulletin, 109*(2), 163-203. doi: 10.1037//0033-2909.109.2.163

Maess, B., Friederici, A. D., Damian, M., Meyer, A. S., & Levelt, W. J. M. (2002). Semantic category interference in overt picture naming: Sharpening current density localization by PCA. *Journal of Cognitive Neuroscience, 14*(3), 455-462. doi: 10.1162/089892902317361967

Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(3), 503-535. doi: 10.1037/0278-7393.33.3.503

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods, 164*(1), 177-190. doi: 10.1016/j.jneumeth.2007.03.024

Martin-Rhee, M. M., & Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Bilingualism-Language and Cognition, 11*(1), 81-93.

Marzecova, A., Asanowicz, D., Kriva, L., & Wodniecka, Z. (2013). The effects of bilingualism on efficiency and lateralization of attentional networks. *Bilingualism-Language and Cognition, 16*(3), 608-623. doi: 10.1017/s1366728912000569

Mass Univariate ERP Toolbox. (2015). from http://openwetware.org/index.php?title=Mass_Univariate_ERP_Toolbox&oldid=867176

Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16*(5), 798-817. doi: 10.3758/pbr.16.5.798

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*(5), 1124-1140. doi: 10.1037//0278-7393.26.5.1124

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93*(3), 568-592. doi: 10.1037/0021-9010.93.3.568

Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development, 75*(5), 1373-1386. doi: 10.1111/j.1467-8624.2004.00746.x

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology-General, 132*(2), 228-252. doi: 10.1037/0096-3445.132.2.228

Mishra, R. K., Hilchey, M. D., Singh, N., & Klein, R. M. (2012). On the time course of exogenous cueing effects in bilinguals: Higher proficiency in a second language is associated with more rapid endogenous disengagement. *Quarterly Journal of Experimental Psychology, 65*(8), 1502-1510. doi: 10.1080/17470218.2012.657656

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. C. (2004). Inner speech as a retrieval aid for task goals: the effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*(2-3), 123-142. doi: 10.1016/j.actpsy.2003.12.004

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science, 21*(1), 8-14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Mohades, S. G., Struys, E., Van Schuerbeek, P., Baeken, C., Van De Craen, P., & Luypaert, R. (2014). Age of second language acquisition affects nonverbal conflict processing in children: an fMRI study. *Brain and Behavior, 4*(5), 626-642. doi: 10.1002/brb3.246

Morales, J., Yudes, C., Gomez-Ariza, C. J., & Bajo, M. T. (2015). Bilingualism modulates dual mechanisms of cognitive control: Evidence from ERPs. *Neuropsychologia, 66*, 157-169. doi: 10.1016/j.neuropsychologia.2014.11.014

Morton, J. B., & Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Developmental Science, 10*(6), 719-726.

Morris, N., & Jones, D. M. (1990). Memory Updating in Working Memory - the Role of the Central Executive. *British Journal of Psychology, 81*, 111-121.

Neill, W. T. (1977). Inhibitory and facilitatory processes in selective attention. *Journal of Experimental Psychology-Human Perception and Performance, 3*(3), 444-450. doi: 10.1037//0096-1523.3.3.444

Nicolay, A. C., & Poncelet, M. (2013). Cognitive advantage in children enrolled in a second-language immersion elementary school program for three years. *Bilingualism-Language and Cognition, 16*(3), 597-607. doi: 10.1017/s1366728912000375

Norman, D. A., & Shallice, T. (1983). Attention to action: Willed and automatic control of behavior. *Bulletin of the Psychonomic Society, 21*(5), 354-354.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113. doi: http://dx.doi.org/10.1016/0028-3932(71)90067-4

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28*(2), 191-230. doi: 10.1017/s0142716407070117

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology, 112*(4), 713-719. doi: http://dx.doi.org/10.1016/S1388-2457(00)00527-7

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114*(2), 227-252.

Paap, K. R. (2014). The role of componential analysis, categorical hypothesising, replicability and confirmation bias in testing for bilingual advantages in executive functioning. *Journal of Cognitive Psychology, 26*(3), 242-255. doi: 10.1080/20445911.2014.891597

Paap, K. R., Darrow, J., Dalibar, C., & Johnson, H. A. (2015). Effects of script similarity on bilingual advantages in executive control are likely to be negligible or null. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01539

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology, 66*(2), 232-258. doi: 10.1016/j.cogpsych.2012.12.002

Piai, V., Roelofs, A., & van der Meij, R. (2012). Event-related potentials and oscillatory brain responses associated with semantic and Stroop-like interference effects in overt naming. *Brain Research, 1450*, 87-101. doi: 10.1016/j.brainres.2012.02.050

Pelham, S. D., & Abrams, L. (2014). Cognitive Advantages and Disadvantages in Early and Late Bilinguals. *Journal of Experimental Psychology-Learning Memory and Cognition, 40*(2), 313-325. doi: 10.1037/a0035224

Peal, E., & Lambert, W. E. (1962). The Relation of Bilingualism to Intelligence. *Psychological Monographs, 76*(27), 1-23.

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 113*(4), 535-551. doi: 10.1016/j.jecp.2012.06.013

Porcaro, C., Medaglia, M. T., & Krott, A. (2015). Removing speech artifacts from electroencephalographic recordings during overt picture naming. *NeuroImage, 105*, 171-180. doi: 10.1016/j.neuroimage.2014.10.049

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology, 22*(3), 415-422. doi: 10.1016/j.acn.2007.01.015

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annu Rev Neurosci, 13*, 25-42. doi: 10.1146/annurev.ne.13.030190.000325

Prior, A., & Gollan, T. H. (2013). The elusive link between language control and executive control: A case of limited transfer. *Journal of Cognitive Psychology, 25*(5), 622-645. doi: 10.1080/20445911.2013.821993

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism-Language and Cognition, 13*(2), 253-262.

Rabbitt, P. (1997). Introduction: Methodologies and Models in the Study of Executive Function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function*. Hove, UK: Psychology Press.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*(3), 446-461. doi: 10.1037/0033-2909.86.3.446

Ratiu, I., & Azuma, T. (2015). Working memory capacity: Is there a bilingual advantage? *Journal of Cognitive Psychology, 27*(1), 1-11. doi: 10.1080/20445911.2014.976226

Raven, J., Raven, J. C., & Court, J. H. (2000, updated 2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Section 3: The Standard Progressive Matrices. San Antonio, TX: Harcourt Assessment.

Ridderinkhof, K. R. (2002a). *Activation and suppression in conflict tasks: empirical clarification through distributional analyses*.

Ridderinkhof, K. R. (2002b). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychological Research-Psychologische Forschung, 66*(4), 312-323.

Ridderinkhof, K. R., Wildenberg, v., Wijnen, J., & Burle, B. s. (2004). *Response Inhibition in Conflict Tasks Is Revealed in Delta Plots*. Paper presented at the British Journal of Mathematical & Statistical Psychology.

Riès, S., Janssen, N., Burle, B., & Alario, F. X. (2013). Response-Locked Brain Dynamics of Word Production. *Plos One, 8*(3), e58197. doi: 10.1371/journal.pone.0058197

Roberts, R. J., Hager, L. D., & Heron, C. (1994). Prefrontal Cognitive Processes - Working Memory and Inhibition in the Antisaccade Task. *Journal of Experimental Psychology-General, 123*(4), 374-393. doi: 10.1037//0096-3445.123.4.374

Robertson, I. H., & Garavan, H. (2004). Vigilant Attention. *Cognitive Neurosciences Iii, Third Edition*, 631-640.

Rodriguez-Pujadas, A., Sanjuan, A., Fuentes, P., Ventura-Campos, N., Barros-Loscertales, A., & Avila, C. (2014). Differential neural control in early bilinguals and monolinguals during response inhibition. *Brain and Language, 132*, 43-51. doi: 10.1016/j.bandl.2014.03.003

Rodriguez-Pujadas, A., Sanjuan, A., Ventura-Campos, N., Roman, P., Martin, C., Barcelo, F., . . . Avila, C. (2013). Bilinguals Use Language-Control Brain Areas More Than Monolinguals to Perform Non-Linguistic Switching Tasks. *Plos One, 8*(9), 8. doi: 10.1371/journal.pone.0073028

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition, 42*(1-3), 107-142. doi: 10.1016/0010-0277(92)90041-f

Roelofs, A. (2003). Goal-referenced selection of verbal action: Modelling attentional control in the Stroop task. *Psychological Review*, *110*(1), 88-125. doi: 10.1037/0033-295x.110.1.88

Roelofs, A., Piai, V. r., & Garrido Rodriguez, G. (2011). Attentional inhibition in bilingual naming performance: Evidence from delta-plot analyses. *Frontiers in Psychology, 2*. doi: 10.3389/fpsyg.2011.00184

Roelofs, A., van Turennout, M., & Coles, M. G. H. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. *Proceedings of the National Academy of Sciences of the United States of America, 103*(37), 13884-13889. doi: 10.1073/pnas.0606265103

Rogers, R. D., & Monsell, S. (1995). Costs of a Predictable Switch Between Simple Cognitive Tasks. *Journal of Experimental Psychology-General, 124*(2), 207-231. doi: 10.1037//0096-3445.124.2.207

Romaine, S. (1989). *Bilingualism*: Basil Blackwell.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36.

Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Matos, L., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal of Neuroscience, 112*(6), 759-776. doi: 10.1080/00207450290025752

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*(4), 850-869. doi: 10.1016/j.jml.2012.02.007

Saer, D. J. (1923). The Effect of Bilingualism on Intelligence. *British Journal of Psychology-General Section, 14*, 25-38.

Salvatierra, J. L., & Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *International Journal of Bilingualism, 15*(1), 26-37. doi: 10.1177/1367006910371021

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., & Salmon, D. P. (2010). What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition, 13*(2), 231-252.

Schmiedek, F., Oberauer, K., Wilhelm, O., Suss, H. M., & Wittmann, W. W. (2007). Individual differences in components their relations to working of reaction time distributions and memory and intelligence. *Journal of Experimental Psychology-General, 136*(3), 414-429. doi: 10.1037/0096-3445.136.3.414

Shao, Z. S., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology, 5*, 10. doi: 10.3389/fpsyg.2014.00772

Shao, Z. S., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *Quarterly*

Journal of Experimental Psychology, 65(10), 1927-1944. doi: 10.1080/17470218.2012.670252

Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language, 51*(3), 381-398.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R Compatibility - Effect of An Irrelevant Cue on Information Processing. *Journal of Applied Psychology, 51*(3), 300-304.

Singh, N., & Mishra, R. K. (2015). The modulatory role of second language proficiency on performance monitoring: evidence from a saccadic countermanding task in high and low proficient bilinguals. *Frontiers in Psychology, 5*, 14. doi: 10.3389/fpsyg.2014.01481

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology, 2*, 8. doi: 10.3389/fpsyg.2011.00183

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology-Human Perception and Performance, 22*(2), 461-479. doi: 10.1037/0096-1523.22.2.461

Spieler, D. H., Balota, D. A., & Faust, M. E. (2000). Levels of selective attention revealed through analyses of response time distributions. *Journal of Experimental Psychology-Human Perception and Performance, 26*(2), 506-526. doi: 10.1037//0096-1523.26.2.506

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*(3), 281-284.

Starreveld, P. A., & la Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*(4), 896-918. doi: 10.1037/0278-7393.22.4.896

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*(4), 745-759. doi: 10.1080/17470210500162854

Steinhauser, M., & Hubner, R. (2009). Distinguishing Response Conflict and Task Conflict in the Stroop Task: Evidence From Ex-Gaussian Distribution Analysis. *Journal of Experimental Psychology-Human Perception and Performance, 35*(5), 1398-1412. doi: 10.1037/a0016467

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662. doi: 10.1037/0096-3445.121.1.15

Tao, L., Marzecova, A., Taft, M., Asanowicz, D., & Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Frontiers in Psychology, 2*, 19. doi: 10.3389/fpsyg.2011.00123

Thierry, G., & Wu, Y.J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proceedings of the National Academy of Sciences of the USA.* **104**(30), 12530-12535.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*(26), 14792-14797.

Tipper, S. P. (1985). The negative priming effect - inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 37*(4), 571-590.

Tipper, S. P., & Driver, J. (1988). Negative priming between pictures and words in a selective attention task - evidence for semantic processing of ignored stimuli. *Memory & Cognition, 16*(1), 64-70. doi: 10.3758/bf03197746

Townsend , J. T., & Ashby , F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200-239). Hillsdale, NJ: Erlbaum.

Tse, C. S., & Altarriba, J. (2012). The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism-Language and Cognition, 15*(3), 663-676. doi: 10.1017/s1366728912000077

Tse, C. S., & Altarriba, J. (2014). The relationship between language proficiency and attentional control in Cantonese-English bilingual children: evidence from Simon, Simon switching, and working memory tasks. *Frontiers in Psychology, 5,* 18. doi: 10.3389/fpsyg.2014.00954

Tse, C. S., & Altarriba, J. (2015). Local and Global Task Switching Costs in Bilinguals Who Vary in Second Language Proficiency. *American Journal of Psychology*, *128*(1), 89-106. doi: 10.5406/amerjpsyc.128.1.0089

Tse, C. S., Balota, D. A., Yap, M. J., Duchek, J. M., & McCabe, D. P. (2010). Effects of Healthy Aging and Early Stage Dementia of the Alzheimer's Type on Components of Response Time Distributions in Three Attention Tasks. *Neuropsychology*, *24*(3), 300-315. doi: 10.1037/a0018274

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. *Intelligence, 38*(1), 111-122. doi: 10.1016/j.intell.2009.08.002

Valian, V. (2015). Bilingualism and cognition. *Bilingualism-Language and Cognition, 18*(1), 3-24. doi: 10.1017/s1366728914000522

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition, 110*(1), 84-99. doi: 10.1016/j.cognition.2008.10.013

Wu, Y.J.; Cristino, F., Leek, E.C., & Thierry, G. (2013). Non-selective lexical access in bilinguals is spontaneous and independent of input monitoring: Evidence from eye tracking. *Cognition.* **129**(2), 418-425.

Wu, Y.J., & Thierry, G. (2010). Chinese-English bilinguals reading English hear Chinese. *Journal of Neuroscience.***30**(22), 7646-7651.

Wu, Y.J.; & Thierry, G. (2012) Unconscious translation during incidental foreign language processing.*NeuroImage.* **59**(4), 3468-3473.

Wu, Y.J., & Thierry, G. (2013). Fast Modulation of Executive Function by Language Context in Bilinguals. *The Journal of Neuroscience.* **33**(33), 13533-13537.

Wuhr, P., & Ansorge, U. (2005). Exploring trial-by-trial modulations of the Simon effect. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 58*(4), 705-731. doi: 10.1080/02724980443000269

West, S., Taylor, A., & Wu, W. (2012). Model Fit and Model Selection in Structural Equation Modeling. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 209-231). New York: Guilford Press.

Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A., & Olson, R. K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110*(1), 157-172. doi: 10.1037//0021-843x.1001.1.157

Yang, S. J., Yang, H. J., & Lust, B. (2011). Early childhood bilingualism leads to advances in executive attention: Dissociating culture and language. *Bilingualism-Language and Cognition, 14*(3), 412-422. doi: 10.1017/s1366728910000611

Yntema, D. B. (1963). Keeping Track of Several Things At Once. *Human Factors, 5*(1), 7-17.

Yow, W. Q., & Li, X. Q. (2015). Balanced bilingualism and early age of second language acquisition as the underlying mechanisms of a bilingual executive control advantage: why variations in bilingual experiences matter. *Frontiers in Psychology, 6*, 12. doi: 10.3389/fpsyg.2015.00164

Yudes, C., Macizo, P., & Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Frontiers in Psychology, 2*, 9. doi: 10.3389/fpsyg.2011.00309

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time-course of phonological encoding. *Journal of Memory and Language, 34*(3), 311-334. doi: 10.1006/jmla.1995.1014

APPENDICES

APPENDIX 1

Language History Questionnaire Version 1

Please indicate your self-perceived proficiency in **English** by drawing a **vertical line** on the scale below. The far left end stands for no knowledge in English, and the far right end stands for 100% native-like proficiency.

		Little/No knowledge	Native-like
Speaking	**0**		**100**
Understanding speech	**0**		**100**
Reading	**0**		**100**
Writing	**0**		**100**

1. Please indicate your self-perceived proficiency in **Chinese** by drawing a **vertical line** on the scale below. The far left end stands for no knowledge in Chinese, and the far right end stands for 100% native-like proficiency.

		Little/No knowledge	Native-like
Speaking	**0**		**100**
Understanding speech	**0**		**100**
Reading	**0**		**100**
Writing	**0**		**100**

2. Please draw a vertical line on the scale to indicate the current use of both **English and Chinese** in oral communications, at home and outside home.

		No English	All English
At home	**0**		**100**
Outside home	**0**		**100**
		No Chinese	**All Chinese**
At home	**0**		**100**

Outside home **0** ├──────────────┤ **100**

3. Do you speak Chinese (Mandarin or Cantonese) fluently? (By fluently we mean that for everyday conversations, you are able to converse with native speakers without having to consciously translate).

4. Please list all languages you speak (which reached **native or native-like competence**) in the order you began to acquire them (since born). Indicate at what age you began to learn each and at what age (approximately) you mastered each:

	Language	Age began to learn	Age mastered
1			
2			
3			
4			

5. In what setting did you acquire your second (and third, if applicable) language? (E.g., at home, through school, living abroad, other)

Second language
- At home
- Through school
- Living abroad
- Other (please specify)

____

Third language
- At home
- Through school
- Living abroad
- Other (please specify)

____

Forth language
- At home
- Through school
- Living abroad
- Other (please specify)

____

6. **Language(s) of parents** (or primary caretaker, guardian, etc):___________________

7. Please roughly describe your previous language use history using the table provided.

Please specify your own language use experience at different stages of your life till now.

Age :			
Only use **first** language regularly			
Only use **second** language regularly			
Use both languages regularly, but in **different** settings (e.g. use Chinese at home and English outside of home)			
Use both languages regularly, but in **the same** setting (e.g. use both languages both at home and outside of home)			

8. **Current language use** (check the one that applies)

Do you now:

_______use primarily one language? If so, which one? __________

_______use both languages regularly but in different settings (i.e., one at home and one at school, one with friends and one with family, etc.)

_______use both languages every day within the same setting (i.e., use both at home)

9. Do you have friends or family who are also bilingual in the two languages you speak?

Yes No

10. When speaking with these bilingual friends/family members, do you ever find yourself using both languages within the same conversation or even in the same sentence?

_____________ Yes, frequently

_____________ Yes, but only rarely

_____________ No, never

11. **True/False:**

T F I mix languages only when talking to friends or family.

T F I mix languages in conversations with other bilinguals because this enables me to express myself better.

T F I mix languages because of other reasons (please specify):

____

T F I try not to mix languages in the same conversation.

12. What is the highest level of certificate that you have got now?

1. GCSE or below		6. Others(please specify)	
2. A level			
3. Bachelor's degree			
4. Master's degree			
5. Doctorates or above			

13. If one of your languages is Chinese, please indicate which dialect you speak (e.g., Mandarin, Cantonese etc. . .)

____

APPENDIX 2

Vocabulary List

Below are four vocabulary lists. Each list contains words from the same category. Please tick the words that you know. Only tick when you are **100% sure** what the word refers to.

List 1									
toga		thongs		robe		turban		anorak	
mortarboard		cardigan		brief		nylons		tutu	
sabot		dressing gown		tunic		slouch		muff	
deel		vest		waistcoat		doublet		plimsolls	
dungarees		romper suit		nappy		hijab		bearskin	
fez		kilt		sari		bonnet		corset	
cut offs		leotard		sarong		beret		dirndl	
sneakers		sombrero		swimsuit		negligee		pullover	
kimono		petticoat		Stetson		trilby		swim goggles	
catsuit		mitten		sandals		casquette		bowler	
swimming trunk		wetsuit		slippers		ballerina			
pyjama		girdle		nightie		cheongsam			
tracksuit		tanktop		hula skirt		y-fronts			
sarafan		boxers		stola		moccasins			

List 2									
wardrobe		vitrine		bureau		sleeper sofa		highchair	
hassock		cabinet		shoji screen		couch		light	
pew		cradle		hope chest		cupboard		bassinet	
clothes valet		ottoman		rocking chair		bergere		chandelier	
hutch		fauteuil		sconce		valet		cot	
cubbies		murphy bed		tuffet		settee		lounger	
chest		sofa bed		stool		wing chair		tansu	
sofa		bentwood rocker		porch swing		chair		bookcase	
footstool		futon		shelf		armchair		throne	
divan		nightstand		lawn chair		lift chair		love seat	
Windsor chair		hall tree		lamp		bookshelf			
recliner		armoire		desk		sideboard			
chaise lounge		credenza		table		curio			

List 3

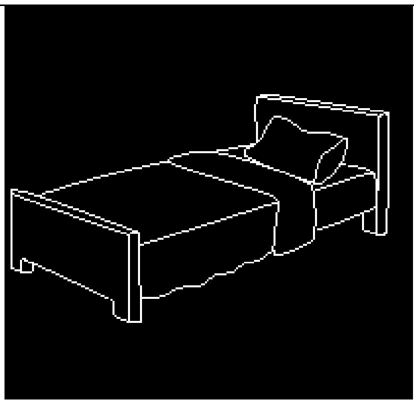
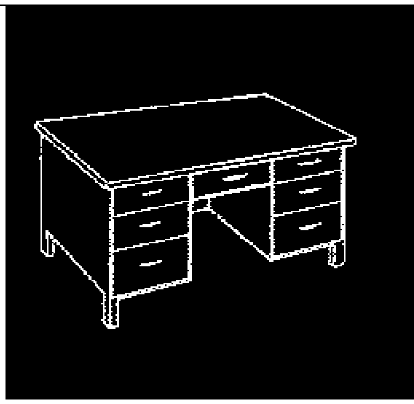
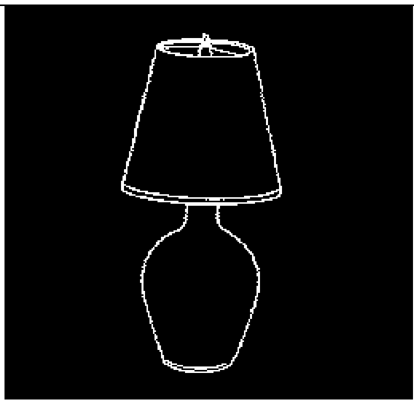
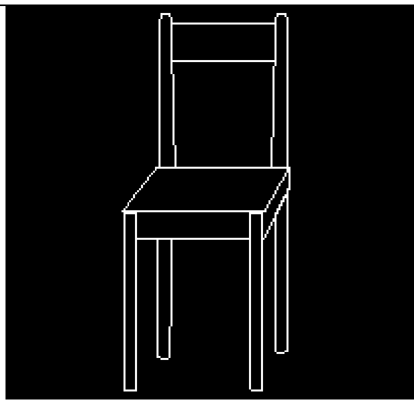
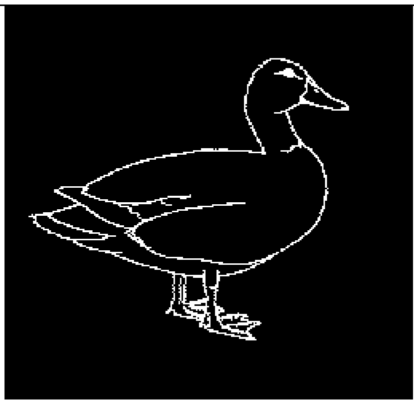
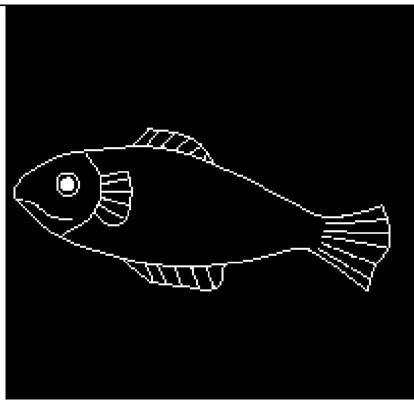
wheelbarrows		vise		clamp		spring		plane	
pick		cleaver		spanner		fishing rod		sander	
scraper		garden hoe		pliers		hammer		azada	
file		scoop		shear		brace		dibber	
plunger		locking		shovel		bank stick		fishing reel	
gauge		scalpel		tongs		snips		saw	
drill		stapler		stiletto		wrench		threader	
sprocket holder		bevel square		screwdriver		cultivator		brush	
machete		mallet		spade		anvil		tomahawk	
ladder		scissors		rake		rotisserie		scale	
wire stripper		hoe		spear		mattock			
secateurs		trowel		punch		auger			
daisy grubber		tape measure		jigsaw		garden loppers			

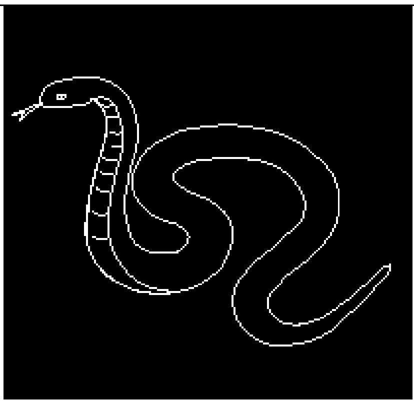
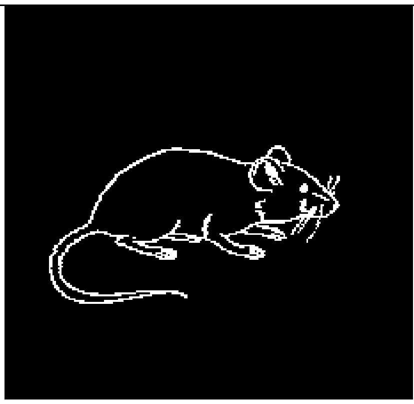
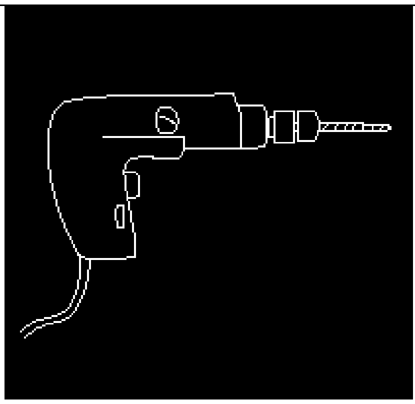
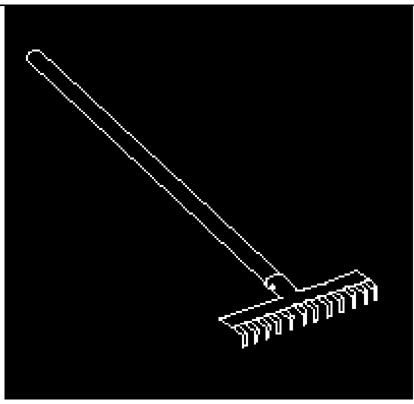
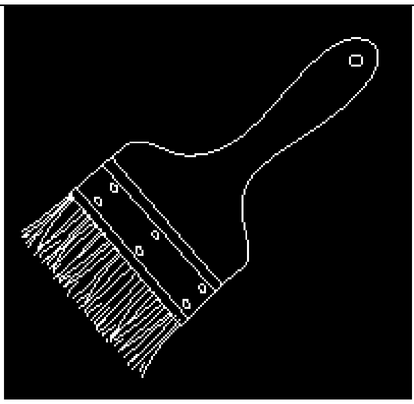
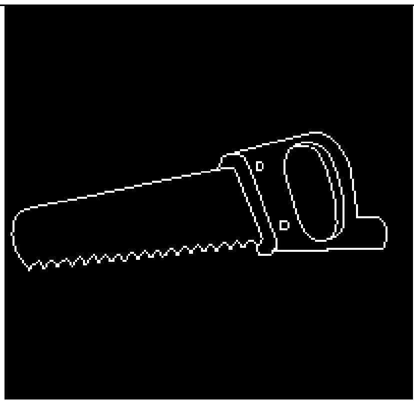
List 4

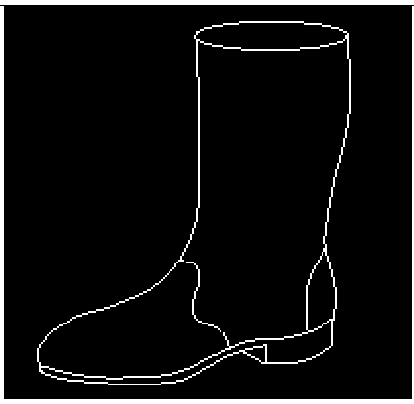
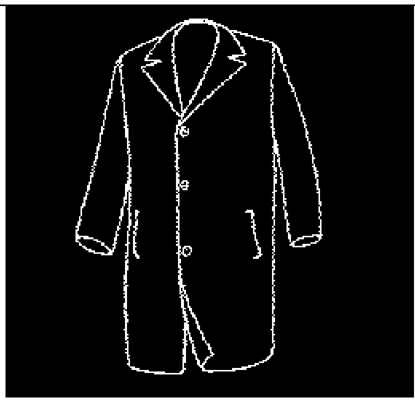
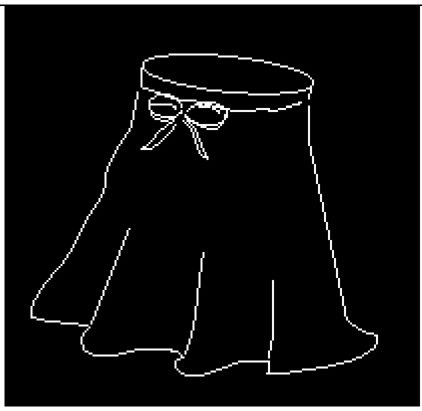
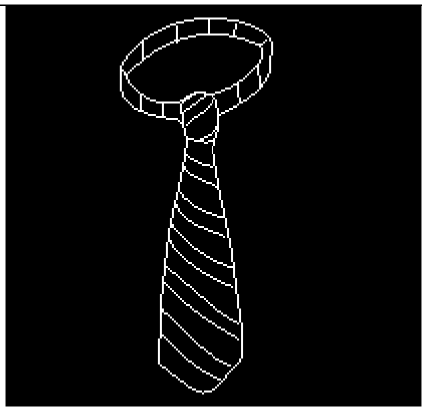
springhare		sparrow		manatee		pheasant		lark	
hamster		colubrid		cockatoo		lacertid		hare	
newt		flamingo		walrus		puffin		lemur	
eagle		whale		pocket gopher		ferret		anchovy	
hummingbirds		mole		hippopotamus		gannet		crow	
frog		gourami		paca		bobcat		eel	
dormice		loon		pelican		baboon		lingcod	
python		raven		shark		petrel		rook	
groundhog		barracuda		rays		swallows		budgerigar	
gnatcatcher		gundi		viper		boas		sea anemone	
lizard		gecko		potamotrigon		halibut			
stoat		skink		kiwi		toadfish			
otter		caribou		gerbil		chameleon			
crocodile		lizard		rhinoceros		flounder			

APPENDIX 3

Pictures Used in the Semantic Blocking Task

Bed	Desk
	
Lamp	Chair
	
Duck	Fish
	
Snake	Mouse

	
Drill	Rake
	
Brush	Saw
	
Boot	Coat

	
Skirt	Tie
	

APPENDIX 4

Language Questionnaire version 2

Please be aware that responses to the questionnaire will be kept in confidence and will be anonymous. The information sought for this study will be used as part of a PhD project.
Please read following questions carefully and respond to the best of your knowledge.

Section A. Proficiency Rating

1. Is English your first language?
   Yes (Go to No. 2) No (Go to No. 4)

2. Do you speak any other language(s) fluently?
   Yes (Go to No. 3) No (Go to No. 5)

3. What is the other language that you speak fluently? Do you use it Daily?
   __________________ __________________
   At what age did you start to learn that language?
   _____ (Go to No.5)

4. What is your first language?
   ____
   At what age did you start to learn English?
   _____

5. Please indicate your self-perceived proficiency in **English** on a scale of 1 to 7. (1 = poor, 4= average (i.e., you can get by OK), 7 = excellent).

	poor			average			excellent
Speaking	1	2	3	4	5	6	7
Understanding speech	1	2	3	4	5	6	7
Reading	1	2	3	4	5	6	7
Writing	1	2	3	4	5	6	7

6. Please indicate your self-perceived proficiency of **the other language** on a scale a scale of 1 to 7. (1 = poor, 4= average (i.e., you can get by OK), 7 = excellent).

	poor			average			excellent
Speaking	1	2	3	4	5	6	7
Understanding speech	1	2	3	4	5	6	7
Reading	1	2	3	4	5	6	7

Writing	1	2	3	4	5	6	7

Section B. Language Use

7. Please list all languages that you speak, **in the order** you began to acquire them (since born). Indicate at what age you began to learn each. Also state the overall level of proficiency on a scale of 7. (1 = poor, 4= average (i.e., you can get by OK), 7 = excellent).

	Language	Age began to learn	Overall proficiency
1			
2			
3			
4			

8. In what setting did you acquire the languages stated above? (E.g., at home, through school, living abroad, other)

First language
- At home
- Through school
- Living abroad
- Other (please specify) ____________________

Second language
- At home
- Through school
- Living abroad
- Other (please specify) ____________________

Third language
- At home
- Through school
- Living abroad
- Other (please specify) ____________________

Forth language
- At home
- Through school
- Living abroad
- Other (please specify) ____________________

9. Please indicate the percentage of the time you are currently using each language in **oral communications** at home and outside home. (Your percentages should add up to 100%)

Overall

List language here				
List percentage here				

At home

List language here				
List percentage here				

Outside home

List language here				
List percentage here				

10. **Language(s) of parents** (or primary caretaker, guardian, etc):

____

11. Please roughly describe your previous language use history using the table provided. Please specify your own language use experience at different stages of your life till now.

Example:

Age:	***0-3***	***4-8***	***9-12***	**…**	**…**	
*Only use **first** language regularly*	√					
*Only use **second** language regularly*		√				
*Use both languages regularly, but in **different** settings (e.g. use Chinese at home and English outside of home)*			√			

Please fill in the following table

Age :						
Only use **first** language regularly						
Only use **second** language regularly						
Use both languages regularly, but in **different** settings (e.g. use Chinese at home and English outside of home)						
Use both languages regularly, but in **the same** setting (e.g. use both languages both at home and outside of home)						

12. **Current language use** (check the one that applies)

Do you now:

_______use primarily one language? If so, which one? _________

_______use both languages regularly but in different settings (i.e., one at home and one at school, one with friends and one with family, etc.)

_______use both languages every day within the same setting (i.e., use both at home)

13. Do you have friends or family who are also bilingual in the two languages you speak?

Yes No

14. When speaking with these bilingual friends/family members, do you ever find yourself using both languages within the same conversation or even in the same sentence?

_____________ Yes, frequently

_____________ Yes, but only rarely

_____________ No, never

15. **True/False:**

T F I mix languages only when talking to friends or family.

T F I mix languages in conversations with other bilinguals because this enables me to express myself better.

T F I mix languages because of other reasons (please specify):

____

T F I try not to mix languages in the same conversation.

Section C General Information
Education

16. What is the highest level of certificate that you have got now?

1. GCSE or below	
2. A level	
3. Bachelor's degree	
4. Master's degree	
5. Doctorates or above	
6. Others(please specify)	

17. Education of parents

Father		**Mother**	
1. GCSE or below		1. GCSE or below	
2. A level		2. A level	
3. Bachelor's degree		3. Bachelor's degree	
4. Master's degree		4. Master's degree	
5. Doctorates or above		5. Doctorates or above	
6. Others(please specify)		6. Others(please specify)	

18. Family income

Of these income groups, can you tell me which letter best represents your family total household income per week?

1	< £200		6	£600 - £700	
2	£200 - £300		7	£700 - £800	
3	£300 - £400		8	£800 - £900	
4	£400 - £500		9	£900 - £1,000	
5	£500 - £600		10	> £1,000	

19. Which socioeconomic class do you belong to

1	lower middle class	
2	middle class	
3	upper middle class	