frontiers RESEARCH TOPICS

BILINGUALISM AND COGNITIVE CONTROL

Topic Editors
Judith F. Kroll, Ingrid Christoffels and Teresa Bajo

frontiers in PSYCHOLOGY

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

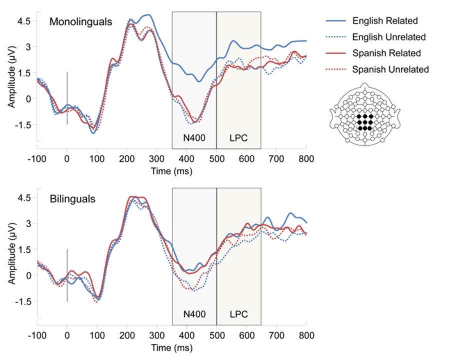
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-154-3
DOI 10.3389/978-2-88919-154-3

BILINGUALISM AND COGNITIVE CONTROL

Topic Editors:
Judith F. Kroll, Penn State University, USA
Ingrid Christoffels, University of Leiden, Netherlands
Teresa Bajo, Universidad de Granada, Spain



ERPs measured over the centroparietal electrodes where N400 and LPC were maximal (a liner derivation of C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2) as a function of relatedness, homograph language, and language group. Figure taken from Hoshino N and 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. *Front. Psychol.* 3:9. doi: 10.3389/fpsyg.2012.00009.

Research on bilingual language processing reveals an important role for control processes that enable bilinguals to negotiate the potential competition across their two languages. The requirement for control that enables bilinguals to speak the intended language and to switch between languages has also been suggested to confer a set of cognitive consequences for executive function that extend beyond language to domain general cognitive skills. Many recent studies have examined aspects of how cognitive control is manifest during bilingual language processing, how individual differences in cognitive resources influence second language learning and performance, and the range of cognitive tasks that appear to be influenced by bilingualism. However, not all studies demonstrate a bilingual advantage in all tasks that tap into cognitive control. Indeed, many questions are unanswered that are critical to our understanding of bilingual control: What aspects of cognitive control are enhanced for proficient bilinguals? How are individual differences in cognitive control related to language acquisition, proficiency, or professional translation skill? How does the language environment affect concurrent processing? How exactly does language control come about in tasks such as speech production, switching between languages, or translation? When and how does inhibitory processing support language control?

The focus of this Research Topic is on executive control and bilingualism. The goal is to have a broad scope that includes all of these issues. We seek empirical contributions using different methodologies including behavioral, computational and neuroscience approaches. We also welcome theoretical contributions that provide detailed discussion of models or mechanisms that account for the relationship between bilingualism and cognitive control. We aim to provide a platform for new contributions that represent a state-of-the art overview of approaches to cognitive control in bilingualism. We hope that this Research Topic will enable the field to formulate more precise hypotheses and causal models on the relation between individual differences, cognitive control and bilingual language processing.

Table of Contents

frontiers in PSYCHOLOGY

EDITORIAL
published: 17 April 2013
doi: 10.3389/fpsyg.2013.00199

Introduction to *Bilingualism and Cognitive Control*

I. K. Christoffels[1]*, J. F. Kroll[2] and M. T. Bajo[3]

[1] *Department of Psychology, Faculty of Psychology and Leiden Institute for Brain and Cognition, University of Leiden, Leiden, Netherlands*
[2] *Department of Psychology, Pennsylvania State University, University Park, PA, USA*
[3] *Departamento de Psicología Experimental, Universidad de Granada, Granada, Spain*
**Correspondence: ichristoffels@fsw.leidenuniv.nl*

Edited by:
Lorenza S. Colzato, Leiden University, Netherlands

Reviewed by:
Lorenza S. Colzato, Leiden University, Netherlands

In the past decade, there has been an upsurge of research on bilingualism. A theme in this work is that the bilingual's two languages are always active, at times converging with one another to produce benefits to comprehension and production, but at other times conflicting, with the requirement to negotiate cross-language competition. A goal in the recent work has been to characterize the cognitive processes that enable bilinguals to negotiate the cross-talk between their two languages. The ease with which highly proficient bilinguals are able to speak each of their languages without frequent errors or intrusions and, at the same time, switch between the two languages in contexts in which code switching is appropriate or encouraged, suggests the presence of a high level of cognitive control. At the same time, behavioral and neurocognitive studies have shown that bilinguals differ from monolinguals in their performance on tasks that are purely cognitive, often showing advantages relative to monolinguals, and clear differences in neural function and structure. A key question is how we might begin to relate the findings on language control to the documented cognitive consequences of bilingualism. The papers in this special issue on *Bilingualism and Cognitive Control* represent the best of the new research on each of these issues to understand how control in language processing is achieved and how domain-general cognitive processes are themselves affected by language experience.

In what follows, we review and summarize the main goals of the papers that comprise this effort. We note that the questions about cognitive control are increasingly exploiting sophisticated methods (e.g., see the work using delta plot analysis, Roelofs et al., 2011) and extending analyses of executive function to different populations of language users (e.g., Tao et al., 2011). We set out in this special topic issue to ask a number of questions concerning cognitive control in bilingualism. The contributing authors addressed these questions in very different and interesting ways.

A still controversial issue is: *How is cognitive control manifest during bilingual language processing?* Calabria et al. (2012) asked whether bilingual language control is the same as other types of cognitive control. The pattern of symmetrical switch costs that is often obtained for proficient bilinguals did not replicate across domain in a non-linguistic switch task in the same group. Therefore, the authors argue that bilingual language control is not completely subsidiary to domain general control. Their conclusions contrast to those of Roelofs et al. (2011) who used delta-plot analyses to show that telltale characteristics of the reaction time distribution in bilingual naming performance shows clear similarities with performance in other domains. This finding supports the view that inhibition is a mechanism of attentional control in bilingual language performance, and that it is a domain general mechanism. In their review of patient studies, intracranial electrical, and transcranial magnetic stimulation, Hervais-Adelman et al. (2011) proposed a distinction between two distinct networks contributing to the executive control of language. A fronto-basal-ganglia loop is implicated in the inhibition of the irrelevant language during production, and may be crucial for access to translation equivalents. A fronto-parietal network appears to subserve more general switching mechanisms. Perhaps such a distinction may be helpful in resolving controversies that have arisen in behavioral studies of switching and control.

van Heuven et al. (2011) used the Stroop task to address the effects of cross-language similarity. Three groups of trilingual's systematically differed on whether their languages use the same or different scripts. They obtained similar within-language Stroop interference across groups, but between-language Stroop interference was modulated by cross-language similarity, in particular by differences in script between languages. Hoshino and Thierry (2012) used Event Related Potentials (ERP) to investigate interlingual homographs. These stimuli induce between language semantic conflict because they have identical form, but different meanings in both languages. Both readings of interlingual homographs were processed, even in a single language context. Interlingual homographs modulated the N400 time window during word reading. Interestingly, there was no effect in later time windows suggesting that the activated semantic information of the non-target language is not explicitly processed. Adaptive performance during learning was addressed by Davidson and Indefrey (2011) who also used ERPs to investigate how learning of grammatical categories leads to changing error-related electrophysiological activity over time. They showed that learning only took place when performance feedback was provided. Finally, Morales et al. (2011) report a study in which the manifestation of cognitive control is investigated in the hitherto unexplored direction of grammatical gender. Their study suggested the presence of an inhibitory mechanism that suppresses grammatical gender when it is a source of competition between languages.

The idea that bilingualism may result in cognitive advantages is a topic that has received a great deal of attention in the recent scientific and popular literature: *What aspects of cognitive control are enhanced for proficient bilinguals?* Recent evidence suggests that (only) specific aspects of executive control are related to bilingualism. Papers in the special issue support this claim by demonstrating specific cognitive advantages associated with bilingualism. Tao et al. (2011) found bilingual enhancement of executive functions for early and late bilinguals. They investigated how age of L2 acquisition and relative balance of the two languages influenced performance on a lateralized attention network test (ANT) for executive function. Monolinguals were less efficient in the resolution of conflict than both early and late bilinguals. Although both early and late bilinguals were found to have more efficient attentional networks, late bilinguals showed the greatest advantage in conflict resolution, whereas early bilinguals were advantaged in monitoring. By testing simultaneous interpreters, Yudes et al. (2011) explored non-verbal executive processes in a bilingual group known to have exceptional working memory abilities. They compared interpreters, bilinguals and monolinguals on the Wisconsin Card Sorting Task and the Simon task, taken to reflect cognitive flexibility and inhibitory control, respectively. Interpreters showed higher mental flexibility than the other groups. However, a similar Simon effect indicated that interpreters do not out perform other groups on inhibitory control of executive functioning. Investigating a rather different aspect of cognitive functioning, Hommel et al. (2011) addressed the relation between bilingualism and creativity. They showed a specific advantage for high proficient participants for convergent thinking. In contrast, low proficient bilinguals were better at a divergent thinking task. This suggests that bilingualism supports a relative focused cognitive control style, with strong top-down control.

A number of papers in the special issue addressed language switching habits as an index of individual differences rather than focusing on the more general patterns of switch costs themselves. Festman and Münte (2012) used performance on a switch task to identify participants as switchers or non-switchers based on the degree of unintentional switching during naming. They found that non-switchers were advantaged on aspects of the Wisconsin Card Sorting Task and the Flanker task, indicating that individual differences in language control and executive control function are related. Rodriguez-Fornells et al. (2012) developed a questionnaire to psychometrically assess self-perceived individual differences in language switching. Soveri et al. (2011) used this questionnaire combined with a multiple regression approach to investigate whether performance on tasks measuring different executive functions could be predicted by the frequency of language switches in everyday life.

A related focus on individual differences in another set of papers concerned the question of *How individual differences in cognitive resources influence second language learning and performance.* Bartolotti et al. (2011) asked whether cognitive control and bilingual experiences influence success in learning a new language. They tested groups of participants with high and low cognitive control abilities and high and low bilingual experience. Results indicated that both factors may influence learning success; their relative importance depends of the amount of overlap between languages. In the high interference condition cognitive control abilities influenced learning success. Pivneva et al. (2012) addressed fluency and nativeness in the L1 and L2 spontaneous monologue and dialogue. Not only proficiency levels influenced speech planning and production, but these processes were also more efficient for bilinguals with high inhibitory capacity, in particular for highly proficient bilinguals. Finally, Reiterer et al. (2011) examined EEG gamma band phase synchrony measures for high and low proficient participants. Processing in the second language required significantly higher synchronization strength than in the first language. Lower proficiency was related to a stronger synchronized network than higher proficiency, which was more widely distributed in left fronto-parietal areas.

As should be clear, this special topic has generated many exiting new contributions to some of the most important questions that relate to bilingualism and cognitive control. An issue that has not been addressed in this set of papers, is the question how language environment affects concurrent processing. We anticipate that the context of language use and language learning will become a topic of interest and investigation in the next wave of research on bilingualism and control. That said, the current set of papers offers a fresh and novel perspective on how cognitive control is engaged to enable proficient language performance and how skilled bilingual performance changes cognition in ways that suggest much greater optimism about the plasticity of the adult mind and brain than previously understood.

REFERENCES

Bartolotti, J., Marian, V., Schroeder, S. R., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Psychol.* 2:324. doi: 10.3389/fpsyg.2011.00324

Calabria, M., Hernández, M., Branzi, F. M., and Costa, A. (2012). Qualitative differences between bilingual language control and executive control: evidence from task-switching. *Front. Psychol.* 2:399. doi: 10.3389/fpsyg.2011.00399

Davidson, D. J., and Indefrey, P. (2011). Error-related activity and correlates of grammatical plasticity. *Front. Psychol.* 2:219. doi: 10.3389/fpsyg.2011.00219

Festman, J., and Münte, T. F. (2012). Cognitive control in Russian–German bilinguals. *Front. Psychol.* 3:115. doi: 10.3389/fpsyg.2012.00115

Hervais-Adelman, A. G., Moser-Mercer, B., and Golestani, N. (2011). Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. *Front. Psychol.* 2:234. doi: 10.3389/fpsyg.2011.00234

Hommel, B., Colzato, L. S., Fischer, R., and Christoffels, I. K. (2011). Bilingualism and creativity: benefits in convergent thinking come with losses in divergent thinking. *Front. Psychol.* 2:273. doi: 10.3389/fpsyg.2011.00273

Hoshino, N., and 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. *Front. Psychol.* 3:9. doi: 10.3389/fpsyg.2012.00009

Morales, L., Paolieri, D., and Bajo, T. (2011). Grammatical gender inhibition in bilinguals. *Front. Psychol.* 2:284. doi: 10.3389/fpsyg.2011.00284

Pivneva, I., Palmer, C., and Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. *Front. Psychol.* 3:57. doi: 10.3389/fpsyg.2012.00057

Reiterer, S., Pereda, E., and Bhattacharya, J. (2011). On a possible relationship between linguistic expertise and EEG gamma band phase synchrony. *Front. Psychol.* 2:334. doi: 10.3389/fpsyg.2011.00334

Rodriguez-Fornells, A., Krämer, U. M., Lorenzo-Seva, U., Festman, J., and Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Front. Psychol.* 2:388. doi: 10.3389/fpsyg.2011.00388

Roelofs, A., Piai, V., and Rodriguez, G. G. (2011). Attentional inhibition in bilingual naming performance: evidence from delta-plot analyses. *Front. Psychol.* 2:184. doi: 10.3389/fpsyg.2011.00184

Soveri, A., Rodriguez-Fornells, A., and Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. *Front. Psychol.* 2:183. doi: 10.3389/fpsyg.2011.00183

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., and Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Front. Psychol.* 2:123. doi: 10.3389/fpsyg.2011.00123

van Heuven, W. J. B., Conklin, K., Coderre, E. L., Guo, T., and Dijkstra, T. (2011). The influence of cross-language similarity on within- and between-language Stroop effects in trilinguals. *Front. Psychol.* 2:374. doi: 10.3389/fpsyg.2011.00374

Yudes, C., Macizo, P., and Bajo, T. (2011). The influence of expertise in simultaneous interpreting on non-verbal executive processes. *Front. Psychol.* 2:309. doi: 10.3389/fpsyg.2011.00309

Received: 21 March 2013; accepted: 02 April 2013; published online: 17 April 2013.

Citation: Christoffels IK, Kroll JF and Bajo MT (2013) Introduction to Bilingualism and Cognitive Control. Front. Psychol. ***4****:199. doi: 10.3389/fpsyg.2013.00199*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 13 January 2012
doi: 10.3389/fpsyg.2011.00399

Qualitative differences between bilingual language control and executive control: evidence from task-switching

Marco Calabria[1], Mireia Hernández[1,2,3], Francesca M. Branzi[1] and Albert Costa[1,4]*

[1] Department of Technology, Universitat Pompeu Fabra, Barcelona, Spain
[2] Center for Mind/Brain Sciences, University of Trento, Trento, Italy
[3] Cognitive Neuropsychology Laboratory, Department of Psychology, Harvard University, Cambridge, MA, USA
[4] Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

Edited by:
Ingrid Christoffels, University of Leiden, Netherlands

Reviewed by:
Andrea Kiesel, Julius-Maximilians-University, Germany
Rico Fischer, Technische Universität Dresden, Germany

****Correspondence:***
Albert Costa, Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, C/Tanger, 122-140, 08014 Barcelona, Spain.
e-mail: costalbert@gmail.com

Previous research has shown that highly proficient bilinguals have comparable switch costs in both directions when they switch between languages (L1 and L2), the so-called "symmetrical switch cost" effect. Interestingly, the same symmetry is also present when they switch between L1 and a much weaker L3. These findings suggest that highly proficient bilinguals develop a language control system that seems to be insensitive to language proficiency. In the present study, we explore whether the pattern of symmetrical switch costs in language switching tasks generalizes to a non-linguistic switching task in the same group of highly proficient bilinguals. The end goal of this is to assess whether bilingual language control (bLC) can be considered as subsidiary to domain-general executive control (EC). We tested highly proficient Catalan–Spanish bilinguals both in a linguistic switching task and in a non-linguistic switching task. In the linguistic task, participants named pictures in L1 and L2 (Experiment 1) or L3 (Experiment 2) depending on a cue presented with the picture (a flag). In the non-linguistic task, the same participants had to switch between two card sorting rule-sets (color and shape). Overall, participants showed symmetrical switch costs in the linguistic switching task, but not in the non-linguistic switching task. In a further analysis, we observed that in the linguistic switching task the asymmetry of the switch costs changed across blocks, while in the non-linguistic switching task an asymmetrical switch cost was observed throughout the task. The observation of different patterns of switch costs in the linguistic and the non-linguistic switching tasks suggest that the bLC system is not completely subsidiary to the domain-general EC system.

Keywords: bilingualism, executive control, language control, task-switching, language switching

INTRODUCTION

A remarkable skill of bilingual speakers is the ability to confine speech to one language while preventing interference from the unintended language. The cognitive process underlying this ability is often referred to as bilingual language control (bLC; e.g., Green, 1998; Costa and Santesteban, 2004; Crinion et al., 2006; Abutalebi and Green, 2007; Christoffels et al., 2007). Although there is disagreement regarding the nature of the bLC mechanisms, there is a general consensus that certain aspects of domain-general executive control (EC) functions mediate this ability (Abutalebi et al., 2008). However, it is still unclear whether bLC is completely subsidiary to the domain-general EC system or whether it also involves mechanisms specific to language.

In fact, the relationship between bLC and domain-general EC processes can be characterized in at least two different ways. First, one could think of bLC as a set of processes that are fully subsidiary to the domain-general EC functioning. That is, a bilingual speaker producing language would engage the very same set of EC processes that are involved in other non-linguistic activities requiring EC. Under this hypothesis, when switching language as a function of the interlocutor, individuals would engage the very same control mechanisms as when they are asked to switch between different non-linguistic tasks in everyday life. Alternatively, the bLC system may be only partially subsidiary to domain-general EC processes. That is, it is possible that the continuous control that bilingual speakers exert over their two languages results in the development of control processes specific to language (Costa and Santesteban, 2004). Although they probably make use of certain aspects of the EC system, additional processes may become specifically engaged in language switch related tasks. From this viewpoint, the crosstalk between the bLC and domain-general EC would still be present, leading to the repeatedly reported bilingual advantages in EC (e.g., Bialystok et al., 2004; Costa et al., 2008, 2009; Hernández et al., 2010). At the same time, however, some aspects of the bLC system would be specific to language and not necessarily related to the EC system.

Here, we set out to gain some initial insights on this issue by exploring a phenomenon observed both in language switching and task-switching, namely, the "asymmetrical switch cost" (see below). By doing this, we hope to shed some light on the crosstalk between the processes involved in bLC and those involved in domain-general EC.

ON THE FUNCTIONING OF EC SYSTEM IN BILINGUALS AND MONOLINGUALS

A first indication revealing that bilingualism affects the EC functioning can be found in those studies comparing monolinguals and bilinguals performing EC tasks. An increasing body of literature reveals that the continuous use of two languages seems to enhance processes related to domain-general EC such as those put at play in Stroop-like tasks and non-linguistic task-switching. This has been indexed through the observation of reduced Stroop-like interference and switch costs for bilinguals relative to monolinguals (e.g., Bialystok et al., 2004, 2006, 2008, 2010; Colzato et al., 2008; Costa et al., 2008, 2009; Bialystok and Viswanathan, 2009; Hernández et al., 2010). In particular, Prior and MacWhinney (2010) assessed whether bilinguals would show an advantage over monolinguals in non-linguistic task-switching with two sorting rules (sorting by shape or by color). They found that bilinguals had a reduced switch cost compared to monolinguals. Of the multiple components involved in task-switching (e.g., goal shifting, rule activation, etc., see Rubinstein et al., 2001), the authors hypothesized that the bilingual advantage in task-switching might be related to a more efficient goal shifting. The reasoning behind this hypothesis was that bilinguals' lifelong use of language switching may lead to an enhancement of the abilities of goal shifting also in the non-linguistic cognitive control mechanisms[1].

Other indications of the crosstalk between EC and bLC come from neuroimaging studies comparing monolinguals and bilinguals. Recently, Abutalebi et al. (2011) found differences in the way the dorsal anterior cingulate cortex (ACC) was recruited during conflict resolution in the flanker task. Specifically, bilinguals revealed a smaller activation of this area than monolinguals during conflict resolution. This pattern of brain activation was consistent with the fact that behaviorally bilinguals showed a reduced magnitude of the conflict effect compared to monolinguals. These results suggest that the ACC, one area within the cognitive control network, is engaged to a different extent in bilinguals and monolinguals during EC tasks.

There are also some indications of qualitative differences in brain activation between monolinguals and bilinguals during EC tasks (Garbin et al., 2010). In the study of Garbin et al. (2010), monolinguals and bilinguals completed a task-switching experiment using two sorting rules determined by stimulus color and shape. The authors found that bilinguals recruited brain areas normally engaged during language control (left inferior frontal gyrus), whereas monolinguals did not. This suggests that bilinguals recruit different neural structures relative to monolinguals in tasks involving the EC system.

Overall, these results indicate that bilingualism has an impact on the development of EC. However, they do not exclude the possibility that bLC involves certain processes that are outside the EC system. One way to explore the crosstalk between bLC and EC is to look at the qualitative difference of performance in tasks that engaged these two systems. Let us explain in more detail these qualitative aspects, specifically the asymmetry of the switch costs in linguistic and non-linguistic task-switching.

QUALITATIVE DIFFERENCES IN SWITCH COSTS BETWEEN LINGUISTIC AND NON-LINGUISTIC TASK-SWITCHING

Abutalebi and Green (2007), in a review of neuroimaging studies, suggested that the same neural regions (the dorsolateral prefrontal cortex, the ACC and the caudate nucleus) are engaged during both language switching tasks (e.g., Price et al., 1999; Hernandez et al., 2000, 2001; for a review see Hervais-Adelman et al., 2011) and non-linguistic task-switching (e.g., Botvinick et al., 1999; Crone et al., 2006). This indirect evidence supports the hypothesis that the mechanisms for language control are subsidiary to those of the domain-general EC.

However, an fMRI study conducted by Abutalebi et al. (2008) may actually be interpreted as going against the claim of functional overlap between bLC and EC. The authors demonstrated the existence of a neural network that is specifically recruited to switch between two different linguistic registers but not between two intra-linguistic tasks. This suggests that some processes at play during bLC are "language-specific" and not recruited for any other switching task.

In this article we further explore the issue of the crosstalk between bLC and EC by assessing qualitative aspects of these two systems (see below). To do so, we employ tasks involving bLC (language switching task) and EC (non-linguistic switching task) to compare the patterns of switch costs observed within the same population of highly proficient bilinguals. These two tasks share many different cognitive components and one can argue that in fact, the language switching task is just a specific instantiation of the more general task-switching paradigm (see for example, Abutalebi and Green, 2008). If so, and according to the first hypothesis put forward above, the pattern of results in the two tasks should be similar. In contrast, if bLC is not fully subsidiary to the EC processes, one could predict that the pattern of results in the two tasks may not be identical. Let us be more specific about the pattern of results we are referring to.

One of the most robust effects in task-switching is the so-called "local switch cost" (e.g., Meiran, 1996; Monsell, 2003; Koch et al., 2010; Schneider and Anderson, 2010; Martin et al., 2011). This cost refers to the observation of slower reaction times (RTs) for trials that require a task-switch in comparison to trials that do not require such a switch. For our present purposes, it is interesting that the magnitude of the local switch cost is not constant for any given task, but rather depends on the relative difficulty of the two tasks at hand during the experiment. Given differences in task difficulty, local switch costs tend to be larger when switching into the easier task than when switching into the more difficult one. For example, consider a switching task where task 1 consists

[1] The question of which EC processes are involved in task-switching is a complex issue that goes beyond the purposes of the present article. Several theories have exemplified how task-switching might be mediated by separable executive control processes [e.g., attention-to-action (ATA) model by Norman and Shallice, 1986; the frontal-lobe executive (FLE) model by Duncan, 1986; and the strategic response–deferment (SRD) model, Meyer and Kieras, 1997]. For a detailed description of such theories see reviews by Rubinstein et al. (2001) and Monsell (2003). Here, we refer to Rubinstein et al.'s (2001) account discussed in Prior and MacWhinney's (2010) study on the bilingual advantage in task-switching. Rubinstein et al. (2001) proposed that at least two processes of the EC system are involved in task-switching, namely "goal shifting" and "rule activation." "Goal shifting" updates the content of the declarative working memory about the two task-sets; whereas rule activation enables the selection of the current task and disables the rules of the previous one.

in sorting cards by color and task 2 consists in sorting cards by shape, with unpredictable switches from one task (e.g., color) to the other (shape). The switch cost observed when switching to the more difficult task "sorting by shape" are usually smaller than when switching to the easier task "sorting by color" (e.g., Nagahama et al., 2001; Rubinstein et al., 2001; Martin et al., 2011). This phenomenon, often referred to as the asymmetrical switch cost, has received many different explanations in the task-switching literature (for a review see Koch et al., 2010; Schneider and Anderson, 2010). Given the focus of this article, we will only discuss briefly what is, perhaps, the most influential account of this asymmetrical switch cost.

According to Allport et al. (1994), the "task-set inertia hypothesis", part of the switching cost stems from the need to retrieve a task-set that has been inhibited in the previous trial. Furthermore, the amount of inhibition applied to a given task-set (e.g., sorting by color or shape) depends on the relative strength of the task. That is, the easier task is inhibited more strongly than the more difficult one. Given this imbalance, the asymmetrical switch cost comes about in the following way: when performing the more difficult task (i.e., sorting by shape), the system has to strongly inhibit the task-set corresponding to the easier task (sorting by color). Hence, in the following trial, retrieving the strongly inhibited task-set will incur in a large switching cost. In contrast, when performing the easier task (i.e., sorting by color), the system has to inhibit with less strength the task-set corresponding to the more difficult task (sorting by shape). Consequently, in the following trial, retrieving the not-very-much inhibited task-set will incur in a small switching cost. Therefore, switching from the easier to the more difficult task will incur in a smaller switch cost (from color to shape) than switching from the more difficult to the easier task (from shape to color)[2].

Similarly, when the task-switching involves two languages, low-proficient bilinguals show asymmetrical switch costs (i.e., larger switch costs when switching into the easier language), which parallels the pattern of the non-linguistic task-switching paradigms. That is, for low-proficient bilinguals switching into the less proficient (and hence, the more difficult task) language (L2) is easier (in terms of RTs and errors) than switching into the more proficient (and hence, the easier task) language (L1; e.g., Meuter and Allport, 1999). This linguistic asymmetrical switch cost can be explained in the same manner as domain-general asymmetrical switching costs. In fact, Meuter and Allport (1999) argued that the magnitude of the inhibition applied to two languages is dependent on the relative strength of the two languages. Therefore, when the less proficient L2 needs to be produced, the more proficient L1 needs to be inhibited more than the other way around. Thus, an asymmetrical switch cost arises because the amount of inhibition that needs to be overcome during the switch into L1 is larger than when switching into L2. This pattern of asymmetries in low-proficient bilinguals fits very well with the notion that the same control processes involved in bLC are the ones that are also at play in domain-general EC.

The framework described above makes a straightforward prediction: whenever there is a difference in the difficulty of the tasks (or languages) involved in the switching task, there should be an asymmetrical switching cost, being such cost larger when switching into the easier task. Along the same lines, symmetrical switch costs are expected for switching tasks involving tasks of similar difficulty.

Crucial for present purposes is the fact that several studies conducted with highly proficient bilinguals have given only partial support to this prediction. Highly proficient bilinguals do not seem to show asymmetrical language switching costs regardless of the difficulty of the languages involved in the task. Let us be more specific and describe the pattern of language switching cost for highly proficient bilinguals in some detail.

As expected, when highly proficient bilinguals are asked to switch between their two proficient languages (hence little difference in difficulty between the two tasks), the switching costs are comparable in both directions (from L1 to L2 and vice versa; Costa and Santesteban, 2004; Costa et al., 2006). However, and crucial for present purposes, when these bilinguals are asked to switch between languages of different difficulties (e.g., switching between their L1 and their L3), the predicted asymmetrical switch cost is not present. In a series of experiments Costa et al. (2006) showed that in highly proficient bilinguals the symmetrical switch cost was present irrespective of the age of acquisition of L2, the similarities of two languages involved in the switching task and language proficiency. Given this pattern, two questions emerge:

(a) Why highly proficient bilinguals do not show the predicted asymmetrical switch cost when switching between languages of different proficiency, as the low-proficient bilinguals do?
(b) Would these bilinguals be sensitive to task difficulty when performing a non-linguistic switching task (e.g., would they show asymmetrical switch costs)? Answering this second question is the goal of the present article.

In trying to answer the first question, Costa and Santesteban (2004) hypothesized that highly proficient bilinguals might recruit a qualitatively different bLC when performing the language switching task compared to low-proficient bilinguals. As proposed by Costa and Santesteban (2004), there might be a shift in the type of mechanisms responsible for the selection of the intended language once a certain level of proficiency is attained in an L2. That is, it is possible that at some point highly proficient bilinguals do not make use of inhibition (as low-proficient ones probably do), but instead they make use of a mechanism that restricts lexical competition to the intended language. Importantly, once highly proficient bilinguals develop such as a mechanism it would be applied also to other languages (e.g., a weaker L3).

This explanation contains the implicit assumption that bLC might be to some extent different from EC processes in general, and hence the "task-set inertia" hypothesis (Allport et al., 1994) for the performance of highly proficient bilinguals is not granted. Note

[2]Other authors have proposed different accounts based on long-term memory retrieval processes (e.g., Allport and Wylie, 2000; Mayr and Kliegl, 2000; Bryck and Mayr, 2008). One assumption is that the retrieval of irrelevant task traces interferes with selection of the relevant task and that more instances of the more difficult task would be encoded/retrieved into long-term memory than in the case of the easier task. Since the amount of interference is proportional to the number of irrelevant task traces in long-term memory, the interference will be larger when switching into the easier task than into the more difficult one. This leads to a larger switch cost when switching from the more difficult to the easier task than vice versa.

that this hypothesis would predict asymmetrical switch costs when switching from L3 into L1 for highly proficient bilinguals, given that one language (L3) is harder than the other (L1) – similarly to what happens when low-proficient bilinguals switch between L1 and L2. Thus, according to this hypothesis, the difference in the relative strength between L1 and L3 should involve a different amount of inhibition when speaking in one language or the other and therefore produce asymmetries in switch costs as well.

Regardless these explanations, what is relevant here is the potential generalizability of such a lack of asymmetrical switch costs of highly proficient bilinguals to non-linguistic tasks. That is, the question is whether the crosstalk between bLC and EC systems is such that the relative insensitivity of highly proficient bilinguals to task difficulty in the language switching task will also be present in a non-linguistic switching task.

If the bLC system is fully subsidiary to the EC system, it is reasonable to predict that whichever pattern is observed in the language switching task will also be present in a non-linguistic switching task. Hence, we predict that differences in task difficulty should not lead to asymmetrical switch costs in these bilinguals, in the same way that differences in language difficulty do not lead to asymmetrical switch costs for this group. On the other hand, if bLC is governed by processes that are, to some extent, independent of the EC system, then it is possible that the symmetrical switch costs observed for language switching do not generalize to non-linguistic task-switching.

We put these predictions to test by comparing the performance of highly proficient Catalan–Spanish bilinguals in a linguistic and non-linguistic switching paradigm and examining the qualitative pattern of the switch costs. Specifically, we compared the symmetry/asymmetry of the switch costs between tasks differing in their level of difficulty. We used an adaptation of the linguistic switching task previously employed by Costa and Santesteban (2004), through which we expected to replicate the typical symmetrical switch cost of highly proficient bilinguals between L1 and L2 and also between L1 and L3. Note that for the sake of completeness we present two experiments: in Experiment 1 highly proficient bilinguals switched between L1 and L2, and in Experiment 2 between L1 and L3.

Concerning the non-linguistic task, we used a task-switching where participants had to switch between two rule-sets of a card sorting task (color and shape). As previously described, sorting by color is easier than sorting by shape. This effect of task difficulty permitted us to compare the non-linguistic switching task with the language switching task. We defined the non-linguistic switching task such that it did not require changing languages and it did not require explicit verbalization of the response.

To recapitulate, we will examine the issue of the crosstalk between bLC and EC in two ways:

(a) From a qualitative point of view: by examining the pattern of the switch costs in terms of the symmetry/asymmetry in the linguistic and non-linguistic switching tasks. If highly proficient bilinguals show a symmetrical switch cost in the language switching task, the same symmetrical pattern is expected in the non-linguistic switching task if the mechanisms of bLC are completely subsidiary to the EC system.

(b) From a quantitative point of view: by examining any potential correlations between linguistic and non-linguistic switch costs. Significant correlations between switch costs in linguistic and non-linguistic switching tasks could indicate that the bilinguals' behavior in the bLC generalizes to a non-verbal domain, such as domain-general EC.

PARTICIPANTS

Fourteen bilinguals (mean age = 23.2, range = 18–27 years old) took part in Experiment 1, and 15 bilinguals did it in Experiment 2 (mean age = 20.3, range = 18–23 years old). All participants in both experiments were early and highly proficient Catalan–Spanish bilinguals. All participants had Catalan as L1 and they learned Spanish before the age of 6. Their proficiency in the two languages was tested by means of a questionnaire. Each participant self-rated on a four-point scale the abilities of speaking, comprehension, writing and reading for each language (1 = poor, 2 = regular, 3 = good, 4 = perfect). All the participants were highly proficient in both L1 and L2 (see **Table 1**). In addition, participants in Experiment 2 were low-proficient in English (L3).

EXPERIMENT 1: LINGUISTIC SWITCHING BETWEEN L1 AND L2 AND NON-LINGUISTIC SWITCHING TASK

MATERIALS AND PROCEDURE

Linguistic switching task

Eight pictures of objects were selected from Snodgrass and Vanderwart (1980). Half of them referred to cognate words [Spanish/Catalan names: "Caracol"/"Cargol" (in English, snail); "Escoba"/"Escombra" (broom); "Martillo"/"Martell" (hammer); "Reloj"/"Rellotge" (watch)], and the other half to non-cognate words ["Calcetín"/"Mitjó" (sock); "Manzana"/"Poma" (apple); "Silla"/"Cadira" (chair); "Tenedor"/"Forquilla" (fork)].

Participants were required to name the picture in Catalan or in Spanish. A Catalan or Spanish flag, which was presented along with the picture, acted as a cue to indicate in which language subjects had to name the picture.

Table 1 | Language proficiency (mean and SD) of speaking, comprehension, writing, and reading abilities for each language, self-rated on a four-point scale (1 = poor, 2 = regular, 3 = good, 4 = perfect).

Experiment 1	Catalan, mean (SD)	Spanish, mean (SD)
Speaking	4.0 (0.0)	3.9 (0.3)
Comprehension	4.0 (0.0)	4.0 (0.0)
Pronunciation	4.0 (0.0)	3.9 (0.3)
Reading	4.0 (0.0)	4.0 (0.0)
Writing	4.0 (0.0)	3.9 (0.3)
Experiment 2	**Catalan, mean (SD)**	**English, mean (SD)**
Speaking	4.0 (0.0)	2.1 (0.5)
Comprehension	4.0 (0.0)	2.9 (0.7)
Pronunciation	4.0 (0.0)	2.1 (0.7)
Reading	4.0 (0.0)	3.0 (0.4)
Writing	4.0 (0.0)	2.7 (0.5)

There were two types of trials: (a) those in which participants were required to name the picture in the same language as the preceding trial (repeat trial), (b) those in which participants were required to name in a different language with respect to the previous trial (switch trial). There were a total of 320 trials divided in two blocks with 160 trials each. The total distribution of trials was: 128 repeat trials in Catalan, 128 repeat trials in Spanish, 64 switch trials in Catalan, and 64 in Spanish.

Participants were asked to name the picture as fast as possible and they were informed that the language to be used was indicated by a flag, presented on the top of the picture. At the beginning of each series a word cue was presented for 1000 ms indicating in which language participants had to start to name ("CATALÀ," for Catalan; "ESPAÑOL," Spanish). Then the picture appeared for 1700 ms and the timeout to respond was 5000 ms. The pictures were presented in a series of three to seven trials and at the end of each series an asterisk appeared and the participants pressed the spacebar to start the next series. The experiment started with a practice session of 80 trials.

Non-linguistic switching task

Three shapes (square, circle, and triangle) and three colors (green, blue, and red) were selected for the task. The three shapes were combined with the three colors, resulting in a total of nine colored shapes (e.g., green square, blue square etc.). Participants were presented with an array containing three shapes, two at the top of the screen and one at the bottom. They were instructed to match the shape at the bottom with one of the two at the top of the display according to two possible criteria (shape or color). The criterion was indicated by a cue ("COLOR," for Color; "FORMA," for Shape) appearing in the center of the array. As in the linguistic version of the task, there were two types of trials: repeat and switch trials.

At the beginning of each series a word cue was presented for 1000 ms indicating by which rule participants must start matching each item ("COLOR," for Color; "FORMA," for Shape). Then the array appeared for 2500 ms and the timeout to respond was 3000 ms.

Participants gave the response by pressing the two keys "M" or "V" according to the position of the matched picture at the top of the array. Specifically, they had to press "M" key when the correct answer was at the top-right part of the array and the "V" key when the correct response was at the top-left part of the array. The experiment started with a practice session of 80 trials.

The experiments were controlled by the software DMDX (Forster and Forster, 2003), which recorded participants' vocal and manual responses. Responses were analyzed off-line and naming latencies were measured from the onset of the word trough Checkvocal, a program of data analysis of naming tasks in DMDX (Protopapas, 2007). Participants always performed the linguistic switching and then the non-linguistic switching task. The order of the two tasks was not counterbalanced.

RESULTS

Linguistic switching cost

The variables considered in the analyses were "type of trial" (switch vs. repeat) and "response language" (L1 and L2) which were included as within-subject factors in a repeated-measure ANOVA on naming latencies. Naming latencies 3 SD above or below a given participant's mean were excluded from the analyses. Also the naming latencies in which the participants produced a different name from what was expected were excluded from the analyses.

Reaction times. Overall participants were slower in switch trials (886 ms) compared to repeat trials [801 ms; $F(1, 13) = 55.11$, $MSE = 1822.67$, $p < 0.0001$, $\eta_p^2 = 0.81$], and faster to name in L1 (829 ms) than in L2 [857 ms; $F(1, 13) = 4.81$, $MSE = 2318.88$, $p = 0.05$, $\eta_p^2 = 0.27$]. But the cost to switch to L1 (87 ms) and to L2 (82 ms) was the same, as indexed by a non-significant "type of trial" × "response language" interaction [$F(1, 13) = 0.15$, $MSE = 741.59$, $p = 0.70$; see **Figure 1A**]. That is, there was a symmetrical switch cost.

Accuracy. No difference in accuracy was found between switch and repeat trials [Type of trial: $F(1, 13) = 2.29$, $MSE = 9.65$, $p = 0.15$] and between L1 and L2 [Response language: $F(1, 13) = 0.40$, $MSE = 22.76$, $p = 0.54$]. The interaction between type of trial and response language was not significant either [$F(1, 13) = 0.19$, $MSE = 6.64$, $p = 0.66$; see **Table 1**].

Non-linguistic switching cost

The variables considered in the analysis were "type of trial" (switch vs. repeat) and "sorting criteria" (color and shape), which were included as a within-subject factor in a repeated-measure ANOVA using RTs as a dependent variable.

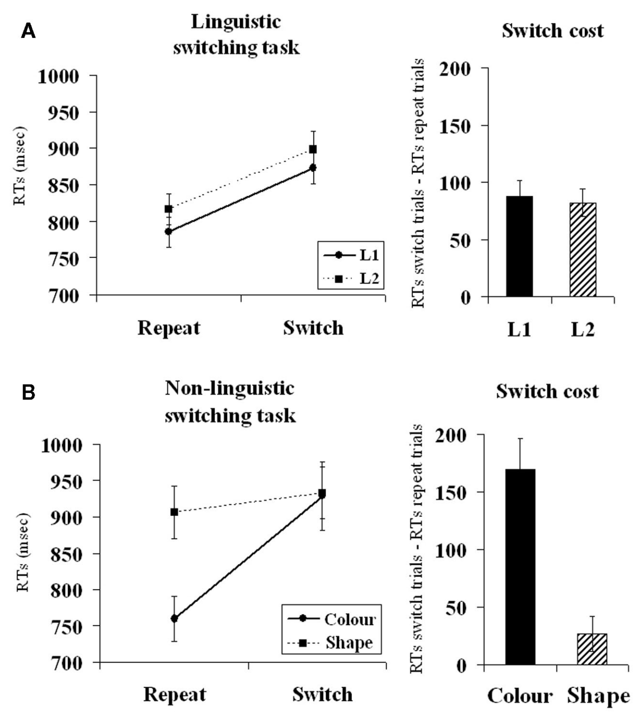


FIGURE 1 | (A) Performances on the linguistic switching task (left) and magnitude of the switch cost for L1 and L2 (right). Error bars represent SE. **(B)** Performances on the non-linguistic switching task (left) and magnitude of the switch cost for color and shape (right). Error bars represent the SE.

Reaction times. Overall participants were slower in switch trials (931 ms) compared to repeat trials [833 ms; $F(1, 13) = 38.42$, $MSE = 3505.52$, $p < 0.0001$, $\eta_p^2 = 0.75$], and faster to sort by color (843 ms) than to sort by shape [920 ms; $F(1, 13) = 40.32$, $p < 0.0001$, $MSE = 2011.41$, $\eta_p^2 = 0.76$]. In this case the switch cost interacted with "type of trial" [$F(1, 13) = 19.88$, $MSE = 3592.72$, $p = 0.001$, $\eta_p^2 = 0.61$]. That is, participants showed a cost when they switched from shape to color [169 ms, $F(1, 13) = 37.57$, $MSE = 5353.39$, $p < 0.0001$, $\eta_p^2 = 0.74$], but not from color to shape [27 ms, $F(1, 13) = 2.85$, $MSE = 1744.86$, $p = 0.11$; see **Figure 1B**].

Accuracy. There was a tendency toward lower accuracy for switch trials (91.25%) over repeat ones [94.75%; Type of trial: $F(1, 13) = 3.64$, $MSE = 17.80$, $p = 0.08$]. Also, participants were less accurate in sorting by shape (90.0%) than by color [94.7%; $F(1, 13) = 14.22$, $MSE = 22.40$, $p < 0.01$, $\eta_p^2 = 0.52$; see **Table 2**].

To summarize, we found that bilingual participants showed symmetrical switch costs in the linguistic task-switching, but in the non-linguistic one we found asymmetrical switch costs since only switching into color resulted in a cost.

EXPERIMENT 2: LINGUISTIC SWITCHING BETWEEN L1 AND L3 AND NON-LINGUISTIC SWITCHING TASK

As advanced in the Introduction, one could argue that the symmetrical switch costs between L1 and L2 of highly proficient bilinguals are due to both tasks (naming in L1 and naming in L2) being equally easy for highly proficient bilinguals. In other words, we would have a difference in difficulty between color and shape in the non-linguistic task-switching but not between L1 and L2 in the language switching task. Thus, in this experiment, bilinguals (who were still highly proficient in both Catalan and Spanish) conducted the language switching task between their L1 (Catalan) and L3 (English) for which they were low-proficient.

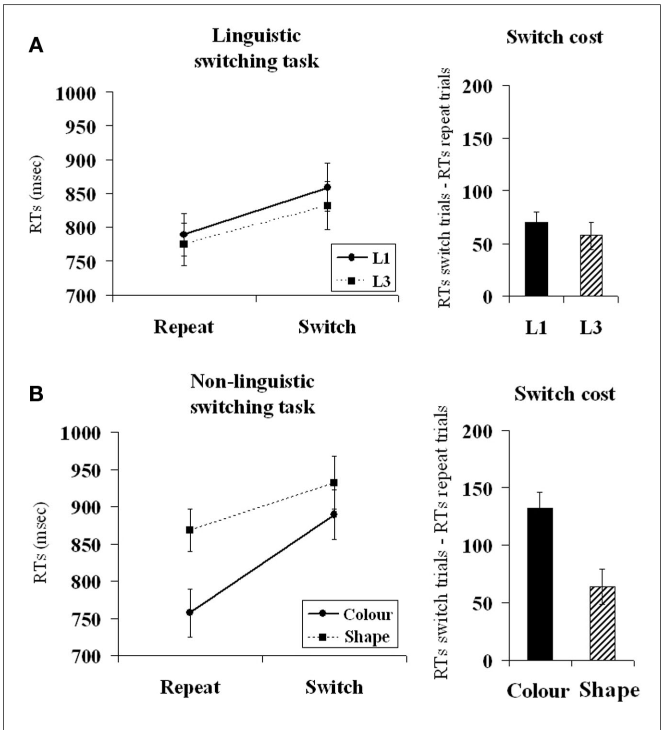


FIGURE 2 | (A) Performances on the linguistic switching task (left) and magnitude of the switch cost for L1 and L3 (right). Error bars represent SE. **(B)** Performances on the non-linguistic switching task (left) and magnitude of the switch cost for color and shape (right). Error bars represent the SE.

MATERIALS AND PROCEDURE

The procedure for the linguistic and non-linguistic switching tasks was the same as that reported for the Experiment 1. The only difference with Experiment 1 was that participants were required to name in Catalan and English, instead of Catalan and Spanish in the language switching task. The material was the same as in Experiment 1.

RESULTS

Linguistic switching cost

The variables considered in the analyses were "type of trial" (switch vs. repeat) and "response language" (L1 and L3), which were included as within-subject factor in a repeated-measure ANOVA on naming latencies.

Reaction times. Overall participants were slower in switch trials (846 ms) compared to repeat trials [783 ms; $F(1, 14) = 75.85$, $MSE = 799.13$, $p < 0.0001$, $\eta_p^2 = 0.84$], but there was no difference in naming latencies between L1 (824 ms) and L3 [804 ms; $F(1, 14) = 2.12$, $MSE = 2914.51$, $p = 0.17$]. The cost to switch to L1 (70 ms) and to L3 (57 ms) was equivalent, as indexed by a non-significant effect of "type of trial" × "response language" interaction [$F(1, 14) = 0.56$, $MSE = 1211.89$, $p = 0.47$; see **Figure 2A**], revealing a symmetrical switch cost.

Accuracy. No difference in accuracy was found between switch and repeat trials [Type of trial: $F(1, 14) = 2.81$, $MSE = 11.99$, $p = 0.12$] and between L1 and L3 [Response language: $F(1, 14) = 0.59$, $MSE = 10.92$, $p = 0.46$]. The interaction between type of trial and response language was not significant either [$F(1, 14) = 0.09$, $MSE = 13.93$, $p = 0.77$; see **Table 3**].

Table 2 | Accuracy (%) and SE in the linguistic and non-linguistic versions of the task-switching broken for trial types for the Experiment 1.

Experiment 1	Accuracy (%)	SE	Accuracy (%)	SE
	L1		**L2**	
LINGUISTIC VERSION				
Repeat	97.8	0.5	97.3	0.6
Switch	96.8	1.0	95.7	1.5
Total	97.3	0.7	96.5	1.0
	Color		**Shape**	
NON-LINGUISTIC VERSION				
Repeat	96.0	0.6	90.9	0.8
Switch	93.5	1.0	89.0	1.9
Total	94.7	0.8	90.0	1.3

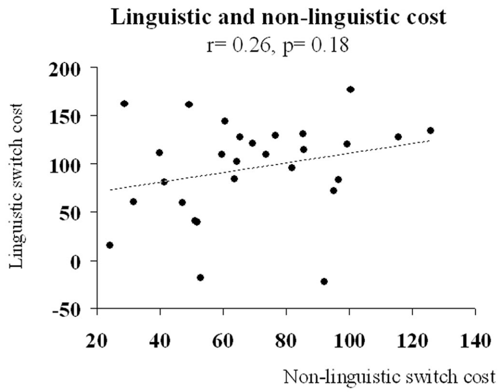


FIGURE 3 | Correlation of individuals' performances between the linguistic and non-linguistic switching tasks, for Experiment 1 and Experiment 2 ($n = 28$). In this graph we excluded one participant from Experiment 1 because his language switching cost was 2 SD above the group's mean.

Table 3 | Accuracy (%) and SE in the linguistic and non-linguistic versions of the task-switching broken for trial types for the Experiment 2.

Experiment 2	Accuracy (%)	SE	Accuracy (%)	SE
	L1		**L3**	
LINGUISTIC VERSION				
Repeat	94.5	1.1	93.4	2.1
Switch	92.6	2.1	92.2	2.1
Total	93.4	1.6	92.4	2.1
	Color		**Shape**	
NON-LINGUISTIC VERSION				
Repeat	96.0	0.8	91.2	0.8
Switch	92.2	1.5	91.7	1.3
Total	93.6	1.1	91.9	1.2

Non-linguistic switching cost

The variables considered in the analysis were "type of trial" (switch vs. repeat) and "sorting criteria" (color and form), which were included as within-subject factors in a repeated-measure ANOVA on the RTs.

Reaction times. Overall participants were slower in switch trials (911 ms) compared to repeat trials [812 ms; $F(1, 14) = 69.38$, $MSE = 2104.36$, $p < 0.0001$, $\eta_p^2 = 0.83$], and faster sorting by color (823 ms) than sorting by shape [900 ms; $F(1, 14) = 42.81$, $p < 0.0001$, $MSE = 2085.37$, $\eta_p^2 = 0.75$]. In this case the switch cost interacted with "type of trial" [$F(1, 14) = 14.11$, $MSE = 1221.76$, $p = 0.002$, $\eta_p^2 = 0.50$]. That is participants showed a larger cost when they switched from shape to color [132 ms, $F(1, 14) = 82.34$, $MSE = 1600.58$, $p < 0.0001$, $\eta_p^2 = 0.85$], than from color to shape [64 ms, $F(1, 14) = 18.22$, $MSE = 1725.55$, $p = 0.001$; see **Figure 2B**].

Accuracy. Participants were less accurate in switch trials (91.9%) than in repeat trials [93.6%; Type of trial: $F(1, 14) = 7.59$, $MSE = 5.54$, $p = 0.01$, $\eta_p^2 = 0.35$], and less accurate to sort by shape (91.4%) than by color [94.1%; $F(1, 14) = 9.44$, $MSE = 11.58$, $p < 0.01$, $\eta_p^2 = 0.40$]. A significant interaction between "type of trial" and "sorting criteria" [$F(1, 14) = 7.38$, $MSE = 9.34$, $p = 0.02$, $\eta_p^2 = 0.34$], indicated an increase of errors when participants switched from shape to color [$F(1, 14) = 12.76$, $MSE = 8.57$, $p < 0.01$, $\eta_p^2 = 0.47$] but not from color to shape [$F(1, 14) = 0.26$, $MSE = 6.31$, $p = 0.62$; see **Table 2**].

To summarize, we found that bilingual participants showed symmetrical switch costs in the linguistic version of the task, but asymmetrical switch costs in the non-linguistic version, as we did in Experiment 1.

Individuals' differences in performance: correlations

Additionally, we used a correlation analysis (Pearson's coefficient) to compare the magnitude of the switch cost between the linguistic and non-linguistic switching tasks.

In fact, if we assume that the switch cost reflects to some extent the efficiency of the bLC and EC in the same way, we may expect that the magnitude of the two switch costs (linguistic and non-linguistic) varies in the same manner in participants.

First, we obtained the correlation coefficient of the total switch cost between the linguistic task and the non-linguistic task (collapsing language in one case and the sorting criteria in the other case). In order to gain more statistical power we ran the analysis with participants of both experiments resulting in a total number of 28 (one participant from Experiment 1 was excluded because his performance was 2 SD above the group means). The switch costs of the two tasks were not significantly correlated ($r = 0.26$, $p = 0.18$; see **Figure 3**).

Then, we tested whether the cost of switching into the easier language (L1) correlated with the cost of switching into the easier sorting criteria (i.e., color), and whether the cost of switching into the difficult language (L2/L3) correlated with the cost of switching into the more difficult sorting criteria (shape). Neither the correlation between the cost of switching to L1 and to color ($r = 0.16$, $p = 0.42$), nor the correlation between the cost of switching to L2/L3 and to shape ($r = -0.15$, $p = 0.44$) were significant.

Exploratory analysis of the switch costs across blocks

Considering the overall results, we found that the switch cost was symmetrical in the linguistic switching task and asymmetrical in the non-linguistic switching task. In a further analysis we explored the pattern of the switch costs across the two experimental blocks with the aim of assessing any potential differences in task adaptation.

To do so we calculated the switch costs separately for the two blocks of the two tasks (linguistic and non-linguistic), containing 160 trials each. In the non-linguistic switching task-switch costs were asymmetrical in both blocks[3] (i.e., switching into color

[3]Non-linguistic switching task. In Experiment 1, the switch costs were 149 ms for color and 34 ms for shape in block 1; 162 ms for color and 24 ms for shape in block 2 [Type of trial × Block interaction: $F(1, 13) = 0.34$, $p = 0.57$]. In Experiment 2, the switch costs were 133 ms for color and 49 for shape in block 1; 134 ms for color

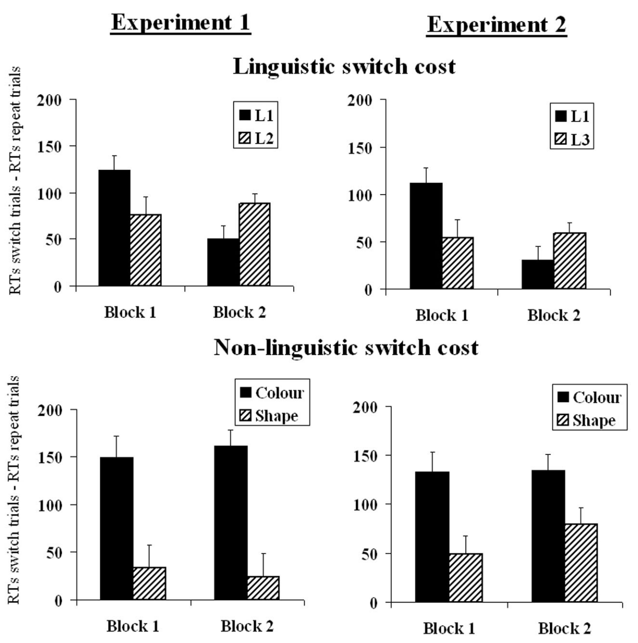


FIGURE 4 | Magnitude of the switch costs in the linguistic and non-linguistic switching tasks broken by blocks and experiments. Error bars represent SE.

was more costly than switching into shape; see **Figure 4**). However, in the linguistic switching task we found a more puzzling result. In the first block, switching into L1 was more costly than switching into L2 or L3, but this pattern reversed in the second block. Interestingly, the cost of switching into L2 or L3 was constant across both blocks, whereas the cost of switching into L1 decreased in the second block. Even though this interaction renders the interpretation of the results more complex, the interesting point here is that it suggests that there are differences between the two types of task-switching also in what regards adaptation to the task.

DISCUSSION AND CONCLUSION

In the present study we examined the relationship between the bLC and EC system. We did so by comparing the pattern of switch costs across linguistic and non-linguistic tasks within a set of highly proficient bilinguals.

We assessed the presence of the symmetrical switch cost in the linguistic task as a starting point, and then we looked at the pattern of switch cost in a non-linguistic switching task. In both experiments, bilinguals showed an asymmetrical non-linguistic switch cost: switching from shape to color was more costly than switching from color to shape. That is, switching from the more difficult task (sorting by shape) to the easier one (sorting by color) resulted in a larger switch cost than vice versa. Additionally, participants committed more errors when they sorted by shape than by color, suggesting that the shape criterion was the most difficult of the two – a finding congruent with previous studies (e.g., Koch, 2001; Martin et al., 2011). In contrast, the same participants showed a symmetrical switch cost in the linguistic task (as previously reported by Costa and Santesteban, 2004; Costa et al., 2006). That is, there seems to be a qualitative difference in the way highly proficient bilinguals perform linguistic and non-linguistic task-switching.

The relationship between the two tasks was also explored by examining the magnitude of the switch costs in the two task versions. The idea behind this analysis was to see whether the efficiency of the bLC abilities could, to some extent, be transferred to the domain-general EC system. Specifically, bilingual individuals that have developed more efficient bLC will probably show relatively small switch costs in the language switching task compared to individuals with less developed bLC. If indeed the bLC functioning depends completely on the EC system, one would expect to find smaller switch costs also in the non-linguistic task. We did not find significant correlations between the linguistic and non-linguistic switch costs, neither between L1 and color nor between L2/L3 and shape. Thus, quantitatively, the magnitude of the switch cost suggests that there is no generalizability from the bLC to the EC system.

Similar results of uncorrelated performance between linguistic and non-linguistic tasks were reported in a study by Bialystok et al. (2008). These authors correlated the performance of bilingual speakers in two language production tasks (fluency and picture naming) with their performance in EC tasks. They did not find any correlation and concluded that their results leave open the possibility that the mechanisms responsible for bLC and those of domain-general EC may have different causes.

Further evidence about differences between the patterns of results in the two versions of the task-switching comes from the different adaptation patterns across the experiment. In the non-linguistic switching task, asymmetrical switching costs (larger switch cost for the easier task) were consistently observed across the whole experiment. However, this was not the case in the language switching task, where a puzzling result was observed. The switch cost for L1, both in Experiment 1 and 2, decreased from block one to block two, whereas the switch cost for L2 and L3 remained constant across blocks. That is, while there is a modulation of the switch cost for the easier task (L1) across the experiment, switch costs for the more difficult task (L2 and L3) remain the same. An interpretation of the L1 adaptation is premature, and future studies need to replicate it. However, our observations highlight the need of exploring language switching costs across the experimental blocks. Besides any kind of interpretation, the interesting point here is that in the two versions of task-switching we found different patterns of switch costs also over time. To some extent, these results indicate that some properties of bLC, for instance a certain degree of flexibility to adapt the behavior, are peculiar to the linguistic domain and they do not transfer to other domains. Once again, this might be evidence for the fact that bLC processes are not fully subsidiary to those

and 79 ms for shape in block 2 [Type of trial × Block interaction: $F(1, 14) = 0.92$, $p = 0.35$]. Linguistic switching task. In Experiment 1, the switch costs were 124 ms for L1 and 76 ms for L2 in block 1; 50 ms for L1 and 88 ms for L2 in block 2 [Type of trial × Block interaction: $F(1, 13) = 19.72$, $p = 0.001$]. In Experiment 2, the switch costs were 112 ms for L1 and 54 for L3 in block 1; 31 ms for L1 and 59 ms for L3 in block 2 [Type of trial × Block interaction: $F(1, 14) = 12.96$, $p = 0.003$].

of the EC system and that there is no transfer from bLC to the domain-general EC system.

Before going into the implications of the results reported here, it is important to note a potential caveat of our study. We have argued that the instantiation of the language switching task in Experiment 2 involves languages of different difficulty, since we compared L1 and L3. In principle, the difference in proficiency between the two languages should be enough to reveal asymmetrical switch costs, as has been shown previously with low-proficient bilinguals (Costa and Santesteban, 2004). However, we do not have any independent evidence that guarantees this difference in proficiency. Indeed, one may be tempted to take the fact that L1 is slower than L3 as an indication against our assumption. However, such interpretation is not without problems. This is because in previous studies we observed a similar pattern of RTs for participants for which we did have independent evidence that L1 was much stronger than L3 (Costa and Santesteban, 2004; Costa et al., 2006). At any rate, we acknowledge that the lack of independent information about the differences in strength between the two languages is a shortcoming of the present study.

The results of the present study suggest that the set of processes engaged in bLC are not fully subsidiary to the domain-general EC processes. That is, a bilingual speaker producing language will not engage the very same set of EC processes that are involved in any other non-linguistic activity in which the executive system is required. As discussed in the Introduction, most of the available evidence from neuroimaging studies is indirect. That is, it is a result of comparing different groups of participants performing either language switching tasks (e.g., Abutalebi and Green, 2007, 2008) or non-linguistic switching tasks (Garbin et al., 2010). One exception is the study of Abutalebi et al. (2011) in which the same group of bilinguals performed a language switching task and a non-linguistic conflict resolution task. The analysis of the brain networks involved in the two tasks showed an overlap over a set of brain areas along the mesial surface, comprising the ACC (BA 32) and the pre-SMA (BA 6). However, some additional areas were recruited during the conflict resolution task that were not active during the language switching task. Thus, the general conclusion from the neuroimaging literature is that some brain areas of the bLC and EC overlap, but the small amount of direct evidence (e.g., the same group of participants tested both on linguistic and non-linguistic tasks involving EC) precludes us from drawing strong conclusions about the extent of this overlap.

Our results fit well with data on brain-damaged individuals. Studies testing bilingual aphasics have reported double dissociations between language control and domain-general control (e.g., Green et al., 2010; see also Abutalebi et al., 2000; Mariën et al., 2005). For example, in Green et al. (2010) found a relatively different impairment of language control and the EC system as a result of the brain lesion, indicating that the brain areas implicated in language control are not totally subsidiary to those implicated in EC and vice versa.

Before concluding, it is worth mentioning the lack of a correlation observed between the magnitudes of the switch costs in the linguistic and non-linguistic tasks. This also points to the direction that one cannot equate the processes involved in bLC with those involved in domain-general EC system. This approach, in which the crosstalk between bLC and EC is assessed in the same group of bilinguals by comparing the magnitude of switch costs, has started to receive some attention. Recently, Prior and Gollan (2011) looked at this issue by testing whether the bilingual advantage in EC was to some extent related to the cost of language switching. They found that those bilinguals who showed less cost in task-switching were also those who showed less cost in language switching. But this was true only for those bilinguals who reported to switch quite often in their everyday life. Second, no direct correlations of the switch costs between the two tasks were performed within the group of participants. Therefore, only based on the results of Prior and Gollan (2011) it is premature to conclude that the mechanisms underlying bLC are fully subsidiary to the EC system. And, in fact, if anything our results indicate otherwise.

To conclude, in this study we found different patterns of switch costs between a language switching task and a non-linguistic switching task. These results suggest that even if there is crosstalk between bLC and domain-general EC, there are some aspects of the bLC system that are specific to the domain of language and not necessarily related to the EC system. The relevance of our results is that they represent an attempt to investigate the crosstalk between the bLC and EC in the same group of participants. Further research is needed to investigate the exact mechanisms underlying the bLC and EC systems in bilinguals in order to eventually gain knowledge about their functional and neural relationship.

ACKNOWLEDGMENTS

This work was supported by grants from the Spanish government (PSI2008-01191, Consolider Ingenio 2010 CSD2007-00012) and the Catalan government (Consolidat SGR 2009-1521). Marco Calabria was supported by a postdoctoral fellowship from the Spanish Government (Juan de la Cierva fellowship). Francesca M. Branzi was supported by a predoctoral fellowship from the Spanish Government (FPU-2009-2013). This research was also supported by a Marie Curie International Outgoing Fellowship within the seventh European Community Framework Programme awarded to Mireia Hernández. The authors are grateful to Ian FitzPatrick, Philip Jaekl, and Elin Runnqvist for comments on previous versions of this manuscript.

REFERENCES

Abutalebi, J., Annoni, J. M., Zimine, I., Pegna, A. J., Seghier, M. L., Lee-Jahnke, H., Lazeyras, F., Cappa, S. F., and Khateb, A. (2008). Language control and lexical competition in bilinguals: an event-related FMRI study. *Cereb. Cortex* 18, 1496–1505.

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurol.* 20, 242–275.

Abutalebi, J., and Green, D. (2008). Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang. Cogn. Process.* 23, 557–582.

Abutalebi, J., Miozzo, A., and Cappa, S. (2000). Do subcortical structures control "language selection" in polyglots? Evidence from pathological language mixing. *Neurocase* 6, 51–56.

Abutalebi, J., Rosa, P. A., Green, D., Hernández, M., Scifo, P., Keim, R., Cappa, S. F., and Costa, A. (2011). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb. Cortex.* doi: 10.1093/cercor/bhr287

Allport, A., and Wylie, G. (2000). "Task-switching, stimulus–response bindings, and negative priming," in *Attention and Performance XVIII: Control of Cognitive Processes*, eds S. Monsell and J. S. Driver (Cambridge, MA: MIT Press), 35–70.

Allport, D. A., Styles, E. A., and Hseih, S. (1994). "Shifting intentional set: exploring the dynamic control of tasks," in *Attention and Performance XV: Conscious and Nonconscious Information Processing*, eds C. Umiltà and M. Moscovitch (Cambridge, MA: MIT Press), 421–452.

Bialystok, E., Barac, R., Blaye, A., and Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *J. Cogn. Dev.* 11, 485–508.

Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873.

Bialystok, E., Craik, F. I., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., Craik, F. I., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354.

Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., and Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402, 179–181.

Bryck, R. L., and Mayr, U. (2008). Task selection cost asymmetry without task s witching. *Psychon. Bull. Rev.* 15, 128–134.

Christoffels, I. K., Firk, C., and Schiller, N. O. (2007). Bilingual language control: an event-related potential study. *Brain Res.* 1147, 192–208.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511.

Costa, A., Santesteban, M., and Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and Language-Specific Selection mechanisms are both functional. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1057–1074.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green, D. W., and Price, C. J. (2006). Language control in the bilingual brain. *Science* 312, 1537–1540.

Crone, E. A., Wendelken, C., Donohue, S. E., and Bunge, S. A. (2006). Neural evidence for dissociable components of task-switching. *Cereb. Cortex* 16, 475–486.

Duncan, J. (1986). Disorganization of behaviour after frontal-lobe damage. *Cogn. Neuropsychol.* 3, 271–290.

Forster, K. I., and Forster, J. C. (2003). A Windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124.

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., Hernández, M., Costa, A., and Avila, C. (2010). Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. *Neuroimage* 53, 1272–1278.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Green, D. W., Grogan, A., Crinion, J., Ali, N., Sutton, C., and Price, C. J. (2010). Language control and parallel recovery of language in individuals with aphasia. *Aphasiology* 24, 188–209.

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

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

Hernández, M., Costa, A., Fuentes, L., and Vivas, A. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Biling. Lang. Cogn.* 13, 315–325.

Hervais-Adelman, A. G., Moser-Mercer, B., and Golestani, N. (2011). Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. *Front. Psychol.* 2:234. doi:10.3389/fpsyg.2011.00234

Koch, I. (2001). Automatic and intentional activation of task sets. *J. Exp. Psychol. Learn. Mem. Cogn.* 27, 1474–1486.

Koch, I., Gade, M., Schuch, S., and Philipp, A. M. (2010). The role of inhibition in task switching: a review. *Psychon. Bull. Rev.* 17, 1–14.

Mariën, P., Abutalebi, J., Engelborghs, S., and De Deyn, P. P. (2005). Pathophysiology of language switching and mixing in an early bilingual child with subcortical aphasia. *Neurocase* 11, 385–398.

Martin, C. D., Barceló, F., Hernández, M., and Costa, A. (2011). The time course of the asymmetrical "local" switch cost: evidence from event-related potentials. *Biol. Psychol.* 86, 210–218.

Mayr, U., and Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1124–1140.

Meiran, N. (1996). Reconfiguration of processing mode to task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 1423–1442.

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40.

Meyer, D. E., and Kieras, D. E. (1997). EPIC – a computational theory of executive cognitive processes and multiple-task performance: part 1. Basic mechanisms. *Psychol. Rev.* 104, 3–65.

Monsell, S. (2003). Task switching. *Trends Cogn. Sci. (Regul. Ed.)* 7, 134–140.

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., Konishi, J., Fukuyama, H., and Shibasaki, H. (2001). Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cereb. Cortex* 11, 85–92.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-Regulation*, Vol. 4, eds R. J. Davidson, G. E. Schwartz, and D. Shapiro (New York: Plenum), 1–18.

Price, C. J., Green, D. W., and Von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain* 122, 2221–2235.

Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish-English and Mandarin-English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 1–10.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. Lang. Cogn.* 13, 253–262.

Protopapas, A. (2007). Checkvocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behav. Res. Methods* 39, 859–862.

Rubinstein, J. S., Meyer, D. E., and Evans, J. E. (2001). Executive control of cognitive processes in task switching. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 763–797.

Schneider, D. W., and Anderson, J. R. (2010). Asymmetric switch costs as sequential difficulty effects. *Q. J. Exp. Psychol.* 63, 1873–1894.

Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, familiarity and visual complexity. *J. Exp. Psychol. Learn. Mem. Cogn.* 6, 174–215.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 June 2011; accepted: 30 December 2011; published online: 13 January 2012.

Citation: Calabria M, Hernández M, Branzi FM and Costa A (2012) Qualitative differences between bilingual language control and executive control: evidence from task-switching. Front. Psychology ***2****:399. doi: 10.3389/fpsyg.2011.00399*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 09 August 2011
doi: 10.3389/fpsyg.2011.00184

Attentional inhibition in bilingual naming performance: evidence from delta-plot analyses

Ardi Roelofs*, Vitória Piai and Gabriela Garrido Rodriguez

Donders Centre for Cognition, Radboud University Nijmegen, Nijmegen, Netherlands

Edited by:
Ingrid Christoffels, University of Leiden, Netherlands

Reviewed by:
K. Richard Ridderinkhof, University of Amsterdam, Netherlands
Julia Festman, Otto-von-Guericke University Magdeburg, Germany

****Correspondence:***
Ardi Roelofs, Centre for Cognition, Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Spinoza Building B.01.08, Montessorilaan 3, 6525 HR Nijmegen, Netherlands.
e-mail: a.roelofs@donders.ru.nl

It has been argued that inhibition is a mechanism of attentional control in bilingual language performance. Evidence suggests that effects of inhibition are largest in the tail of a response time (RT) distribution in non-linguistic and monolingual performance domains. We examined this for bilingual performance by conducting delta-plot analyses of naming RTs. Dutch–English bilingual speakers named pictures using English while trying to ignore superimposed neutral Xs or Dutch distractor words that were semantically related, unrelated, or translations. The mean RTs revealed semantic, translation, and lexicality effects. The delta plots leveled off with increasing RT, more so when the mean distractor effect was smaller as compared with larger. This suggests that the influence of inhibition is largest toward the distribution tail, corresponding to what is observed in other performance domains. Moreover, the delta plots suggested that more inhibition was applied by high- than low-proficiency individuals in the unrelated than the other distractor conditions. These results support the view that inhibition is a domain-general mechanism that may be optionally engaged depending on the prevailing circumstances.

Keywords: attention, bilingualism, delta plots, inhibition, naming, response times

INTRODUCTION

Attentional control includes the ability to formulate goals and plans of action and to follow these while facing distraction. This ability is critical to normal human functioning and it is a hallmark of general intelligence (e.g., Wundt, 1904; Duncan, 2010). Attentional control plays a central role in human performance generally and language performance specifically (e.g., Roelofs, 2003, 2008). Bilingual language performance is an instance of powerful attentional control in a naturalistic situation. Although bilingual speakers can usually choose from at least two words for any given concept (i.e., one in each language), they are able to restrict their utterances to one language only. Even non-balanced bilinguals, whose first language (L1) is stronger than their second language (L2), can speak one language without apparently being much hampered by the other language. However, little is known about the mechanisms of attentional control in bilingual performance. The aim of the present article is to illuminate properties of these mechanisms. A better understanding of how bilinguals achieve attentional control over their languages will be informative not only regarding bilingual language performance, but also regarding efficient attentional control in general.

INHIBITION IN BILINGUAL PERFORMANCE

A prominent account of attentional control in bilingual language performance holds that inhibition is involved (Green, 1998). The inhibition is attentional, because it concerns a top-down goal-dependent modulation of language processes rather than a type of inhibition that is evoked bottom up by language perception (such as lateral inhibition present in several models of word recognition), see Aron (2007) for an extensive discussion. The notion of attentional inhibition has a long history in psychology. In the early days of experimental psychology, Wundt (the founder of modern scientific psychology and psycholinguistics) assumed that an attentional inhibition mechanism, which he located in the frontal lobes of the human brain, modulates activity in a language network, assumed to be centered around perisylvian brain areas (e.g., Wundt, 1904). In contemporary psychology, inhibition has been proposed as a compulsory (Abutalebi and Green, 2007; Kroll et al., 2008) or an optional mechanism (Verhoef et al., 2009) of attentional control in bilingual language performance, although others argue against the assumption of inhibition (Finkbeiner et al., 2006).

The compulsory inhibition hypothesis holds that attentional control over the languages of a bilingual speaker is achieved by inhibiting the irrelevant language. Green (1998) assumed that inhibition in bilingual performance is reactive, that is, evoked in response to lexical activation. Consequently, the amount of inhibition that is applied depends on the magnitude of lexical activation in the non-target language. According to the hypothesis that inhibition is an option, inhibition is not the mechanism that achieves attentional control over the languages, but inhibition may be optionally engaged to increase the speed and accuracy of bilingual performance, depending on the prevailing circumstances (Verhoef et al., 2009).

Regardless of whether inhibition is compulsory or optional, little is known about the nature of the inhibitory mechanism. One possibility is that inhibition is a domain-general mechanism that is shared between linguistic and non-linguistic performance, as proposed by Abutalebi and Green (2007). If so, inhibition in bilingual language performance should share critical properties, such as its dynamics, with inhibition in other performance domains.

DYNAMICS OF INHIBITION IN SIMON, ERIKSEN, AND STROOP TASK PERFORMANCE

As concerns the dynamics of inhibition in non-linguistic domains, Ridderinkhof and colleagues (e.g., Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005) argued that attentional inhibition takes time to build up. Typically, in a situation with targets and distractors, interference from the distractors increases monotonically with response time (RT). Inhibition can decrease interference, but because inhibition is reactive and builds up slowly, its effects will be stronger for longer RTs. Consequently, effects of differential inhibition should be largest toward the tail of an RT distribution.

To assess the effect of differential inhibition, Ridderinkhof et al. (2004, 2005) constructed delta plots, which map out distractor condition differences as a function of RT. Delta plots prototypically have a positive slope, reflecting that the effect of an experimental factor tends to increase as a function of RT (cf. Luce, 1986). That is, the magnitude of factor effects tends to be larger for long than short RTs. However, if conditions differ in the amount of inhibition that is applied, a difference in RT between conditions should not increase linearly as a function of RT, but instead level off and become reduced for slow responses. If larger inhibition results in a more pronounced reduction of condition effects in slow responses, then the leveling off of the delta plot should be more pronounced in conditions that involve more inhibition. That is, the flattening of the delta plot should be more pronounced in experimental conditions that require more stringent inhibition compared to conditions that are less demanding.

Figure 1 illustrates the delta-plot logic. Suppose an experiment includes two conditions, one with conflict (red line) and one without conflict (green line). The left-hand panel of **Figure 1** illustrates the cumulative distributions of the two conditions by plotting the mean RT for each condition per quintile. The horizontal difference between the two distributions represents the difference in interference between the conditions. Typically, interference increases with RT. That is, the difference between the two curves becomes larger for higher quintiles. The right-hand panel of **Figure 1** shows the corresponding delta plots, which plot these condition differences per quintile against the mean of the two conditions for the corresponding quintile. The upper curve shows the situation when no inhibition is applied to resolve the conflict: The interference becomes larger with increasing quintile. The middle curve shows that with weak inhibition, the interference decreases and the difference in RT between the two conditions becomes smaller. However, because inhibition is reactive and builds up slowly, its effect is stronger for longer RTs. Interference will therefore tend to level off for the higher quintiles of the delta plot. The lower curve shows that strong inhibition will decrease the interference further, which may even yield negative slopes for the segments connecting the higher quintiles (e.g., segment q4–5 connecting quintiles 4 and 5). To conclude, differences in delta plots index differences in the amount of inhibition that is applied in experimental conditions.

Ridderinkhof and colleagues (Ridderinkhof, 2002; Ridderinkhof et al., 2005) provided evidence from RT distributional analyses of non-linguistic Simon and Eriksen task performance that delta plots leveled off more with increased inhibition (for reviews, see Proctor et al., 2011; Van den Wildenberg et al., 2011). In the Simon task, participants indicate the identity of a left or right stimulus by pressing a left or right button, whereby the left–right position of the stimulus on a trial can be congruent (e.g., left stimulus requiring a left button press) or incongruent (e.g., left stimulus requiring a right button press). RTs are typically longer on incongruent than congruent trials, called the Simon effect. Ridderinkhof (2002) observed that delta plots leveled off more in participants with relatively small Simon effects (presumed to reflect strong inhibition) than in participants with relatively large Simon effects (presumed to reflect weaker inhibition). In an arrow version of the Eriksen task, participants have to indicate the identity of a left- or right-pointing target arrow flanked by incongruent or congruent distractor arrows on each side (e.g.,>><>> or <<<<<, respectively). RTs are typically longer on incongruent than congruent trials, called the Eriksen flanker effect. Ridderinkhof et al. (2005) observed that the leveling off of the delta plot for the Eriksen flanker effect was more pronounced for normal participants than for participants

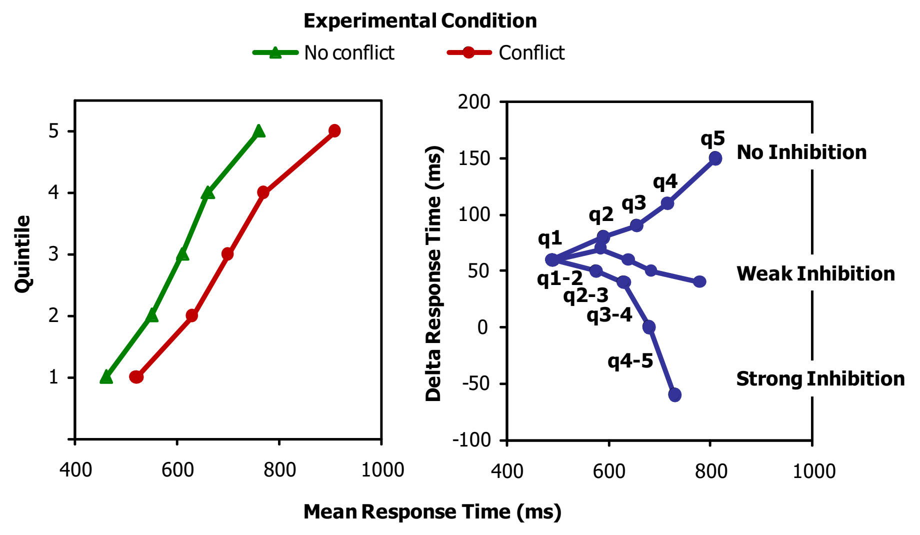


FIGURE 1 | Left panel: cumulative distribution curves for response times in experimental conditions with and without conflict. Right panel: delta plots showing the condition differences (deltas) as a function of quintile (1–5) and amount of inhibition (no, weak, strong). q1, quintile 1, and so forth; q1–2 is the segment connecting quintiles 1 and 2, etc.

diagnosed with ADHD, which were assumed to have an inhibition deficit. The results from the distributional analyses of RTs in Simon and Eriksen task performance support the claim of Ridderinkhof and colleagues that inhibition takes time to develop during a trial in non-linguistic domains.

Similar results have been obtained in monolingual Stroop task performance. In the color–word Stroop task, individuals name the ink color of printed congruent or incongruent color words (e.g., the words GREEN or RED in green color; say "green") or neutral Xs. Mean RT is typically longer on incongruent than neutral trials, and often shorter on congruent than neutral trials. In performing delta-plot analyses on Stroop color naming RTs, Bub et al. (2006) obtained evidence that younger children (7–9 years) engage in stronger inhibition than older children (9–11 years), indicating that inhibition may also be engaged in non-linguistic domains. Moreover, delta-plot analyses of RTs in a manual version of the Stroop task by Sharma et al. (2010) revealed that the presence of a passively observing confederate during task performance leads to stronger inhibition as compared with the absence of a confederate. This influence of social context was observed when the preparation interval between consecutive Stroop stimuli was long (1000 ms), but not when the interval was short (32 ms). These results suggest that Stroop task performance may engage an inhibition mechanism, whose effects build up slowly, in line with the observations of Ridderinkhof et al. (2004, 2005). The effect of social context (Sharma et al., 2010) suggests that inhibition is an optional rather than compulsory mechanism.

Forstmann et al. (2008a) observed a strong link between individual differences in delta-plot parameters for Simon task performance and activity in right inferior frontal cortex (see Aron et al., 2004, for a review of the imaging literature localizing inhibition to right inferior frontal cortex). Moreover, when individual RT distribution parameters were used to classify subgroups of good and poor inhibitors based on a median split of the slowest segment of the delta plots, it was observed that individuals with better inhibition abilities showed higher brain connectivity values for white matter tracts in right inferior frontal cortex than poorer inhibitors. These results corroborate the assumption that delta plots reflect the operation of an attentional inhibition mechanism whose influence builds up slowly and is implemented in right inferior frontal cortex (see also Forstmann et al., 2008b), in line with Wundt's (1904) suggestion.

DYNAMICS OF INHIBITION IN BILINGUAL PICTURE NAMING

Evidence that attentional inhibition builds up slowly in bilingual language performance was provided by Verhoef et al. (2009). They measured event-related electrical brain potentials (ERPs) while Dutch–English bilingual participants named pictures in their first or second language, whereby the naming language on consecutive trials could be the same or different. The target language was indicated by a cue that preceded the picture by a short (500 ms) or long (1250 ms) preparation interval. In the ERPs time-locked to the picture onset, Verhoef et al. (2009) observed a right-lateralized N2 effect, which was argued to reflect activity in right inferior frontal cortex and to index inhibition. The N2 magnitude was modulated by the size of the preparation interval (larger N2 for long than short intervals) but not by language (L1 vs. L2) or language switching. So, the experiment provides evidence for an effect of time to prepare on the amount of inhibition applied during the planning of the picture name. This suggests that inhibition in picture naming takes time to develop, in line with the observations of Ridderinkhof et al. (2004, 2005). Given the timing evidence and the evidence that the N2 was frontal right-lateralized, suggesting a right inferior frontal locus, there is a clear similarity with the claims concerning the time course and right inferior frontal locus of the inhibition derived from delta plots. The effect of preparation interval corresponds to what Sharma et al. (2010) observed for manual Stroop task performance. The absence of an effect of language and switching on the N2 suggests that inhibition is an optional rather than compulsory mechanism.

Whereas Verhoef et al. (2009) argued that inhibition is used depending on the prevailing circumstances, Costa and Santesteban (2004) argued that the use of inhibition depends on language proficiency. According to them, attentional control over the two languages is accomplished by different mechanisms in linguistically balanced (i.e., equal proficiency in L1 and L2) and non-balanced bilinguals. In their view, while non-balanced bilinguals use inhibition of the non-target language to speak the target language, balanced bilinguals do not need to recruit inhibition since they have developed a mechanism allowing language-specific lexical access. This claim of Costa and Santesteban (2004) was based on their observation of asymmetrical RT switch costs in language switching for non-balanced bilinguals and symmetrical switch costs for balanced bilinguals in picture naming. An RT switch cost means longer RTs when the language of the current trial is different from that of the previous trial (switch trials) than when the language is the same (repeat trials). Costa and Santesteban (2004) observed that balanced Spanish–Catalan speakers had equal switch costs in picture naming for switching to L1 and to L2, whereas non-balanced bilinguals had larger switch costs for switching to their L1 than to their L2. This finding suggests that for non-balanced bilinguals, using L2 involves stronger inhibition of L1 than using L1 involves inhibition of L2. Consequently, it will take longer to overcome the previous inhibition in switching to L1 than to L2.

However, Verhoef et al. (2009) observed that both asymmetrical and symmetrical RT switch cost patterns may occur as a function of preparation interval in a single population of non-balanced Dutch–English bilinguals. In their study, short preparation intervals elicited asymmetrical switch costs and long preparation intervals elicited symmetrical switch costs. This suggests that the engagement of inhibition may counteract the negative effect of lower proficiency in L2 than L1, provided that the preparation interval is long enough for inhibition to be engaged. The engagement of inhibition during the long preparation interval was reflected in the right-lateralized N2.

OUTLINE OF THE PRESENT STUDY

We report a bilingual picture–word interference study in which we further examined the nature of inhibition in bilingual language performance. In particular, we tested the prediction derived from the claim by Ridderinkhof et al. (2004, 2005) that effects of differential inhibition should be largest in the tail of an RT distribution. In our study, Dutch–English non-balanced bilingual participants named target pictures in their second language English (L2) while trying to ignore written Dutch distractor words (L1) superimposed onto the pictures. The target pictures and distractor words were semantically related (e.g., picture of a rabbit, say "rabbit"; distractor word HERT, *deer*), unrelated (e.g., word STOEL, *chair*), or translation equivalents (e.g., KONIJN, *rabbit*). In addition, a series of Xs was superimposed on the picture in the control condition. Previous research has demonstrated semantic interference effects (i.e., longer RTs on semantically related than unrelated trials), lexicality effects (i.e., longer RTs on unrelated than control trials), and translation effects (i.e., longer RTs on unrelated than translation trials), see Costa et al. (1999) and Hermans et al. (1998), among others.

The interference effects from distractor words suggest that there are differential needs for inhibition on the different trial types. In particular, if the amount of inhibition depends on the magnitude of lexical activation (cf. Green, 1998), more inhibition is required on semantically related than unrelated trials and on unrelated than neutral and translation trials. If inhibition is a mechanism of attentional control in bilingual performance, and inhibition is a domain-general mechanism with a dynamics as assessed by Ridderinkhof et al. (2004, 2005), then the differential need for inhibition on the different trial types should be reflected in the tail of the corresponding RT distributions. In particular, the magnitude of the semantic, lexicality, and translation effects should decrease with increasing RT, more so with smaller effect sizes (presumed to reflect strong inhibition) than with larger effect sizes (presumed to reflect weaker inhibition), following Ridderinkhof (2002). In contrast, if inhibition is not applied in the present study or has different dynamic properties than inhibition in other performance domains (e.g., in the Simon, Eriksen, and Stroop tasks), then condition differences in RT should monotonically increase with increasing RT, and delta plots should not level off differentially among trial types depending on mean distractor effect.

In addition, we compared the delta plots between high- and low-proficiency individuals. It has been argued that bilingual individuals have enhanced inhibition abilities compared to monolingual individuals, as assessed for the non-linguistic Simon task (Bialystok et al., 2004) and the Eriksen flanker task (e.g., Costa et al., 2008). The idea is that the constant need to inhibit one of the two languages (i.e., the non-target one) provides bilinguals with an enhanced ability to ignore distracting and irrelevant stimuli, not only in linguistic tasks, but also in non-linguistic ones. However, Colzato et al. (2008) found no evidence for enhanced inhibition when comparing monolingual and bilingual individuals with regard to stop signal performance, inhibition of return, and the attentional blink, which all three can be argued to tap into aspects of inhibition. This suggests that the enhanced inhibition is restricted to Stroop-like circumstances, such as present in the Simon and Eriksen tasks. According to Green (1998), lexical competitors are more highly activated for high-proficiency as compared with low-proficiency bilinguals, and so these lexical competitors require a greater degree of inhibition. This implies that the inhibition ability develops along with skill in L2. If the inhibition ability depends on the level of skill in L2 and is reflected in performance on Stroop-like tasks, one would expect enhanced inhibition for high- compared with low-proficiency individuals in the present picture–word interference experiment. In particular, the delta plots should level off more for high- than low-proficiency individuals. Furthermore, if delta plots level off differentially depending on effect type (i.e., semantic, translation, lexicality), this would suggest that individuals engage inhibition differently depending on the prevailing circumstances.

MATERIALS AND METHODS

PARTICIPANTS

The experiment was carried out with a group of 16 students of Radboud University Nijmegen, the Netherlands. Their mean age was 23.3 years (SD = 2.86). All participants were native speakers of Dutch with normal or corrected-to-normal vision, who learned English as a second language. The students participated in return of either payment or course credits.

A 17-item self-rating questionnaire was used to obtain proficiency scores. This questionnaire was composed of three main parts. In the first section (four items), participants indicated how well their English (L2) skills (reading, speaking, writing, and listening) were compared to Dutch (L1). The scores were on a five point scale, in which 1 represents that L2 skills were just as good as L1 skills and 5 represents that L2 skills were worse than L1 skills. On average, participants rated their proficiency for L2 compared to L1 as 2.94 (SD = 0.84). Thus, the participants were bilingually non-balanced. The second section (eight items) measured participants' use of L2 in different situations. Scores for L2 use were also measured at a five point scale, where 1 represents less than 1 h per week and 5 represents more than 10 h per week. On average, participants L2 use score was 1.92 (SD = 0.76). The last part of the questionnaire (five items) evaluated age of onset (which refers to the age at which participants started learning English) and number of years of L2 use. The mean onset of L2 was at 10.63 years of age (SD = 1.86) and L2 was used, on average, for 12.65 years (SD = 3.46).

MATERIALS AND DESIGN

From the picture gallery of the Max Planck Institute for Psycholinguistics (Nijmegen, the Netherlands), 32 pictured objects from eight different semantic categories (i.e., clothing, animals, transportation, buildings, weapons, service, furniture, and body parts) were selected together with their basic-level names in English and Dutch. **Table 1** lists the English target picture names and the Dutch translation equivalents that were used as distractor words. These pictures and distractors were chosen because they yielded clear semantic, lexicality, and identity effects in earlier studies (e.g., Roelofs, 2003, 2008; Roelofs and Verhoef, 2006). We tried to avoid the use of cognates. The pictures were white line drawings on a black background and they were digitized and scaled to fit into a virtual frame of 10 cm × 10 cm. The printed words were presented in white color in 36-point lowercase Arial font.

There was one independent variable concerning the picture–word stimuli, referred to as distractor type, with four levels: semantically related, unrelated, translation, and control. Each picture was combined with a distractor from the same semantic category (the semantically related condition), with a word from another semantic category (the unrelated condition), with a word that was the Dutch translation equivalent of the English picture name (the translation condition), or with a string of Xs (the control condition). The distractor conditions were created by recombining pictures and words. For example, the picture of a car (say "car") was combined with the Dutch word FIETS (*bicycle*) in the semantic condition,

Table 1 | Basic-level names of the pictures in English (the target language) and their Dutch translation equivalents.

English name	Dutch name	English name	Dutch name
car	auto	cup	beker
bicycle	fiets	plate	bord
airplane	vliegtuig	bowl	kom
truck	vrachtwagen	jug	kan
toe	teen	coat	jas
leg	been	sweater	trui
nose	neus	skirt	rok
ear	oor	dress	jurk
deer	hert	castle	kasteel
swan	zwaan	mill	molen
rabbit	konijn	factory	fabriek
turtle	schildpad	church	kerk
table	tafel	dagger	dolk
cupboard	kast	sword	zwaard
desk	bureau	rifle	geweer
chair	stoel	tomahawk	bijl

with the Dutch word STOEL (*chair*) in the unrelated condition, with the Dutch translation equivalent AUTO (*car*) in the translation condition, and the Xs in the control condition. All target pictures and distractor words occurred equally often in each distractor type condition and they were repeated three times, yielding 384 trials in total. The order of presenting the stimuli across trials was randomized for each participant.

PROCEDURE AND APPARATUS

The participants were tested individually. They were seated in front of a CRT computer monitor and a microphone connected to an electronic voice key. The distance between participant and screen was approximately 70 cm, and the distance between participant and microphone was approximately 18 cm. Before the experiment began, participants were given a booklet that contained the set of experimental pictures and their names. They were asked to go through it in order to be familiarized with the pictures and their appropriate English names. After a participant had read the instructions, a block of 32 practice trials was administered in which the experimental pictures, combined with a row of Xs, were presented once and named in English. After this, testing began. A trial started with the presentation of a picture combined with a Dutch distractor word or the Xs for 250 ms followed by a black screen that lasted 2250 ms. The pictures were presented in the center of the screen and the distractor words were presented in the center of the pictures. The vocal response latency was measured to the nearest millisecond from target stimulus onset (with a time-out of 2500 ms). The presentation of stimuli and the recording of responses were controlled by Presentation Software (Neurobehavioral Systems, Albany, CA, USA).

DATA ANALYSIS

For each trial, the experimenter coded the response for errors. Five types of incorrect responses were distinguished: wrong pronunciation of the word, wrong response word (e.g., the response word was given in Dutch instead of English), disfluency, voice key triggered by a non-speech sound, and recording failures. Incorrect responses and RTs below 100 ms were discarded from the statistical analyses of the RTs. The mean RTs were submitted to analyses of variance. The analyses were performed both by participants (F_1) and by items (F_2). Distractor type was tested within participants and within items. In addition, the responses coded as correct or incorrect were submitted to binomial logistic regression analysis with distractor type as predictor (cf. Jaeger, 2008). An alpha level of 0.05 was used for all statistical tests. Following common parlance in the literature, we refer to statistical findings with *p*-values of between 0.05 and 0.10 as "marginally significant."

To obtain the delta plots, the RTs for each participant and distractor type were rank ordered and divided in five RT quintiles (bins) of equal or near-equal size. Next, the mean RT was determined for each quintile in each distractor condition. The delta plots for the semantic, translation, and lexicality effects were obtained by computing, for each quintile, the RT difference between, respectively, the semantically related and unrelated conditions (the semantic effect), unrelated and translation conditions (the translation effect), and the unrelated and control conditions (the lexicality effect).

To assess whether more inhibition leads to smaller distractor effects, we computed the magnitude of the semantic, translation, and lexicality effects for each participant. Next, median splits were made based on distractor effect sizes, referred to as *smaller* and *larger*, and delta plots were derived for each group (smaller, larger) and type of effect (i.e., semantic, translation, and lexicality). Similarly, median splits were made on the basis of proficiency (i.e., the mean scores for the self-ratings of skill in L2 in the first section of the questionnaire), referred to as *higher* vs. *lower*, and delta plots were derived for each group (higher, lower) and type of effect (i.e., semantic, translation, and lexicality). Because of equal skill scores for some participants, the median split yielded two proficiency groups of unequal size: There were nine participants in the high-proficiency group and seven participants in the low-proficiency group. The mean proficiency scores of the high- and low-proficiency groups were 2.31 and 3.75, respectively, which differed significantly, $F(1,14) = 46.52$, $p < 0.001$.

To assess whether the delta plots were different depending on mean distractor effect (smaller, larger) or proficiency (higher, lower), analyses of variance were performed on the slopes of the delta plots as a function of mean distractor effect size and proficiency, following Ridderinkhof and colleagues (Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005), Bub et al. (2006), and Sharma et al. (2010). For these analyses, slopes were computed for the delta-plot segments connecting the data points of consecutive quintiles (q1–2, q2–3, q3–4, and q4–5). The slope of the line segment connecting quintiles 1 and 2 was defined as the delta of quintile 2 minus that of quintile 1 divided by the difference in mean RT (across the two conditions used to compute the delta value) between quintile 2 and quintile 1 (e.g., De Jong et al., 1994; Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005). In a similar manner, the slopes of the other segments were determined. To assess whether the slopes were different depending on mean distractor effect or proficiency, analyses of variance were conducted on the slopes of consecutive quintile pairs with the within-participants factor quintile pair (q1–2, q2–3, q3–4, q4–5) and the between-participants factor effect size (smaller, larger) or proficiency (lower, higher). Following earlier research, the delta plots were constructed such that the RT values on the horizontal axis were the means of the RTs in the two conditions used to compute each delta value (De Jong et al., 1994; Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005; Bub et al., 2006; Sharma et al., 2010).

RESULTS

MEAN PERFORMANCE

Table 2 presents the mean RTs for correct trials, their SD, and mean percentages of errors for each distractor type. RTs were longer for the semantically related and unrelated conditions compared with the translation and control conditions. Interference was found for semantically related distractors relative to unrelated ones (the mean semantic interference effect was 51 ms), and for unrelated distractors relative to control distractors (the mean lexicality effect was 69 ms). Moreover, interference was found for the unrelated condition relative to the translation condition (the mean translation effect was 95 ms). More errors were made when participants had to name a picture combined with a semantically related distractor than when the picture was presented combined with any of the other distractors.

The statistical analysis of the naming RTs yielded main effects of distractor type, $F_1(3,45) = 46.80$, $p < 0.001$, $\eta_p^2 = 0.76$; $F_2(3,93) = 53.97$, $p < 0.001$, $\eta_p^2 = 0.64$. Pairwise comparisons showed significant results for the semantic interference effect, $t_1(15) = 4.71$, $p < 0.001$, $t_2(31) = 4.02$, $p < 0.001$; for the translation effect, $t_1(15) = 5.46$, $p < 0.001$, $t_2(31) = 8.19$, $p < 0.001$; and for the lexicality effect, $t_1(15) = 5.05$, $p < 0.001$, $t_2(31) = 6.90$, $p < 0.001$.

Logistic regression analyses of the error percentages revealed that the log-odds of having a correct response in the unrelated condition were 1.71 higher compared to giving a correct response in the semantically related condition (the semantic effect), β coefficient = 0.54, SE = 0.15, Wald $Z = 3.64$, $p < 0.001$. When a translation equivalent was presented as distractor, the log-odds of having a correct response in comparison to an unrelated distractor (i.e., the translation effect) were 1.85 times higher, β coefficient = 0.62, SE = 0.19, Wald $Z = 3.19$, $p = 0.001$. Finally, when a control distractor was presented, the log-odds of having a correct response were 1.47 times higher than when an unrelated distractor was presented (the lexicality effect), β coefficient = 0.39, SE = 0.18, Wald $Z = 2.14$, $p = 0.032$. These analyses show that the effects in the errors are in the same direction as the RT effects, which indicates that there is no speed–accuracy trade-off in the data.

Table 2 | Mean response time (MRT, in milliseconds), standard deviations (SD), and percentage error (PE) per distractor type.

Distractor type	MRT	SD	PE
Unrelated	860	222	5.2
Semantically related	911	245	8.6
Translation	765	211	2.9
Control	791	215	3.6

DELTA-PLOT ANALYSES

The groups (created by median splits based on distractor effect size) differed in the magnitude of the semantic effect (16 vs. 81 ms), $F(1,14) = 26.43$, $p < 0.001$, the translation effect (48 vs. 115 ms), $F(1,14) = 13.87$, $p = 0.002$, and the lexicality effect (25 vs. 114 ms), $F(1,14) = 28.98$, $p < 0.001$. **Figure 2** gives the delta plots for the semantic, translation, and lexicality effects as a function of relative mean effect size (smaller, larger). The figure shows that the magnitude of the semantic, translation, and lexicality effects increases with increasing RT when the mean effect size is relatively large (presumed to reflect weak inhibition), whereas the magnitude of the effects levels off when the mean effect size is relatively small (presumed to reflect stronger inhibition). The fact that the delta plots for the smaller effect sizes leveled off (i.e., they are negative going) suggests that inhibition was present.

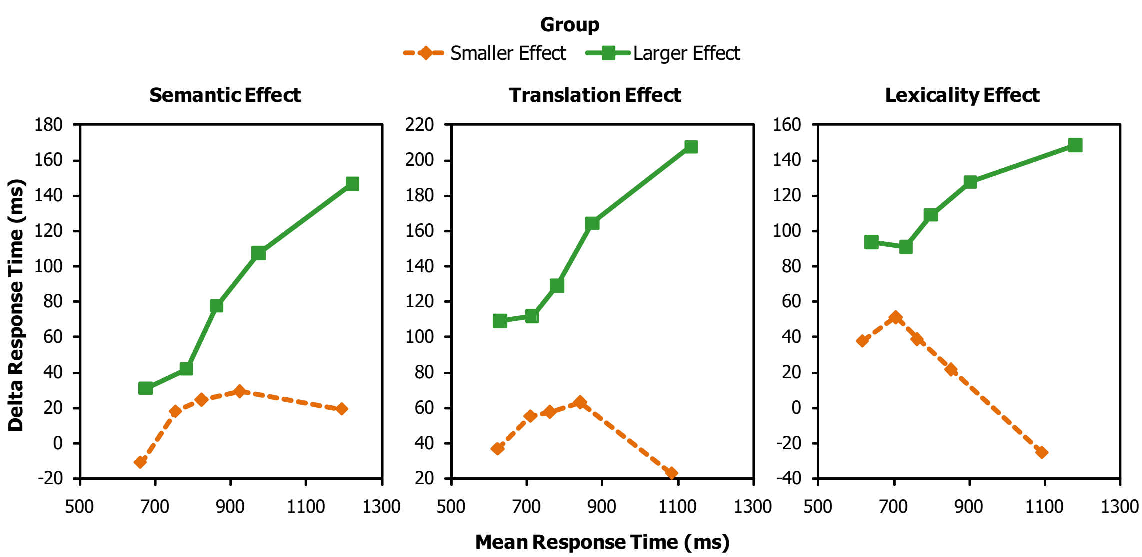


FIGURE 2 | Delta plots for the semantic, translation, and lexicality effects as a function of smaller and larger distractor effect size. The response time values on the horizontal axis for the delta plots are the means of the response times in the two conditions used to compute each delta value.

The difference in delta plots between the larger and smaller mean distractor effects was confirmed by statistical analyses of the slopes. For the semantic effect, the slopes differed between effect sizes significantly for the segment q2–3, $F(1,14) = 7.93$, $p = 0.014$, MSE = 0.056, but not for the other segments, q1–2, $F(1,14) = 1.55$, $p = 0.23$, MSE = 0.088; q3–4, $F(1,14) = 2.20$, $p = 0.16$, MSE = 0.087; q4–5, $F(1,14) = 0.68$, $p = 0.43$, MSE = 0.136. For the translation effect, the difference in slope as a function of effect size was marginally significant for the segment q4–5, $F(1,14) = 4.26$, $p = 0.058$, MSE = 0.085, but not for the other segments, q1–2, $F(1,14) = 1.18$, $p = 0.30$, MSE = 0.058; q2–3, $F(1,14) = 2.19$, $p = 0.16$, MSE = 0.061; q3–4, $F(1,14) = 2.73$, $p = 0.12$, MSE = 0.107. Finally, for the lexicality effect, the slopes differed as a function of effect size for the segment q2–3, $F(1,14) = 14.82$, $p = 0.002$, MSE = 0.053, and the segment q3–4, $F(1,14) = 5.32$, $p = 0.037$, MSE = 0.093, but not for the other segments, q1–2, $F(1,14) = 0.49$, $p = 0.50$, MSE = 0.109; q4–5, $F(1,14) = 2.59$, $p = 0.13$, MSE = 0.086.

Figure 3 gives the delta plots for the semantic, translation, and lexicality effects as a function of relative proficiency (higher, lower) based on the self-ratings of the level of skill in L2. The figure shows that the magnitude of the translation and lexicality effects increases with increasing RT for the low-proficiency individuals but not or much less for the high-proficiency individuals. Moreover, the magnitude of the semantic effect increases with increasing RT for the high-proficiency individuals, but much less for the low-proficiency ones.

The difference in delta plots between the proficiency groups was confirmed by statistical analyses of the slopes. For the semantic effect, the difference in slope between groups was marginally significant for the segment q2–3, $F(1,14) = 4.28$, $p = 0.058$, MSE = 0.068, significant for segment q3–4, $F(1,14) = 8.65$, $p = 0.011$, MSE = 0.062, but not for the other segments q1–2, $F(1,14) = 0.06$, $p = 0.80$, MSE = 0.097; q4–5, $F(1,14) = 0.02$, $p = 0.88$, MSE = 0.142. For the translation effect, the difference in slope between groups was significant for segment q2–3, $F(1,14) = 4.86$, $p = 0.045$, MSE = 0.052, and segment q3–4, $F(1,14) = 6.26$, $p = 0.025$, MSE = 0.088, but not for the other segments, q1–2, $F(1,14) = 1.42$, $p = 0.25$, MSE = 0.057; q4–5, $F(1,14) = 0.83$, $p = 0.38$, MSE = 0.105. Finally, for the lexicality effect, the slopes differed as a function of group for the segment q2–3, $F(1,14) = 12.20$, $p = 0.004$, MSE = 0.059, and the segment q3–4, $F(1,14) = 13.87$, $p = 0.002$, MSE = 0.065, but not for the other segments, q1–2, $F(1,14) = 0.37$, $p = 0.55$, MSE = 0.110; q4–5, $F(1,14) = 1.10$, $p = 0.31$, MSE = 0.094.

The inhibition effects appear ubiquitous. Yet, the observed patterns raise the question of the specificity of effects. If the differential delta-plot effects indicate effects of inhibition, then differential delta-plot effects should *not* be found when delta plots are computed for experimental factor effects that should not involve inhibition. In particular, differential delta plots should not be obtained for median splits based on factors that do not load on inhibition, such as mean RT or age (given that all participants were young). To assess the specificity of the delta-plot effects, we computed delta plots for groups based on median splits of mean RT and age.

The groups created by median splits based on naming speed differed in mean RT (766 vs. 897 ms), $F(1,14) = 17.57$, $p = 0.001$. However, the analyses of the slopes yielded no difference between the mean RT groups for the semantic effect, all $ps > 0.21$, the translation effect, all $ps > 0.17$, and the lexicality effect, all $ps > 0.17$. Moreover, the groups created by median splits based on age differed in mean age (21 vs. 25 years), $F(1,14) = 19.93$, $p = 0.001$. However, the analyses of the slopes yielded no difference between the age groups for the semantic effect, all $ps > 0.25$, the translation effect, all $ps > 0.18$, and the lexicality effect, all $ps > 0.14$. These analyses revealed that although the groups differed statistically in mean RTs and age, the delta plots for the semantic, lexicality, and translation effects did not differ. This suggests that the delta-plot effects are specific to experimental factors differing in inhibition, such as effect size and proficiency.

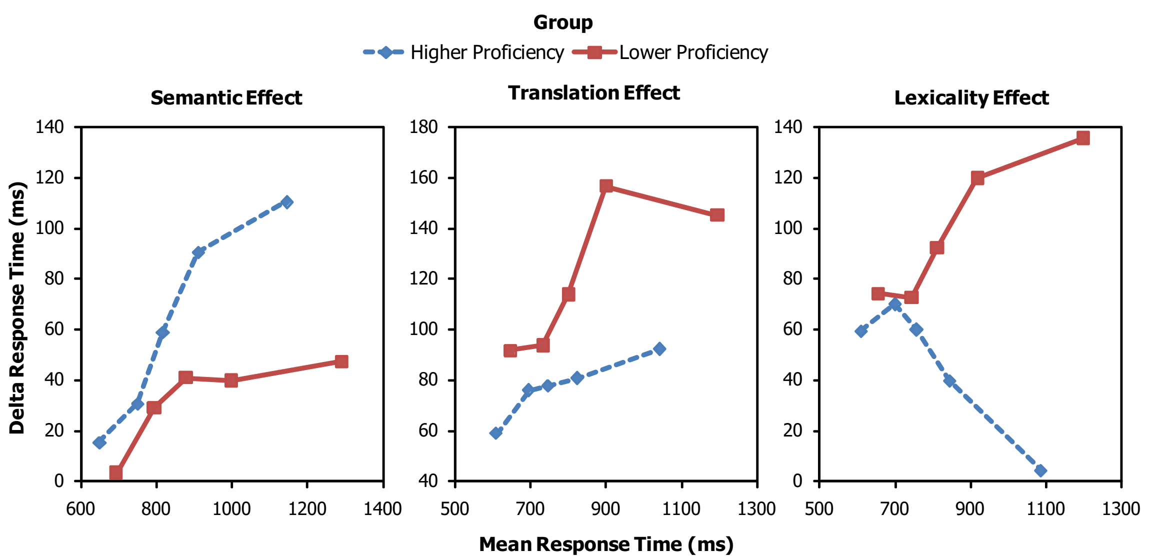


FIGURE 3 | Delta plots for the semantic, translation, and lexicality effects as a function of higher and lower proficiency. The response time values on the horizontal axis for the delta plots are the means of the response times in the two conditions used to compute each delta value.

DISCUSSION

Prior evidence from delta-plot analyses suggests that differential effects of attentional inhibition are largest in the tail of an RT distribution in non-linguistic and monolingual performance domains (e.g., Ridderinkhof, 2002; Ridderinkhof et al., 2004, 2005; Bub et al., 2006; Sharma et al., 2010). The reported experiment examined whether this also holds for bilingual performance by conducting delta-plot analyses of picture naming RTs. Dutch–English bilingual speakers named pictures in English while trying to ignore superimposed Dutch distractor words or neutral series of Xs. Picture name and distractor word were semantically related, unrelated, or translations. The mean RTs revealed semantic, translation, and lexicality effects. The delta pots revealed that the magnitude of the distractor effects flattened with increasing RT, more so when the mean distractor effect was smaller (presumed to reflect strong inhibition) as opposed to larger (presumed to reflect weaker inhibition). Moreover, the delta plots leveled off with increasing RT more for high- than low-proficiency individuals in the unrelated than the control and translation conditions, whereas the reverse held for the semantically related condition.

DYNAMICS OF INHIBITION

The present observation that the magnitude of the distractor effects leveled off with increasing RT, more so when the mean effect size was smaller than when it was larger, corresponds to what Ridderinkhof (2002) observed using the non-linguistic Simon task. In his study, the delta plots leveled off more in participants with relatively small Simon effects than in participants with relatively large Simon effects. According to Ridderinkhof et al. (2004, 2005), this suggests that inhibition builds up slowly during a trial in non-linguistic domains. The present evidence on the dynamics of inhibition also agrees with the evidence of Verhoef et al. (2009) that inhibition takes time to develop in bilingual language performance. In that study, it was observed that the magnitude of a right-lateralized N2 effect was modulated by the size of the preparation interval, but not by language or language switching. The effect of preparation interval suggests that inhibition builds up slowly during a trial, in line with the observations of Ridderinkhof and colleagues. However, the absence of an effect of language and switching on the N2 suggests that inhibition is an optional rather than compulsory mechanism.

The evidence for inhibition in the present experiment raises the question what exactly is inhibited in the bilingual picture–word interference task. Inhibition may concern responding to the distractor or in the irrelevant language. If language rather than distractor were inhibited, differences among distractor word conditions (i.e., semantically related, translation, unrelated) are not to be expected, unless the distractor words activate their language information to different degrees depending on the type of distractor. Thus, it is more likely that the inhibition concerns responding to the distractor word (cf. Bub et al., 2006; Sharma et al., 2010). This inhibition of distractor word processes may not be specific to bilingual performance (i.e., L2 naming), but presumably reflects the fact that distractor words have to be ignored in the picture–word interference paradigm, regardless of their language.

Another issue raised by the present findings is at what level of processing the interference effects occur. Whereas some researchers argue that the effects arise during picture name selection (Roelofs, 2003), others maintain that the effects arise during perceptual/conceptual encoding of the picture (Dell'Acqua et al., 2007; Van Maanen et al., 2009). In a recent study, we examined the time course of semantic, translation, and lexicality effects in overt picture naming by means of ERP recordings (Roelofs et al., submitted). The materials were the same as in the present study.

Predictions for the onset of the distractor effects in the ERPs were based on estimates of the timing of processing stages underlying word production provided by an influential meta-analysis by Indefrey and Levelt (2004). According to these estimations, based on an average naming latency of 600 ms, the stage of perceptual and conceptual encoding is completed around 150–200 ms after picture onset, after which lexical selection starts. As in the present study, the mean naming latencies in our bilingual EEG study were longer than 600 ms, namely around 840 ms in the control condition. Taking 840 ms as the mean naming latency, and scaling the estimates to this mean, gives us 280 ms as the end of the time window of perceptual and conceptual encoding and as the point in time at which the operation of word selection is initiated. If the effects emerge during perceptual and conceptual encoding, they should emerge in the EEG in a time window that extends at most to 280 ms post picture onset, whereas if the effects arise during lexical selection, they should emerge after 280 ms post picture onset. The ERP data revealed that the semantic, translation, and lexicality effects started to emerge 300 ms after picture onset, suggesting that they occurred during the selection of the picture name.

In the present experiment, differences in delta plots as a function of effect size and proficiency tended to occur in segments q2–3 and q3–4, but not in the other segments. The absence of a difference in the first segment q1–2 supports the assumption that inhibition takes some time to develop. Apparently, after some initial build up of inhibition over time, differences in delta-plot slopes as a function of effect size and proficiency started to arise, as reflected in the middle delta-plot segments q2–3 and q3–4. However, differences in slopes tended to be absent in the last segment q4–5. This suggests that participants did not maintain a high level of inhibition throughout a trial. One possibility is that maintaining inhibition requires effort, which participants were willing to invest for some period during a trial but not throughout a whole trial. Alternatively, it may have been impossible for the participants to keep up a high level of inhibition for a longer period (cf. De Jong et al., 1999). Either way, if inhibition diminishes toward the end of a trial, differences in delta-plot slopes as a function of effect size and proficiency will also disappear, as observed in the present experiment.

ROLE OF PROFICIENCY

It has been argued that bilingual individuals have enhanced inhibition capabilities compared to monolingual individuals (Bialystok et al., 2004; Costa et al., 2008), because of their more frequent use of inhibition. According to Green (1998), lexical competitors in the other language are more activated for high- than low-proficiency bilinguals and so require a greater degree of inhibition. As a result, the inhibition ability should improve with increased proficiency. Consequently, delta plots in bilingual picture–word interference should level off more for high- than low-proficiency individuals.

In the present study, the magnitude of the lexicality and translation effects increased with increasing RT for the low-proficiency individuals but not for the high-proficiency ones, in agreement with the hypothesis that inhibition enhances with proficiency. Moreover, the magnitude of the semantic effect increased with increasing RT for the high-proficiency individuals, but much less for the low-proficiency ones. The present data suggest that the high-proficiency participants applied more inhibition than the low-proficiency participants on unrelated trials relative to the other trial types. This leads to a stronger decrease of the distractor effect with increasing RT for the lexicality and translation effects, and also leads to the opposite influence for the semantic effect.

As **Table 2** shows, mean RTs were longer on unrelated than translation and control trials, but shorter on unrelated than semantically related trials. Inhibition of unrelated distractors will reduce RTs on unrelated trials and consequently will reduce the difference in RT between the unrelated and translation trials (i.e., the translation effect), reduce the difference in RT between the unrelated and control trials (i.e., the lexicality effect), but *increase* the difference in RT between the unrelated and semantically related trials (i.e., the semantic effect). If the unrelated distractor words are more strongly inhibited by the high- than the low-proficiency participants, the difference in RT between unrelated and translation trials (i.e., the translation effect) will decrease more with increasing RT for the high- than the low-proficiency participants, as empirically observed. Moreover, the difference in RT between unrelated and control trials (i.e., the lexicality effect) will decrease more with increasing RT for the high- than the low-proficiency participants, also as observed. However, if the unrelated distractor words are more strongly inhibited by the high- than the low-proficiency participants, the difference in RT between unrelated and semantically related trials (i.e., the semantic effect) will *increase* more with increasing RT for the high- than the low-proficiency participants, as empirically observed.

The observation that the delta plots level off differentially depending on the type of effect (i.e., semantic vs. translation and lexicality) suggests that individuals engage inhibition differently depending on the prevailing circumstances. This implies that participants monitored progress on task performance on each trial and allocated attentional inhibition depending on their assessment of the distractor type (cf. Kahneman, 1973; Roelofs, 2007, 2008).

Costa and Santesteban (2004) argued that the use of inhibition depends on language proficiency. In their view, while non-balanced bilinguals use inhibition to achieve language selectivity, balanced bilinguals do not use inhibition to accomplish this. As indicated earlier, this view of Costa and Santesteban (2004) has been challenged by Verhoef et al. (2009), who observed that whether inhibition is engaged may depend on the amount of preparation time in language switching in a single population of non-balanced Dutch–English bilinguals. The present results suggest that the use of inhibition by non-balanced bilinguals may not only depend on the preparation time, but also on the type of distractor. This corroborates the view that inhibition is an optional mechanism that is engaged depending on the prevailing circumstances.

CONCLUSION

The present study provides evidence that inhibition is a mechanism of attentional control in bilingual language performance. In a bilingual picture–word interference experiment, the magnitude of semantic, translation, and lexicality effects decreased with increasing RT, more so when the mean distractor effect is smaller (presumed to reflect strong inhibition) than when it is larger (presumed to reflect weaker inhibition). This suggests that the influence of inhibition is largest toward the RT distribution tail, corresponding to what is observed in non-linguistic domains. Moreover, the delta plots suggested that more inhibition was applied by high- than low-proficiency individuals on unrelated trials than on the other trial types. These results support the view that inhibition is a domain-general mechanism that may be optionally engaged depending on the prevailing circumstances.

ACKNOWLEDGMENT

This research was supported by a grant from the Netherlands Organisation for Scientific Research.

REFERENCES

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *Neuroscientist* 13, 214–228.

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci. (Regul. Ed.)* 8, 170–177.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bub, D. N., Masson, M. E. J., and Lalonde, C. E. (2006). Cognitive control in children: Stroop interference and suppression of word reading. *Psychol. Sci.* 17, 351–357.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S. T., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

Costa, A., Miozzo, M., and Caramazza, A. (1999). Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *J. Mem. Lang.* 41, 365–397.

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511.

De Jong, R., Berendsen, E., and Cools, R. (1999). Goal neglect and inhibitory limitations: dissociable causes of interference effects in conflict situations. *Acta Psychol. (Amst)* 101, 379–394.

De Jong, R., Liang, C.-C., and Lauber, E. (1994). Conditional and unconditional automaticity: a dual-process model of effects of spatial stimulus-response correspondence. *J. Exp. Psychol. Hum. Percept. Perform.* 20, 731–750.

Dell'Acqua, R., Job, R., Peressotti, F., and Pascali, A. (2007). The picture-word interference effect is not a Stroop effect. *Psychon. Bull. Rev.* 14, 717–722.

Duncan, J. (2010). *How Intelligence Happens*. Yale: Yale University Press.

Finkbeiner, M., Almeida, J., Janssen, N., and Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1075–1089.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008a). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *J. Neurosci.* 28, 9790–9796.

Forstmann, B. U., van den Wildenberg, W. P. M., and Ridderinkhof, K. R. (2008b). Neural mechanisms, temporal dynamics, and individual differences in interference control. *J. Cogn. Neurosci.* 20, 1854–1865.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Hermans, D., Bongaerts, T., De Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. Lang. Cogn.* 1, 213–229.

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

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol. (Amst.)* 128, 416–430.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York: Oxford University Press.

Proctor, R. W., Miles, J. D., and Baroni, G. (2011). Reaction time distribution analysis of spatial correspondence effects. *Psychon. Bull. Rev.* 18, 242–266.

Ridderinkhof, K. R. (2002). "Activation and suppression in conflict tasks: empirical clarification through distributional analyses," in *Attention and Performance XIX: Common Mechanisms in Perception and Action*, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 494–519.

Ridderinkhof, K. R., Scheres, A., Oosterlaan, J., and Sergeant, J. (2005). Delta plots in the study of individual differences: new tools reveal response inhibition deficits in AD/HD that are eliminated by methylphenidate treatment. *J. Abnorm. Psychol.* 114, 197–215.

Ridderinkhof, R. K., Van den Wildenberg, W. P. M., Wijnen, J., and Burle, B. (2004). "Response inhibition in conflict tasks is revealed in delta plots," in *Cognitive Neuroscience of Attention*, ed. M. I. Posner (New York: The Guilford Press), 369–377.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psych. Rev.* 110, 88–125.

Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *J. Mem. Lang.* 57, 232–251.

Roelofs, A. (2008). Dynamics of the attentional control of word retrieval: analyses of response time distributions. *J. Exp. Psychol. Gen.* 137, 303–323.

Roelofs, A., and Verhoef, K. (2006). Modeling the control of phonological encoding in bilingual speakers. *Biling. Lang. Cogn.* 9, 167–176.

Sharma, D., Booth, B., Brown, R., and Huguet, P. (2010). Exploring the temporal dynamics of social facilitation in the Stroop task. *Psychon. Bull. Rev.* 17, 52–58.

Van den Wildenberg, W. P. M., Wylie, S. A., Forstmann, B. U., Burle, B., Hasbroucq, T., and Ridderinkhof, R. K. (2011). To head or to heed? Beyond the surface of selective action inhibition: a review. *Front. Hum. Neurosci.* 4:222. doi: 10.3389/fnhum.2010.00222

Van Maanen, L., van Rijn, H., and Borst, J. P. (2009). Stroop and picture-word interference are two sides of the same coin. *Psychon. Bull. Rev.* 16, 987–999.

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

Wundt, W. (1904). *Principles of Physiological Psychology*. London: Swan Sonnenschein.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 February 2011; accepted: 21 July 2011; published online: 09 August 2011.
*Citation: Roelofs A, Piai V and Rodriguez GG (2011) Attentional inhibition in bilingual naming performance: evidence from delta-plot analyses. Front. Psychology **2**:184. doi: 10.3389/fpsyg.2011.00184*
This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

REVIEW ARTICLE
published: 15 September 2011
doi: 10.3389/fpsyg.2011.00234

Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology

Alexis Georges Hervais-Adelman[1]*, Barbara Moser-Mercer[2] and Narly Golestani[1]

[1] Functional Brain Mapping Laboratory, University of Geneva Medical School, University of Geneva, Genève, Switzerland
[2] Ecole de Traduction et d'Interprétation, University of Geneva, Genève, Switzerland

Edited by:
Ingrid Christoffels, University of Leiden, Netherlands

Reviewed by:
Jubin Abutalebi, Vita Salute San Raffaele University, Italy
Sylvain Moreno, The Rotman Research Institute, Canada

***Correspondence:**
Alexis Georges Hervais-Adelman, Functional Brain Mapping Laboratory, University of Geneva Medical School, rue Michel-Servet 1, CH-1211 Genève 4, Switzerland.
e-mail: alexis.adelman@unige.ch

In this review we will focus on delineating the neural substrates of the executive control of language in the bilingual brain, based on the existing neuroimaging, intracranial, transcranial magnetic stimulation, and neuropsychological evidence. We will also offer insights from ongoing brain-imaging studies into the development of expertise in multilingual language control. We will concentrate specifically on evidence regarding how the brain selects and controls languages for comprehension and production. This question has been addressed in a number of ways and using various tasks, including language switching during production or perception, translation, and interpretation. We will attempt to synthesize existing evidence in order to bring to light the neural substrates that are crucial to executive control of language.

Keywords: executive control, bilingualism, dorsolateral prefrontal cortex, parietal lobe, basal ganglia, anterior cingulate cortex, language switching, simultaneous interpretation

INTRODUCTION

There is a considerable behavioral literature demonstrating that multilingualism[1] has benefits in domains extending beyond language (Diamond, 2010). For example, these studies have shown greater cognitive flexibility and control (Bialystok and Senman, 2004; Bialystok and Depape, 2009; Adi-Japha et al., 2010), superior performance on non-verbal switching tasks (Garbin et al., 2010; Prior and MacWhinney, 2010), and advantages on tests of attentional control and flexibility (Costa et al., 2008, 2009; Hernandez et al., 2010) in multilingual compared to monolingual children and adults. Several studies have also used functional brain-imaging to examine the advantages conferred by bilingualism on executive control (Bialystok et al., 2005; Garbin et al., 2010).

On the basis of experiments on bilingual speech production, Abutalebi and Green (2007, 2008) propose a model of language control in the multilingual brain whereby a left hemisphere cortico-subcortical loop comprising the anterior cingulate and prefrontal cortices alongside the left caudate nucleus work to control linguistic functions subtended by a left fronto-parietal network.

In the present paper, we will first briefly describe neuropsychological, intracranial stimulation, and transcranial magnetic stimulation (TMS) studies that contribute to elucidating the neural bases of language control. We will then review functional neuroimaging studies using various paradigms and linguistic stimuli to examine aspects of language control in bilingualism.

[1]We use the terms bilingual and multilingual to refer to individuals who have attained a fluent level of speech in more than one language. We do not attempt to distinguish between individuals who acquired their second (or additional) languages early as opposed to late in life, although we acknowledge that this has a significant impact on language processing. In this paper we use the terms L1 to refer to an individual's first-acquired language and L2 to their second, irrespective of the fluency level of the two languages.

NEUROPSYCHOLOGICAL EVIDENCE

While neuroimaging studies provide correlational evidence for the engagement of brain regions during certain processes, more direct evidence for the involvement of a given structure is provided by neuropsychological reports, by intracranial stimulation studies, and by TMS studies which can directly show that a region is critically involved in a given process. Such lines of evidence are less frequently encountered than functional imaging. Here, we will briefly examine two reports of different speech pathologies, as well as the limited evidence from direct electrical brain stimulation studies. We will also describe a few relevant TMS case studies. Much of the evidence presented here focuses on language switching, i.e., the process of changing the output language. Healthy bilinguals normally select and switch languages as a function of the linguistic knowledge of the interlocutor. This normal process may, however, be disrupted after brain lesions.

Aphasia is a disorder of speech in which comprehension or production of speech is impaired, an extensive discussion of which is beyond the scope of this review. However, aphasia in bilingual patients is of interest in so far as some cases show a loss of appropriate control of language in one or both languages, thereby providing evidence for the involvement of particular brain areas for language control.

A psycholinguistic model of bilingual speech production worth considering in this context is Paradis' activation threshold hypothesis (Paradis, 1993, 2001). It suggests that at any given moment, a stored item in a multilingual lexicon requires a certain amount of activation (its threshold) in order to be accessed. When the threshold is reached, all other alternatives are inhibited (their thresholds are raised). However, thresholds are never raised so high as to totally inhibit the language. Every time a particular item is activated its threshold decreases, but if it is not used for a while, its threshold increases. It has been suggested (Paradis, 1996) that lesions can alter the threshold of languages, which could therefore explain asymmetric impairments in languages in bilingual aphasia, as well as unequal recovery.

Pathological mixing, which refers to the intermingling of languages within a single utterance, has been reported after left temporo-parietal lesions (Fabbro, 1999), suggesting a crucial role for this region in maintaining the appropriate "language set" for output. A related, but different pathology is pathological switching, in which a patient alternates the language of utterance between self-contained speech segments. Fabbro et al. (2000) report a case of pathological language switching following a lesion of the left (and partly of the right) anterior cingulate gyrus, and of the white matter underlying the left inferior, middle, and superior frontal gyri. The patient displayed no aphasic symptoms in either language but was found to switch between his L1 and L2 despite instructions to speak only one language, and despite displaying awareness of the switches.

Abutalebi et al. (2000) report a case of pathological language mixing after a lesion incorporating the head of the left caudate nucleus. They note that the patient always produced output in which noun phrases were complete, and morphological markers were always used appropriately, suggesting that the impairment involved a stage of lexical selection subsequent to specification of syntactic and semantic information.

Marien et al. (2005) report a rare case of bilingual subcortical aphasia in a child. Following a left thalamic hemorrhage, the patient displayed global aphasia in L1 and L2. A few days later the patient displayed fluent aphasia equally affecting L1 and L2, prominently featuring spontaneous pathological language mixing and switching. The patient also displayed significant translation difficulties. Investigation with SPECT[2] showed hypoperfusion in left fronto-parietal and temporal regions as well as the left caudate nucleus. Six months later pathological mixing and switching had remitted but translation difficulties and fluent L1 and L2 aphasia had not. Follow-up SPECT imaging showed relative increases in perfusion in the left frontal regions and left caudate nucleus, but not the left temporal or parietal regions. The pattern of recovery in this case provides compelling evidence for the involvement of a left fronto-subcortical network in language selection.

Abutalebi et al. (2009) studied the recovery of a patient afflicted by bilingual aphasia after a left basal ganglia hemorrhage implicating the globus pallidus and putamen. Initially the patient showed global aphasia in L1 and L2, which changed to fluent aphasia with anomia equally affecting both languages after a few days. The patient was then treated with intensive speech therapy in L2 only, which significantly improved the aphasic symptoms in L2, but not L1. A dynamic causal modeling analysis of functional magnetic resonance imaging (fMRI) data obtained during a bilingual picture naming task showed that improvement in naming performance after treatment was associated with increased functional coupling between a fronto-subcortical network (the "control network") and a fronto-temporal one (the "language network").

Aglioti and Fabbro (1993), followed-up by Aglioti et al. (1996) report a case of subcortical bilingual aphasia in which a lesion of the left basal ganglia asymmetrically more severely impaired the patient's most used language. The impairment included increased difficulty translating into this language. They ascribed this asymmetrical outcome to the role of the basal ganglia in controlling automatized motor and cognitive tasks (for review, see Takakusaki et al., 2004) and in managing behavioral patterns (cf. Graybiel, 1997), arguing that the most used language is more automated than a less used one.

These case studies heavily implicate the basal ganglia in the control of language output in multilingual individuals. However, one report contradicts this view. Fabbro et al. (2000) report a case of pathological language switching following a lesion of the left (and partly of the right) anterior cingulate gyrus, and of the white matter underlying the left inferior, middle, and superior frontal gyri. The patient displayed no aphasic symptoms in either language but was found to switch between his L1 and L2 despite instructions to speak only one language, and despite displaying awareness of the switches. Although Fabbro et al. (2000) argue that language switching is controlled by mechanisms generally involved in task switching (that is, a fronto-parietal network) the published MRI images of the lesion also show damage to the left striatum (see Marien et al., 2005), which could be the cause of the patient's pathological switching. This reinterpretation is consistent with other reports of pathological language mixing and switching in polyglot aphasia following subcortical damage.

INTRACRANIAL ELECTRICAL STIMULATION AND TRANSCRANIAL MAGNETIC STIMULATION EVIDENCE

Less direct evidence for the involvement of the left dorsolateral prefrontal cortex (DLPFC) in language switching comes from a TMS study by Holtzheimer et al. (2005). They reported involuntary language switching in two patients who were treated for drug-resistant major depressive disorder using repetitive TMS to temporarily inhibit various cortical regions. In the first case, an English–German bilingual patient, whose primary spoken and written language was English, reported "thinking in German" after a session of rTMS over the left DLPFC. A second patient, an English–Spanish bilingual, similarly reported the "thinking in Spanish" and the impulse to speak to the tester in Spanish after rTMS of the left DLPFC. Such evidence is not entirely conclusive, since the mechanism of action is unknown, but it further bolsters the evidence that the left DLPFC can play a role in language switching.

Nardone et al. (2011) have recently reported that excitatory TMS of the left DLPFC transiently alleviated pathological language switching in a bilingual patient who had suffered a left frontal stroke.

The most direct evidence for brain regions implicated in language control in bilinguals come from intracranial stimulation studies, although such studies are very rare. Moritz-Gasser and Duffau (2009b) used direct electrical stimulation to map a

[2] Single-photon emission computed tomography, a technique allowing imaging of brain metabolic activity.

language switching network in a bilingual patient. They demonstrated that stimulation of multiple sites could induce language switching, namely stimulation of the left posterior superior temporal sulcus and subcortical stimulation of the left superior longitudinal fasciculus (a white matter tract which connects the left inferior frontal and posterior superior temporal cortices). In an earlier intraoperative study, Kho et al. (2007) induced an involuntary shift from French (L1) to Chinese (L2) during a counting task by stimulating a site in the left inferior frontal gyrus. In another patient they reported involuntary language switching during a Wada test in which the left hemisphere was anesthetized.

Taken together, these reports point to the involvement of a left-lateralized fronto-temporal network in regulating language switching. We suggest that this can be reconciled with the apparently contradictory neuropsychological evidence in the following way. These studies, which show that stimulation or inhibition of cortical areas can lead to language switching, do not necessarily prove that these regions are involved in language selection processes. There is evidence that different languages may be represented in different portions of cortex in multilingual brains (Fabbro, 2001; Sebastian-Galles et al., 2006; Leonard et al., 2010), although this may be a function of proficiency or age of acquisition of L2 (Dehaene et al., 1997; Kim et al., 1997; Golestani et al., 2006). If this holds for the participants of these investigations then it may be the case that these investigations have differentially inhibited or excited the representation of a given language over another, and it is the consequent facilitation or impairment of access that leads to the language switching behavior, without selection mechanisms necessarily being involved. Under such a schema, the subcortical regions implicated by the neuropsychological evidence are likely to be involved in the management of cortical representation for appropriate behavioral output. Such an architecture is in line with much existing data on executive control in other domains. However, the cases reported by Kho et al. (2007) do suggest that a left-lateralized cortical network is part of the switching mechanism.

By examining the functional neuroimaging literature alongside these case reports, we can further delineate the role of these brain areas, *in vivo*, in healthy volunteers.

NEUROIMAGING EVIDENCE

We will focus principally on fMRI studies as they are the most informative with respect to localization of language control processes. We will also review studies that have employed alternative imaging techniques, such as positron emission tomography (PET) and optical imaging (near infrared spectroscopy, NIRS). Numerous electroencephalography (EEG) studies have been carried out to explore these questions. However, these studies mainly focus on the temporal dynamics of language control rather than its localization, and we will therefore address them briefly in a separate section.

LANGUAGE SWITCHING

Language switching tasks can provide direct insight into the substrates of controlling language. Behavioral evidence (Meuter and Allport, 1999) shows that switching between languages is associated with a cost, as manifested by slowed reaction times. The neural manifestations of this cost have been investigated using a variety of tasks in which participants are required either to comprehend or to produce stimuli in multiple languages.

Receptive tasks have included listening to a series of words in either of two languages [Price et al., 1999 (PET); Rodriguez-Fornells et al., 2002 (fMRI)], and listening to sentences with a language switch midway through [Abutalebi et al., 2007 (fMRI)].

Crinion et al. (2006) used a task involving covert reading of words in alternating languages [2006 (PET)]. Explicit, or overt production tasks that have been employed include naming pictures in alternating languages [Hernandez et al., 2000 (fMRI); Hernandez et al., 2001 (fMRI); Khateb et al., 2007 (EEG); Abutalebi et al., 2008 (fMRI); Costa et al., 2009 (fMRI)], digit naming in alternating languages [Wang et al., 2009 (fMRI)], language switching during verbal fluency tasks [Hirshorn and Thompson-Schill, 2006 (fMRI)], and language switching during alternate translation from L1 → L2 and L2 → L1 [Quaresima et al., 2002 (NIRS)].

Neuroimaging evidence for the neural substrates of language switching has implicated a network of predominantly left-hemisphere lateralized cortical regions, including the superior temporal sulcus (Rodriguez-Fornells et al., 2002; Abutalebi et al., 2007, 2008), the superior and inferior parietal lobule (Price et al., 1999; Rodriguez-Fornells et al., 2002; Hirshorn and Thompson-Schill, 2006; Khateb et al., 2007; Costa et al., 2009; Wang et al., 2009), the supplementary motor area (SMA; Wang et al., 2007; Abutalebi et al., 2008), the DLPFC (Hernandez et al., 2000, 2001; Hirshorn and Thompson-Schill, 2006; Khateb et al., 2007; Abutalebi et al., 2008; Wang et al., 2009), the inferior frontal gyrus (Price et al., 1999; Hernandez et al., 2001; Quaresima et al., 2002; Rodriguez-Fornells et al., 2002; Hirshorn and Thompson-Schill, 2006; Abutalebi et al., 2007, 2008), the precentral gyrus (Khateb et al., 2007; Wang et al., 2009), and the right anterior cingulate cortex (Abutalebi et al., 2008). Other right hemisphere activations are reported in the DLPFC, the precentral gyrus, and the SMA by Hernandez (2009) for switching versus not switching during picture naming. Despite the heterogeneity of paradigms used, a consensus does seem to emerge, implicating the left inferior frontal gyrus, left DLPFC, and the left parietal lobule during language switching, consistent with the evidence from TMS and direct stimulation studies presented above. The SMA and precentral gyrus may additionally be engaged in tasks that involve productive switches.

The above described fronto-parietal network overlaps considerably with that ascribed to general executive control which is implicated in diverse processes such as inhibition of prepotent responses, initiation of behavior, planning of action, judgment and decision making, and feedback management (e.g., Collette et al., 2005, 2006; Schumacher et al., 2007; Nagel et al., 2008). Considerable attention has been devoted to the differences between language switching and more general task switching (for discussion see Moritz-Gasser and Duffau, 2009a), and the extent of such differences remains a matter of debate.

TRANSLATION TASKS

Translation requires rapid access to representations of lexical items in two languages. It therefore demands a different type of language control compared to that required during language switching:

selection is still essential, but simply favoring one language over another will not enable faithful translation, beyond the case of isolated words.

Two of the studies described above also included translation tasks (silently mouthing translations of visually presented words: Price et al., 1999; or overtly producing translations of them: Quaresima et al., 2002) in the context of language switching paradigms. Further studies have focused more explicitly on the process of translation. Klein et al. (1995) recorded brain activity using PET while bilingual participants overtly translated single auditorily presented words. Lehtonen et al. (2005) conducted an fMRI investigation in which they asked bilingual individuals to silently translate visually presented sentences. This latter study is of particular interest as it is the only one in which participants are required to tap supra-lexical levels of the speech system in order to successfully carry out the translation task.

Price et al. (1999) showed involvement of the anterior cingulate cortex, the putamen and head of the caudate nucleus, the SMA, the left anterior insula, and the cerebellum bilaterally during silent mouthing of translations. Quaresima et al. (2002) examined only the anterior portion of the left hemisphere during overt translation, and found activation of this region during task performance. Klein et al. (1995) demonstrated engagement of left-lateralized inferior frontal, dorsolateral prefrontal, and inferior temporal cortices, as well as (specifically for translation from L2 → L1) activation of the left putamen. Lehtonen et al. (2005) reported significant activation of the left inferior frontal gyrus and putamen for translation from L2 → L1. We propose that these activations arise from two processes: semantic retrieval in the left inferior frontal gyrus, and control of output in the basal ganglia.

INTERPRETATION TASKS

Simultaneous interpretation places even heavier demands upon the executive control of language than does translation of isolated words or sentences. It requires not just the ongoing retrieval of lexical, terminological, and phraseological units in the appropriate language, but also the maintenance of information in verbal working memory and the continuous monitoring of input and output streams, while constantly executing language and modality switches (Moser, 1978; Moser-Mercer et al., 2000; Christoffels et al., 2006)

There are substantial difficulties in examining overt interpretation of sentences using most imaging techniques as they are highly susceptible to the artifacts arising from speech-related movements. Thus, very few studies have attempted to investigate simultaneous interpretation. Rinne et al. (2000) carried out a PET investigation of eight professional simultaneous interpreters, using overt production. They found that the left premotor and ventrolateral prefrontal cortices were engaged during interpretation both from L2 to L1 and from L1 to L2. In addition, interpreting into L2 engaged the left inferior temporal cortex and the right cerebellum.

In an ongoing fMRI study of the neural substrates of simultaneous interpretation (Hervais-Adelman and colleagues, in preparation), 34 multilingual participants were asked to listen to sentences in a highly proficient second language, and to either shadow (simultaneously repeat) or simultaneously interpret sentences into their L1. Shadowing speech calls for simultaneous speech production and perception, as well as for the simultaneous processing of two streams of speech (one being the sentence being heard, and the other being the feedback from the participants' own output) in a single language, whereas simultaneous interpretation calls for the simultaneous processing of two languages, with the input and output streams being different languages. Comparison of activations arising from these two reveals the substrates underlying the simultaneous processing of two languages during simultaneous interpretation. Preliminary results provide evidence for the engagement of the left premotor and ventrolateral prefrontal cortices, alongside the pre-SMA and caudate nucleus for interpretation into L1 (**Figure 1**). The pattern of the preliminary results is consistent with much of the evidence presented above for the role of these regions in language control.

We have also recently found evidence for brain structural plasticity in individuals training to become simultaneous language interpreters as they develop expertise in this skill. We found that in interpretation students, but not in matched controls, there is an increase in gray matter volume over the course of a 15-month training program in brain regions known to be involved not only in semantic processing but also in aspects of executive function and error monitoring (**Figure 2**; Golestani and colleagues, in preparation). These preliminary results constitute direct, longitudinal

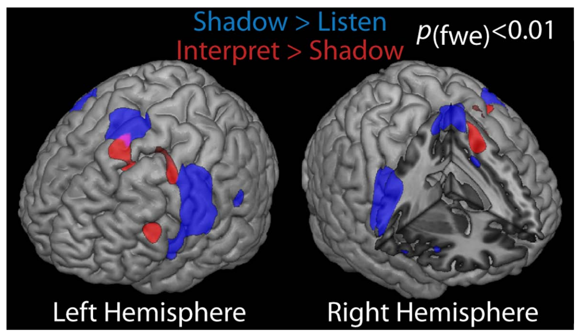


FIGURE 1 | Significant differences in activation levels in 34 non-experts, rendered on canonical single-subject brain. Contrasts shown are speech shadowing in L2 versus listening to L2 (blue) and simultaneous interpretation into L1 versus shadowing (red), at a family wise error corrected significance level of $p < 0.01$.

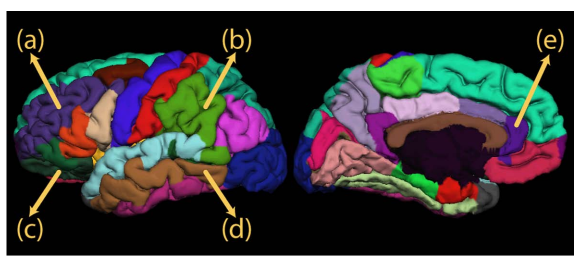


FIGURE 2 | Regions in which we found longitudinal evidence for brain structural plasticity in simultaneous language interpreters: (A) left middle frontal gyrus, (B) left supramarginal gyrus, (C) left pars orbitalis, (D) left middle temporal gyrus, (E) rostral anterior cingulate.

evidence for experience-dependent plasticity. These results, coupled with other functional imaging results, lend further credibility to the hypothesis of a left-lateralized fronto-parieto-subcortical mechanism for controlling language output and comprehension in multilingual individuals.

ELECTROENCEPHALOGRAPHY

The literature on event-related potentials (ERPs) in the study of bilingualism has been thoroughly reviewed by Moreno et al. (2008). ERPs reflect underlying neural activity with a high degree of temporal resolution but in themselves do not provide information about the location of that activity in the brain. Although it is possible to localize the sources of ERPs with an adequate degree of spatial resolution, the articles described in the following section describe only analyses of the temporal dynamics of the neural responses, with varying degrees of topographic accuracy. Even though they do not provide information relating to localization of relevant functions, the information they provide about the time-course of processing is nevertheless illuminating.

In speech production, there is plentiful evidence (see, for example, reviews by Costa, 2005; and Kroll et al., 2008) that languages are simultaneously activated and the inappropriate one suppressed, as a function of task. By providing information at a higher temporal resolution than other imaging modalities, ERP studies can directly address questions such as the psycholinguistic stages of representation and selection at which language interference occurs. For example, Hoshino and Thierry (2011) used EEG in an interference paradigm to determine the timing of language selection in a production task, and showed parallel activation of languages even beyond lexical selection. In comprehension, Van Heuven and Dijkstra (2010) have reviewed the EEG and MRI evidence for various psycholinguistic models of bilingual word recognition, and have found that the evidence favors their bilingual interactive activation+ (BIA+) model, which posits an integrated bilingual lexicon that is accessed in a language non-selective manner (for details of the model see Dijkstra and Van Heuven, 2002).

We will here look at two paradigms that have been used to explore ERPs of language control. These are go/no-go tasks and language switching tasks. We begin by discussing go/no-go tasks.

GO/NO-GO TASKS

In a go/no-go task, participants are required to respond only if certain conditions are met ("go" trials) or otherwise to make no response ("no-go" trials). The magnitude of an ERP component known as the N200 during no-go trials is thought to reflect the control processes relating to suppression of responses. The N200 (or N2) is a negativity observed approximately 200 ms after stimulus onset. The exact role of the N200 is debated; Nieuwenhuis et al. (2003) argue that its presence reflects response inhibition, while Donkers and van Boxtel (2004) argue that it reflects conflict-monitoring. Enhancement of N200 components has been interpreted as reflecting interference effects in bilingual tasks, and has been fairly widely observed (see Rodriguez-Fornells et al., 2006 for review). However, more recent evidence (Huster et al., 2010) suggests that the N200 may in fact reflect response selection, and that a later component, the P300, may reflect inhibitory cognitive components. Nevertheless, the N200 is closely associated with some aspect of response–suppression, in linguistic and non-linguistic tasks. Nieuwenhuis et al. (2003) localized the source of the N200 as the anterior cingulate cortex, and Huster et al. (2010) attributed its source to the left anterior middle cingulate cortex. These localizations are consistent with the neuroimaging evidence presented above.

Moreno et al. (2008) also describe a later ERP component that is systematically greater in amplitude in bilingual than monolingual participants during no-go trials, this being a mid-frontal negativity observed between 400 and 800 ms post-stimulus onset (also reviewed in Rodriguez-Fornells et al., 2006). They suggest that this component reflects enhanced cognitive control mechanisms related to the day-to-day demands of bilingualism.

LANGUAGE SWITCHING TASKS

Language switching is a task that directly calls upon language selection and control mechanisms, and has been extensively used in the study of bilingual control. However, the paradigms and results are rather heterogeneous across studies, and as such, the typical ERP components of language switching during speech production have not been well characterized. Nevertheless, over the studies they review, Moreno et al. (2008) conclude that the data indicate that language switching in production requires active inhibition of a non-target language, and that the ERPs related to language switching and to withholding responses during non-linguistic go/no-go tasks are substantially similar.

For switching during receptive tasks, the data are likewise inconsistent and seem to vary depending on the paradigm. Paradigms requiring participants to make semantic judgments appear to elicit enhanced N400 components for switch trials (e.g., Alvarez et al., 2003; Proverbio et al., 2004). Although there is an ongoing controversy about the exact functional interpretation of the N400, it is generally accepted that the amplitude of the N400 component is sensitive to semantic aspects of word processing, particularly to the cloze probability of a word as it is greater in the case of unexpected words (Kutas et al., 2006; Steinhauer and Connolly, 2008; Friederici and Wartenburger, 2010). While Alvarez et al. (2003) found that N400 was specifically enhanced for L1 to L2 switches, Chauncey et al. (2008) found the reverse. They used masked-priming to examine the ERP correlates of language switching without an overt language switch (the primes were largely invisible) and found enhancements of N250 and N400 components for switch trials. The N400 component was particularly enhanced for L2 to L1 switches and the N250 component was particularly enhanced for L1 to L2 switches.

Code switches are a particular form of language switching, whereby multilingual speakers electively employ words from alternative languages within utterances, while respecting the syntactic structure of the carrier language. It may be expected that listening to such switches might elicit similar ERPs to those described above. However, Moreno et al. (2002) found that code switches within sentences did not elicit enhanced N400 effects while lexical switches did. Instead, the code switch trials produced an enhanced posterior late positivity component (LPC), which is generally observed in response to unexpected or improbable task-relevant events (see, e.g., reviews by Donchin and Coles, 1988; Picton, 1992; Polich and Kok, 1995; Polich, 2007).

A recent study by Kuipers and Thierry (2010) sought to examine the time-course of neural events related to the detection of language changes using an auditory oddball paradigm. They compared ERPs elicited by rare language switch events with those elicited by frequent no-switch trials. They found that bilingual participants showed a response to language switches as early as 200 ms, followed by an N400, while monolingual participants showed only an enhanced N400 in response to switches, suggesting a fundamental difference in the early processing of words in bilinguals. There was also a group difference in the P600 component, which was enhanced for switch trials in bilinguals but not monolinguals. The P600 is associated with stimulus re-evaluation (Osterhout and Holcomb, 1992; Hahne and Friederici, 1999), implying that the bilinguals and not the monolinguals engaged in a process of reinterpreting the stimuli after a switch. The data suggest that bilinguals have a mechanism for rapidly detecting and adapting to language switches.

Overall, the existing work using ERPs to investigate the neural substrates of language control reveals several similarities between bilingual language control and control of other executive functions. Although it is difficult to draw conclusions about the localization of the functions tapped by the variety of tasks and paradigms that have been employed, the findings are complementary to those revealed using methods that offer higher spatial resolution. Additionally, ERP findings contribute to a better understanding of the stages of processing involved in bilingual language control.

CONCLUSION

We have described a number of studies from functional neuroimaging, direct brain stimulation, TMS, and neuropsychology that outline the neural bases of the executive control of language. Beyond the domain of multilingual language control, a fronto-basal-ganglia network has been implicated in the inhibitory control of action and cognition (Aron et al., 2007), and this appears to converge with part of the putative bilingual language control network outlined here. In the context of language switching tasks, the evidence points mainly to a cortical network incorporating the parietal lobe, the posterior superior temporal sulcus and the left inferior frontal gyrus. Tasks involving the conversion of content from one language to another (i.e., translation and interpretation) mainly engage a left-lateralized cortico-subcortical circuit, including the basal ganglia, inferior frontal gyrus, and DLPFC. There is strong anatomical support for functional links between these regions.

We propose that the evidence suggests the presence of two distinct networks contributing to the executive control of language. Although perturbing either may have superficially similar behavioral consequences, they are likely to have differing roles. It seems likely that a fronto-basal-ganglia loop is implicated in the inhibition of inappropriate languages during production. The basal ganglia also play an apparently crucial role in enabling access to translation equivalents, which may reflect inhibitory processes that allow the selection of a term in one language rather than another. Alongside this network, there appears to be another, cortical, fronto-parietal network that sustains more general switching mechanisms. This system, like the fronto-basal-ganglia system delineated above, has a role in other executive functions. These two systems, working in concert with language-specific brain areas, likely manage both inhibitory control as well as language selection, both of which are necessary for the effective management of language in bilingual brains.

The critical components underlying the executive control of language in the multilingual brain seem well delineated, but the exact functional roles of these components and their interactions remain to be fully described. Ongoing work on the acquisition of expertise in interpretation, which is a highly demanding linguistic task involving rapid language switching and handling multiple simultaneous linguistic streams, will shed further light on the executive control of language in the multilingual brain.

REFERENCES

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

Abutalebi, J., Brambati, S. M., Annoni, J. M., Moro, A., Cappa, S. F., and Perani, D. (2007). The neural cost of the auditory perception of language switches: an event-related functional magnetic resonance imaging study in bilinguals. *J. Neurosci.* 27, 13762–13769.

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

Abutalebi, J., and Green, D. (2008). Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang. Cogn. Process.* 23, 557–582.

Abutalebi, J., Miozzo, A., and Cappa, S. F. (2000). Do subcortical structures control "language selection" in polyglots? Evidence from pathological language mixing. *Neurocase* 6, 51–56.

Abutalebi, J., Rosa, P. A., Tettamanti, M., Green, D. W., and Cappa, S. F. (2009). Bilingual aphasia and language control: a follow-up fMRI and intrinsic connectivity study. *Brain Lang.* 109, 141–156.

Adi-Japha, E., Berberich-Artzi, J., and Libnawi, A. (2010). Cognitive flexibility in drawings of bilingual children. *Child Dev.* 81, 1356–1366.

Aglioti, S., Beltramello, A., Girardi, F., and Fabbro, F. (1996). Neurolinguistic and follow-up study of an unusual pattern of recovery from bilingual subcortical aphasia. *Brain* 119, 1551.

Aglioti, S., and Fabbro, F. (1993). Paradoxical selective recovery in a bilingual aphasic following subcortical lesions. *Neuroreport* 4, 1359.

Alvarez, R. P., Holcomb, P. J., and Grainger, J. (2003). Accessing word meaning in two languages: an event-related brain potential study of beginning bilinguals. *Brain Lang.* 87, 290–304.

Aron, A. R., Durston, S., Eagle, D. M., Logan, G. D., Stinear, C. M., and Stuphorn, V. (2007). Converging evidence for a fronto-basal-ganglia network for inhibitory control of action and cognition. *J. Neurosci.* 27, 11860–11864.

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

Bialystok, E., and Depape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 565–574.

Bialystok, E., and Senman, L. (2004). Executive processes in appearance-reality tasks: the role of inhibition of attention and symbolic representation. *Child Dev.* 75, 562–579.

Chauncey, K., Grainger, J., and Holcomb, P. J. (2008). Code-switching effects in bilingual word recognition: a masked priming study with event-related potentials. *Brain Lang.* 105, 161–174.

Christoffels, I. K., De Groot, A., and Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. *J. Mem. Lang.* 54, 324–345.

Collette, F., Hogge, M., Salmon, E., and Van Der Linden, M. (2006). Exploration of the neural substrates of executive functioning by functional neuroimaging. *Neuroscience* 139, 209–221.

Collette, F., Van Der Linden, M., Laureys, S., Delfiore, G., Degueldre, C., Luxen, A., and Salmon, E. (2005). Exploring the unity and diversity of the neural substrates of executive functioning. *Hum. Brain Mapp.* 25, 409–423.

Costa, A. (2005). "Lexical access in bilingual production," in *Handbook of Bilingualism: Psycholinguistic Approaches,* eds J. Kroll and A. M. B. De Groot (Oxford: Oxford University Press), 308–325.

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

Costa, A., Hernandez, M., and Sebastian-Galls, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green, D. W., and Price, C. J. (2006). Language control in the bilingual brain. *Science* 312, 1537–1540.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., Van De Moortele, P. F., Lehèricy, S., and Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8, 3809.

Diamond, J. (2010). The benefits of multilingualism. *Science* 330, 332–333.

Dijkstra, T., and Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. Lang. Cogn.* 5, 175–197.

Donchin, E., and Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357–374.

Donkers, F. C. L., and van Boxtel, G. J. M. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain Cogn.* 56, 165–176.

Fabbro, F. (2001). The bilingual brain: cerebral representation of languages. *Brain Lang.* 79, 211–222.

Fabbro, F. (ed.). (1999). *The Neurolinguistics of Bilingualism: An Introduction.* Hove: Psychology Press.

Fabbro, F., Skrap, M., and Aglioti, S. (2000). Pathological switching between languages after frontal lesions in a bilingual patient. *J. Neurol. Neurosurg. Psychiatr.* 68, 650.

Friederici, A. D., and Wartenburger, I. (2010). Language and brain. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 150–159.

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., Hernandez, M., Costa, A., and Avila, C. (2010). Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. *Neuroimage* 53, 1272–1278.

Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., and Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia* 44, 1029–1040.

Graybiel, A. M. (1997). The basal ganglia and cognitive pattern generators. *Schizophr. Bull.* 23, 459–469.

Hahne, A., and Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. *J. Cogn. Neurosci.* 11, 194–205.

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

Hernandez, A. E. (2009). Language switching in the bilingual brain: what's next? *Brain Lang.* 109, 133–140.

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

Hernandez, M., Costa, A., Fuentes, L. J., Vivas, A. B., and Sebastian-Galles, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Biling. Lang. Cogn.* 13, 315–325.

Hirshorn, E. A., and Thompson-Schill, S. L. (2006). Role of the left inferior frontal gyrus in covert word retrieval: neural correlates of switching during verbal fluency. *Neuropsychologia* 44, 2547–2557.

Holtzheimer, P., Fawaz, W., Wilson, C., and Avery, D. (2005). Repetitive transcranial magnetic stimulation may induce language switching in bilingual patients. *Brain Lang.* 94, 274–277.

Hoshino, N., and Thierry, G. (2011). Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain Res.* 1371, 100–109.

Huster, R. J., Westerhausen, R., Pantev, C., and Konrad, C. (2010). The role of the cingulate cortex as neural generator of the N200 and P300 in a tactile response inhibition task. *Hum. Brain Mapp.* 31, 1260–1271.

Khateb, A., Abutalebi, J., Michel, C. M., Pegna, A. J., Lee-Jahnke, H., and Annoni, J. M. (2007). Language selection in bilinguals: a spatio-temporal analysis of electric brain activity. *Int. J. Psychophysiol.* 65, 201–213.

Kho, K. H., Duffau, H., Gatignol, P., Leijten, F. S. S., Ramsey, N. F., Van Rijen, P. C., and Rutten, G. J. M. (2007). Involuntary language switching in two bilingual patients during the Wada test and intraoperative electrocortical stimulation. *Brain Lang.* 101, 31–37.

Kim, K. H. S., Relkin, N. R., Lee, K. M., and Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature* 388, 171–174.

Klein, D., Milner, B., Zatorre, R. J., Meyer, E., and Evans, A. C. (1995). The neural substrates underlying word generation – a bilingual functional-imaging study. *Proc. Natl. Acad. Sci. U.S.A.* 92, 2899–2903.

Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol. (Amst.)* 128, 416–430.

Kuipers, J. R., and Thierry, G. (2010). Event-related brain potentials reveal the time-course of language change detection in early bilinguals. *Neuroimage* 50, 1633–1638.

Kutas, M., Van Petten, C., and Kluender, R. (2006). "Psycholinguistics electrified II: 1994–2005," in *Handbook of Psycholinguistics,* 2nd Edn, eds M. A. Gernsbacher and M. Traxler (New York, NY: Elsevier Press), 659–724.

Lehtonen, M. H., Laine, M., Niemi, J., Thomsen, T., Vorobyev, V. A., and Hugdahl, K. (2005). Brain correlates of sentence translation in Finnish-Norwegian bilinguals. *Neuroreport* 16, 607–610.

Leonard, M. K., Brown, T. T., Travis, K. E., Gharapetian, L., Hagler, D. J. Jr., Dale, A. M., Elman, J. L., and Halgren, E. (2010). Spatiotemporal dynamics of bilingual word processing. *Neuroimage* 49, 3286–3294.

Marien, P., Abutalebi, J., Engelborghs, S., and De Deyn, P. P. (2005). Pathophysiology of language switching and mixing in an early bilingual child with subcortical aphasia. *Neurocase* 11, 385–398.

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40.

Moreno, E. M., Federmeier, K. D., and Kutas, M. (2002). Switching languages, switching palabras (words): an electrophysiological study of code switching. *Brain Lang.* 80, 188–207.

Moreno, E. M., Rodríguez-Fornells, A., and Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *J. Neurolinguistics* 21, 477–508.

Moritz-Gasser, S., and Duffau, H. (2009a). Cognitive processes and neural basis of language switching: proposal of a new model. *Neuroreport* 20, 1577–1580.

Moritz-Gasser, S., and Duffau, H. (2009b). Evidence of a large-scale network underlying language switching: a brain stimulation study. *J. Neurosurg.* 111, 729–732.

Moser, B. (1978). "Simultaneous interpretation: a hypothetical model and its practical application," in *Language, Interpretation and Communication,* eds D. Gerver and H. W. Sinaiko (New York, NY: Plenum Press), 353–368.

Moser-Mercer, B., Frauenfelder, U. H., Casado, B., and Kunzli, A. (2000). "Searching to define expertise in interpreting," in *Language Processing and Simultaneous Interpreting: Interdisciplinary Perspectives,* eds B. E. Dimitrova and K. Hyltenstam (Amsterdam: John Benjamins), 107–132.

Nagel, I. E., Schumacher, E. H., Goebel, R., and D'esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage* 43, 801–807.

Nardone, R., De Blasi, P., Bergmann, J., Caleri, F., Tezzon, F., Ladurner, G., Golaszewski, S., and Trinka, E. (2011). Theta burst stimulation of dorsolateral prefrontal cortex modulates pathological language switching: a case report. *Neurosci. Lett.* 487, 378–382.

Nieuwenhuis, S., Yeung, N., Van Den Wildenberg, W., and Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cogn. Affect. Behav. Neurosci.* 3, 17–26.

Osterhout, L., and Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *J. Mem. Lang.* 31, 785–806.

Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of "interference" in bilingual speakers: the activation threshold hypothesis. *Int. J. Psycholinguistics* 9, 133–145.

Paradis, M. (1996). Selective deficit in one language is not a demonstration of different anatomical representation: comments on Gomez-Tortosa et al. (1995). *Brain Lang.* 54, 170–173; discussion 174–175.

Paradis, M. (2001). An integrated neurolinguistic theory of bilingualism (1976–2000). *Paper Presented at the LACUS Forum 27,* Houston, TX.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.* 9, 456.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148.

Polich, J., and Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biol. Psychol.* 41, 103–146.

Price, C. J., Green, D. W., and Von Studnitz, R. (1999). A functional imaging study of translation and language switching. *Brain* 122(Pt 12), 2221–2235.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. Lang. Cogn.* 13, 253–262.

Proverbio, A. M., Leoni, G., and Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: an ERP study. *Neuropsychologia* 42, 1636–1656.

Quaresima, V., Ferrari, M., Van Der Sluijs, M. C. P., Menssen, J., and Colier, W. (2002). Lateral frontal cortex oxygenation changes during translation and language switching revealed by non-invasive near-infrared multi-point measurements. *Brain Res. Bull.* 59, 235–243.

Rinne, J. O., Tommola, J., Laine, M., Krause, B. J., Schmidt, D., Kaasinen, V., Teras, M., Sipila, H., and Sunnari, M. (2000). The translating brain: cerebral activation patterns during simultaneous interpreting. *Neurosci. Lett.* 294, 85–88.

Rodriguez-Fornells, A., Balaguer, R. D. D., and Münte, T. F. (2006). Executive control in bilingual language processing. *Lang. Learn.* 56, 133–190.

Rodriguez-Fornells, A., Rotte, M., Heinze, H. J., Nösselt, T., and Münte, T. F. (2002). Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature* 415, 1026–1029.

Schumacher, E. H., Cole, M. W., and D'esposito, M. (2007). Selection and maintenance of stimulus-response rules during preparation and performance of a spatial choice-reaction task. *Brain Res.* 1136, 77–87.

Sebastian-Galles, N., Rodriguez-Fornells, A., De Diego-Balaguer, R., and Diaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *J. Cogn. Neurosci.* 18, 1277–1291.

Steinhauer, K., and Connolly, J. F. (2008). "Event-related potentials in the study of language," in *Handbook of the Neuroscience of Language*, eds B. Stemmer and H. Whitaker (New York, NY: Elsevier), 91–104.

Takakusaki, K., Saitoh, K., Harada, H., and Kashiwayanagi, M. (2004). Role of basal ganglia-brainstem pathways in the control of motor behaviors. *Neurosci. Res.* 50, 137–151.

Van Heuven, W., and Dijkstra, T. (2010). Language comprehension in the bilingual brain: fMRI and ERP support for psycholinguistic models. *Brain Res. Rev.* 64, 104–122.

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

Wang, Y., Xue, G., Chen, C., Xue, F., and Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: an ER-fMRI study. *Neuroimage* 35, 862–870.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 March 2011; accepted: 29 August 2011; published online: 15 September 2011.

Citation: Hervais-Adelman AG, Moser-Mercer B and Golestani N (2011) Executive control of language in the bilingual brain: integrating the evidence from neuroimaging to neuropsychology. Front. Psychology ***2****:234. doi: 10.3389/fpsyg.2011.00234*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 13 December 2011
doi: 10.3389/fpsyg.2011.00374

The influence of cross-language similarity on within- and between-language Stroop effects in trilinguals

Walter J. B. van Heuven[1]*, Kathy Conklin[2], Emily L. Coderre[1], Taomei Guo[3] and Ton Dijkstra[4]

[1] School of Psychology, University of Nottingham, Nottingham, UK
[2] School of English Studies, University of Nottingham, Nottingham, UK
[3] State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China
[4] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, Netherlands

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
Noriko Hoshino, Kobe City University of Foreign Studies, Japan
Pedro Macizo, University of Granada, Spain

***Correspondence:**
Walter J. B. van Heuven, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK.
e-mail: walter.vanheuven@nottingham.ac.uk

This study investigated effects of cross-language similarity on within- and between-language Stroop interference and facilitation in three groups of trilinguals. Trilinguals were either proficient in three languages that use the same-script (alphabetic in German–English–Dutch trilinguals), two similar scripts and one different script (Chinese and alphabetic scripts in Chinese–English–Malay trilinguals), or three completely different scripts (Arabic, Chinese, and alphabetic in Uyghur–Chinese–English trilinguals). The results revealed a similar magnitude of within-language Stroop interference for the three groups, whereas between-language interference was modulated by cross-language similarity. For the same-script trilinguals, the within- and between-language interference was similar, whereas the between-language Stroop interference was reduced for trilinguals with languages written in different scripts. The magnitude of within-language Stroop facilitation was similar across the three groups of trilinguals, but smaller than within-language Stroop interference. Between-language Stroop facilitation was also modulated by cross-language similarity such that these effects became negative for trilinguals with languages written in different scripts. The overall pattern of Stroop interference and facilitation effects can be explained in terms of diverging and converging color and word information across languages.

Keywords: trilinguals, Stroop, interference, facilitation, script

INTRODUCTION

Proficient bilinguals are able to communicate in both of their languages without much difficulty. This is true whether the languages they speak are highly similar in terms of orthography and phonology (e.g., German and Dutch) or dissimilar (e.g., Chinese and English). The ease of bilingual communication is surprising in light of a large body of research that has demonstrated that lexical access is non-selective with respect to both language comprehension (e.g., Dijkstra et al., 1998; van Hell and de Groot, 1998; van Heuven et al., 1998; de Groot et al., 2000; Jared and Kroll, 2001; for a review, see Dijkstra and van Heuven, 2002) and language production (e.g., Hermans et al., 1998; Colomé, 2001; Guo and Peng, 2006; Costa et al., 2008; Hoshino and Kroll, 2008). Language non-selective access implies that word representations from both languages are active during processing, even when only one language is relevant to the situation or task at hand. Because an irrelevant language is often coactivated during processing, bilinguals must rely on cognitive control to respond in the appropriate language.

An important issue on the bilingual research agenda is how the interaction between languages is affected by their similarity or dissimilarity. It is important to explore, for instance, how cross-language similarity, in terms of phonological and/or orthographic/script overlap, influences the bilingual/multilingual language processing system and whether potential differences between languages, such as script, can be exploited to reduce the amount of cross-language interference, thereby influencing how much cognitive control is required to speak exclusively in the target language.

A task that is well-suited to investigate issues of cognitive control and cross-language similarity in bilingual processing is the Stroop task (Stroop, 1935). In a color naming Stroop task, color words are presented in colored ink and participants are asked to ignore the printed word and instead name the color of the ink. To avoid reading the printed word aloud, this task requires averting the highly practiced reading process. In incongruent conditions (where word and ink color do not match; e.g., "red" printed in green ink), the conflicting word and color information requires cognitive control and conflict resolution processes to be engaged, leading to a delay in response times (RTs) compared to control conditions (typically a non-linguistic or non-response set stimulus printed in colored ink; e.g., "XXXX" printed in blue). This delay is referred to as *Stroop Interference*. In contrast, *Stroop Facilitation* refers to the faster RTs in congruent conditions (where word and color match; e.g., "blue" printed in blue ink) than in control conditions. In a multilingual version of the Stroop task, input and output languages can be manipulated so that within- and between-language interference and facilitation effects can be investigated. In what follows, we will first discuss within- and between-language interference and then facilitation effects in the Stroop task.

STROOP INTERFERENCE: WITHIN- AND BETWEEN-LANGUAGE

In a traditional monolingual Stroop task, interference is generally thought to be due to conflicting color and word information (Roelofs, 2010). Thus, seeing "red" printed in green ink leads to long RTs for naming the ink color due to the diverging available information from the color and word. An important question is how this interference is modulated within- and between-languages. Does seeing "red" in blue ink when /blu/ is the required response in an English task, produce a similar amount of interference for German–English bilinguals as seeing "rot" ("red" in German)? In terms of semantics, both *red* and *rot* provide information that diverges from that of the ink color (blue), so one might expect similar degrees of within-language (intralingual) and between-language (interlingual) interference. However, based on his survey of the bilingual Stroop literature, MacLeod concludes that "*Interference between the two languages of a bilingual, although not as great as that within either one of the languages, is very robust: Between-language interference typically is about* 75% *of within-language interference*... There may also be differences in the processing of orthographic and idiographic languages..." (1991, p. 187).

MacLeod's view is supported by early research from Preston and Lambert (1969), who found that between-language interference was only 68% of the within-language interference for English–Hungarian bilinguals, but 95% for French–English bilinguals. Similarly, a study by Dyer (1971) with Spanish–English bilinguals showed that between-language effects were 63% of within-language ones. In a study with Chinese–English, Spanish–English, and Japanese–English bilinguals, Fang et al. (1981) also found greater within- than between-language interference. Interestingly, the between-language effect was modulated by the orthographic similarity of the two languages, such that more overlap lead to stronger Stroop interference in the between-language condition. Crucially, if orthographic similarity underpins the modulation of between-language interference, there should be a larger amount of Stroop interference when the two languages of bilinguals have similar scripts (e.g., alphabetic) than when the languages are written in different scripts (e.g., logographic and alphabetic).

The finding of larger within- than between-language Stroop interference in bilinguals (e.g., Fang et al., 1981; Mägiste, 1984; Tzelgov et al., 1990; Lee et al., 1992; Brauer, 1998) and trilinguals (Abunuwara, 1992) has been termed the within-language Stroop superiority effect (WLSSE; Goldfarb and Tzelgov, 2007). Research has demonstrated that the WLSSE is modulated by cross-language similarity and the proficiency of the participants (e.g., Preston and Lambert, 1969; Fang et al., 1981; Mägiste, 1984; Chen and Ho, 1986; Brauer, 1998; Sumiya and Healy, 2004, 2008). For example, Chen and Ho (1986) conducted a Stroop task with Chinese–English bilinguals in five different age groups. When responses were in the first language (L1) Chinese, all age groups showed greater within- than between-language interference. When responses were in the second language (L2) English, there was a shift from greater between-language interference for the youngest group to greater within-language interference for the oldest three groups.

Similarly, Brauer (1998) conducted two Stroop studies with high and low proficiency bilinguals in languages with high (German, English) and low (English–Greek or English–Chinese) overlap. He found that the low proficiency bilinguals, regardless of how much the languages overlapped, showed more within- than between-language interference when they were required to respond in their L1, and the opposite pattern when they responded in their L2. In the case of high proficiency participants speaking languages with no overlap, there was greater within- than between-language interference when they responded both in the L1 and in the L2. Finally, in high proficiency participants of languages with high overlap, there was no difference between within- and between-language interference effects.

In Sumiya and Healy (2004), Japanese–English bilinguals engaged in a Stroop task in Japanese and English. Color words were used that were phonologically similar across the two languages (/bru:/ and /blu/, with Katakana and English scripts, respectively) or different (/ao/ and /blu/, with Hiragana and English scripts, respectively). Even though script provided a strong cue about the task-relevant language, a significant between-language Stroop effect arose that was larger for phonologically similar words. In a similar study with English–Japanese bilinguals, Sumiya and Healy (2008) found a between-language interference effect that was larger for phonologically similar words, in particular when responses were in L2 Japanese. Additionally, the size of the phonological effect increased with proficiency in Japanese, which was taken as an indication of increased phonological processing when speakers were more proficient in their L2. Such results suggest that not only the degree of form overlap (orthographic and phonological) may modulate interference effects in trilinguals, but proficiency and response language (L1 or L2) as well.

With respect to the WLSSE, it must be considered that when the language of the written word is different from the response language, the influence of the irrelevant language might be minimized through inhibitory control (Green, 1998) or decision criteria (Dijkstra and van Heuven, 2002). Alternatively, response set competition might be involved (Roelofs, 2003; Goldfarb and Tzelgov, 2007). Goldfarb and Tzelgov (2007) examined the cause of the WLSSE by having Hebrew–English bilinguals name an ink color when the distractor was either a color word (*red*, *green*, *blue*) or a color associated word (*tomato*, *grass*, *sky*). In the between-language condition, both the color and associated words were in the irrelevant language. However, color words, but not color associated words, demonstrated larger within- than between-language effects. It was proposed that in the case of color words in the between-language condition, activation at the conceptual level provides activation for items in the response set, while activation at the lexical level does not. This would then induce less interference than in the within-language condition, where the color word activates items in the response set at both the semantic and lexical levels, thereby increasing competition. In the case of color associated words, neither the words in the within-language condition nor in the between-language condition are part of the response set; therefore the WLSSE was not observed.

STROOP FACILITATION: WITHIN- AND BETWEEN-LANGUAGE

Let us now focus on Stroop facilitation, which arises from the difference between responses in the congruent condition (e.g., the word "red" written in red ink) and the control condition (e.g., row of X's in red ink). Typically, congruent trials are processed faster than control trials. There is disagreement in the Stroop literature about the locus of Stroop facilitation. According to the *converging information hypothesis*, in congruent trials the information available from the ink color and the word "converge": they support the correct response, which leads to faster RTs (e.g., Cohen et al., 1990; Melara and Algom, 2003; Roelofs, 2003, 2010). According to the *inadvertent reading hypothesis*, there are attentional lapses on some trials that result in the color word being read out instead of the ink color being named (MacLeod and MacDonald, 2000; Kane and Engle, 2003). In incongruent trials, this inadvertent reading yields an incorrect response. However, in congruent trials such errors are undetectable. They therefore contaminate RTs with incorrect responses to the color word and may lead to apparent but invalid facilitation effects. Research with bilinguals offers a way of testing the two hypotheses (Roelofs, 2010), because a between-language version of the Stroop task allows such reading errors to be detected (e.g., reading the German "blau" printed in blue ink when producing /blu/ is the required response). If previously observed Stroop facilitation effects arise from undetected covert reading errors, they can be eliminated in a between-language task where such errors are apparent. Thus, any facilitation in the between-language condition is not underpinned by invalid facilitation and would be due to converging information (Roelofs, 2010).

Neither the inadvertent reading hypothesis nor the converging information hypothesis makes explicit predictions about how cross-language similarity might affect Stroop facilitation effects (however, see Roelofs, 2010, for an account of how diverging information at the word form level affects Stroop facilitation). It could be the case that inadvertent reading would not occur when script provides a strong cue. However, even if inadvertent reading occurs less when scripts are different, the claim that there will be more within- than between-facilitation still stands. Because facilitation is underpinned by inadvertent reading, these trials are removed in the between-language condition regardless of script or language overlap (unless the color words are absolutely identical in pronunciation) and thus the invalid facilitation is removed. To summarize, according to the inadvertent reading hypothesis, bilinguals and multilinguals should show within- but not between-language facilitation. In contrast, if facilitation stems from converging information, it should occur whenever word and color information converge. However, the question remains as to whether the degree of cross-language convergence modulates the facilitation effect.

Even though bilingual and multilingual research can shed light on whether Stroop facilitation is caused by inadvertent reading or converging information, not many studies in the Stroop literature have focused on facilitation effects in bilinguals and multilinguals. The limited research indicates that between-language Stroop facilitation is modulated by cross-language similarity, such that a negative or interference effect is apparent when languages are more dissimilar. Thus, when the color and the meaning of the word are congruent and the input and output languages differ (i.e., the presented word is a translation equivalent of the word that has to be produced), responses are delayed relative to a control condition when the languages are different, whereas responses are faster when they are similar. For example, Abunuwara (1992) conducted a Stroop task with Arabic–Hebrew–English trilinguals. Although not reported or analyzed as such, congruent trials in the within-language condition yielded a 45-ms facilitation effect relative to the control condition. In contrast, congruent trials in the between-language condition lead to an interference effect of 58 ms. The presence of interference supports the view that the irrelevant language is activated and slows RTs to the ink color. Furthermore, in Experiment 4 of Roelofs (2010) with Dutch–English bilinguals, the between-language facilitation effect appears to be absent at a stimulus onset asynchrony (SOA) of 0 ms. In contrast, MacLeod and MacDonald (2000) reported interference in French–English bilinguals in the between-language congruent condition.

In sum, the review of previous Stroop-research in bilinguals and multilinguals suggests that within-language facilitation should arise in all languages of trilinguals, irrespective of the script involved. However, the picture is less clear for the between-language congruent condition, which may or may not elicit faster RTs compared to the appropriate control condition and might even yield slower RTs. There is, as far as we know, only one study in the literature that has looked at Stroop effects in trilinguals (Abunuwara, 1992). However, this study only focused on Stroop interference and involved only a group of different script trilinguals. Thus, the current research is the first to consider the nature of between-language facilitation in trilinguals whose languages overlap to varying degrees in terms of their script and the orthographic/phonological similarity of their color word translations.

THE PRESENT STUDY

The above literature review suggests that within-language Stroop interference should be apparent in all three languages of trilinguals irrespective of the script involved. Between-language interference should overall be less than within-language interference. In addition, it may be modulated by factors such as script similarity and/or the form overlap (orthographic and phonological) of the color word translations, such that increased similarity may lead to more between-language interference. In terms of Stroop facilitation, there should be evidence of within-language facilitation that is unaffected by script in all languages of the trilinguals. However, previous research is equivocal on whether faster naming responses would be expected in the between-language congruent condition. If between-language Stroop facilitation arises, it might be modulated by language similarity, such that there is more Stroop facilitation when the languages have greater overlap.

In the present study, three groups of trilinguals performed a Stroop color naming task that involved three colors (red, green, and blue). The response language was blocked and the stimulus language was manipulated within each block. Two control conditions were included in each block: a color patch and a control stimulus (e.g., %). The results were analyzed in terms of within- and between-language Stroop interference (incongruent–control stimulus) and facilitation (control stimulus–congruent) to investigate whether cross-language similarity modulates between-language interactions.

Experiment 1 was conducted with German–English–Dutch (GED) trilinguals. In German, English, and Dutch, the color word translations (e.g., *rot*–*red*–*rood*) overlapped not only in terms of semantics but also in script (all alphabetic), orthography, and phonology (same orthographic/phonological onset). Experiment 2 involved Chinese–English–Malay (CEM) trilinguals. In Chinese, English, and Malay, the script is shared in English and Malay (both alphabetic) but differs from Chinese. Furthermore, orthography and phonology of the color word translations differ across all three languages [e.g., 红 (hong)–*red*–*merah*]. Finally, Experiment 3 was conducted with Uyghur–Chinese–English (UCE) trilinguals. The color translations between Uyghur, Chinese, and English are completely different in terms of script, orthography, and phonology [e.g., قىزىل (gizil)–红 (hong)–*red*].

EXPERIMENT 1: GERMAN–ENGLISH–DUTCH TRILINGUALS

METHOD

Participants

Thirty GED trilinguals (eight males) participated in the experiment. All participants studied at the Radboud University in Nijmegen, the Netherlands. They were first language German speakers proficient in English and Dutch. Furthermore, most of them were also proficient in one or more other languages (e.g., French, Spanish, Italian). **Table 1** provides an overview of their mean age and their subjective proficiency scores for each language (scale: from 1 = very poor to 7 = fluent), as well as the year of the first contact with each language and the number of years of experience with each language. In this and the following experiments, the order of the year of first contact with each of the three languages was used to determine the first (L1), second (L2), and third (L3) language of the trilinguals.

MATERIALS AND DESIGN

The stimuli used in the Stroop task were the English color words *red*, *green*, and *blue*, the corresponding color words in Dutch (*rood*, *groen*, *blauw*) and German (*rot*, *grün*, *blau*), control stimuli (row of percent signs), and color patches of red, green, and blue. Colored rectangles about 10 cm × 5 cm (248 × 142 pixels) were used to present the colors. The center of each colored rectangle contained a small black rectangle (142 × 42 pixels) with a color word or control stimulus presented in a white lowercase Courier font (32-point). For each language, a control stimulus was constructed that matched the length of the color word in the languages (e.g., %%% for *red* and *rot*; %%%% for *rood*). The color patch controls were fully colored rectangles (248 × 142 pixels). In total 39 stimuli were created (3 color patches, 9 control stimuli, 9 congruent stimuli, and 18 incongruent stimuli) that were repeated a number of times in such a way that for each output language there were 108 trials: 36 congruent, 36 incongruent, 18 control stimuli, and 18 color patches.

PROCEDURE AND APPARATUS

A Sennheiser headset (PC 161) was connected to a PC and DMDX (Forster and Forster, 2003) was used to present the stimuli, to measure the voice onset latency and to record the vocal response. Each trial started with a fixation sign (+) presented for 500 ms at the center of the 17″ monitor (1024 × 768 pixels, 85 Hz). Next, a blank screen appeared for 300 ms and then the stimulus was presented until the participant responded vocally or for 2000 ms. After 1000 ms the next trial started. Participants were instructed to ignore the letter strings (color words and control stimuli) and to overtly name the color of the rectangle as fast as possible without making any errors. Participants performed the Stroop task in each of their three languages separately. Thus, output language was blocked. At the beginning of each block the required output language was indicated. The order of the output language was counterbalanced across participants. Within blocks all stimuli were randomized differently for each participant so that in contrast to the output language the input language was randomized within blocks. After the experiment participants filled out a language background questionnaire.

ANALYSIS

CheckVocal (Protopapas, 2007) was used to check whether vocal responses were correct and to find and correct voice key errors. Responses outside ±2.5 SD of each subject mean across all trial types were excluded. For the RT analysis, erroneous responses were removed and the mean RTs were calculated. In all ANOVAs a Greenhouse–Geisser correction was applied when the assumption of sphericity was violated, and all reported *p*-values from *post*

Table 1 | Subjective proficiency scores (scale: 1 = very poor to 7 = fluent) and subject demographics for the trilinguals in Experiments 1–3.

Trilinguals	*n*	Age	Language	Subjective proficiency scores					First contact	Years of experience
				Speaking	Listening	Reading	Writing	Overall		
Experiment 1:	30	23.2	German	6.9	7.0	7.0	6.7	6.9	0.0	22.9
German–English–Dutch (GED)			English	4.4	5.5	5.8	4.4	5.0	9.4	12.4
			Dutch	4.8	5.7	6.1	4.7	5.3	19.1	3.7
Experiment 2:	24	21.8	Chinese	6.6	6.6	5.8	5.0	6.0	1.6	18.4
Chinese–English–Malay (CEM)			English	5.3	5.7	5.8	5.4	5.5	3.7	17.0
			Malay	3.8	5.0	5.1	4.0	4.5	5.6	14.3
Experiment 3:	32	22.4	Uyghur	6.3	6.6	6.1	5.8	6.2	0.2	22.1
Uyghur–Chinese–English (UCE)			Chinese	5.2	6.3	5.7	5.0	5.5	8.7	14.1
			English	3.7	4.5	4.7	3.8	4.2	15.2	7.3

Table 2 | Mean RTs and SE of the congruent, incongruent, and control conditions for each input and output language combination in Experiments 1–3.

	Input	Output language					
		German		**English**		**Dutch**	
		Congruent	**Incongruent**	**Congruent**	**Incongruent**	**Congruent**	**Incongruent**
Experiment 1, GED	German	544 (12)	653 (17)	569 (14)	651 (14)	599 (12)	695 (13)
	English	571 (13)	625 (11)	559 (14)	648 (14)	644 (17)	687 (15)
	Dutch	556 (13)	669 (18)	572 (11)	653 (16)	592 (13)	685 (15)
	%%%%	574 (12)		577 (13)		618 (12)	
	Patch	558 (10)		577 (13)		596 (11)	
		Chinese		**English**		**Malay**	
		Congruent	**Incongruent**	**Congruent**	**Incongruent**	**Congruent**	**Incongruent**
Experiment, 2, CEM	Chinese	545 (18)	618 (20)	564 (16)	603 (22)	596 (20)	596 (18)
	English	584 (20)	609 (23)	537 (19)	623 (20)	593 (22)	644 (24)
	Malay	573 (19)	617 (15)	571 (20)	640 (25)	549 (18)	665 (21)
	%%%%	573 (21)		563 (19)		577 (15)	
	Patch	577 (20)		552 (18)		566 (14)	
		Uyghur		**Chinese**		**English**	
		Congruent	**Incongruent**	**Congruent**	**Incongruent**	**Congruent**	**Incongruent**
Experiment 3, UCE	Uyghur	637 (20)	724 (27)	711 (20)	761 (27)	761 (24)	789 (30)
	Chinese	675 (20)	704 (27)	690 (19)	769 (23)	734 (22)	777 (26)
	English	683 (21)	705 (24)	760 (26)	766 (22)	695 (20)	822 (29)
	%%%%	642 (19)		716 (23)		731 (19)	
	Patch	651 (21)		692 (20)		742 (24)	

GED, German–English–Dutch; CEM, Chinese–English–Malay; UCE, Uyghur–Chinese–English.

hoc t-tests were Bonferroni-corrected. We first calculated Stroop interference and facilitation effects based on the raw means and then analyzed whether the magnitude of interference and facilitation effects were modulated by input and output languages. Because error rates were very low (<1.7%) no error analyses were conducted.

RESULTS AND DISCUSSION

The total percentage of errors was 1.66%, and the percentage of outliers was 0.79%. The mean RTs for all conditions are presented in **Table 2**. In all subsequent analyses, we treated the control character as the control condition when calculating Stroop interference and facilitation effects. Separate analysis of the control conditions (control characters vs. color patch) by means of a two-way ANOVA on the mean RTs across all language outputs revealed a significant effect of control type, $F(1,29) = 12.27$, $p < 0.01$, $\eta^2 = 0.01$. Analyses using the control patch yielded the same main effects and interactions as the analyses using the control character.[1]

STROOP INTERFERENCE (INCONGRUENT VS. CONTROL CONDITION)

The means of Stroop interference for all input and output language combinations are presented in **Table 3**. Bonferroni-corrected paired-sample t-tests revealed significant Stroop interference, with all ps < 0.0001, for each input and output language combination.

To investigate whether the magnitude of interference varied across input and output languages, a 3 (input language) × 3 (output language) repeated-measures ANOVA was conducted. This analysis revealed a significant interaction between input and output language, $F(4,116) = 2.52$, $p < 0.05$, $\eta^2 = 0.03$. For the German output, there was a significant main effect of input language, $F(2,58) = 6.95$, $p < 0.01$, $\eta^2 = 0.09$. *Post hoc* paired-sample t-tests showed a smaller Stroop interference for English input (51 ms, SE = 7 ms) than for German (79 ms, SE = 12 ms), $t(29) = 2.54$,

[1] A 3 (input) × 3 (output) repeated-measures ANOVA of the Stroop interference comparison showed there was an interaction of input and output language, $F(4,116) = 2.52$, $p < 0.05$. A significant effect was found of language input on German output, $F(2,58) = 6.95$, $p < 0.01$, but not in English or Dutch output. A similar Stroop facilitation analysis revealed an interaction between input and output, $F(4,116) = 10.77$, $p < 0.0001$, due to a significant effect of language input for German output, $F(2,58) = 6.12$, $p < 0.01$, and Dutch output, $F(2,58) = 27.42$, $p < 0.0001$, but not English output. A 2 (type: within or between) × 3 (output) repeated-measures ANOVA showed a significant effect of type on the magnitude of the Stroop facilitation effect, $F(1,29) = 40.49$, $p < 0.0001$, but not on the magnitude of the Stroop interference effect. As mentioned, all these effects do not differ from those found with the control character.

Table 3 | Magnitude of the Stroop interference and facilitation effects in Experiments 1–3.

	Input language	Interference (incongruent–control)			Facilitation (control–congruent)		
		Output language			Output language		
		German	**English**	**Dutch**	**German**	**English**	**Dutch**
Experiment 1, GED	German	79 (12)***	74 (9)***	77 (7)***	30 (8)*	7 (7)	20 (6)¥
	English	51 (7)***	72 (7)***	69 (8)***	3 (9)	18 (7)	−26 (10)
	Dutch	95 (13)***	77 (8)***	67 (10)***	18 (8)	5 (7)	26 (8)¥
		75	74	71	17	10	7
		Chinese	**English**	**Malay**	**Chinese**	**English**	**Malay**
Experiment 2, CEM	Chinese	45 (9)**	40 (11)¥	19 (12)	28 (10)§	−1 (7)	−19 (11)
	English	36 (9)*	60 (10)***	67 (15)*	−11 (12)	26 (8)¥	−16 (12)
	Malay	44 (11)*	77 (12)***	88 (15)***	1 (14)	−9 (10)	28 (11)
		42	59	58	6	5	−2
		Uyghur	**Chinese**	**English**	**Uyghur**	**Chinese**	**English**
Experiment 3, UCE	Uyghur	81 (14)***	45 (12)*	57 (18)¥	6 (7)	5 (8)	−29 (12)
	Chinese	62 (16)*	52 (11)**	45 (12)*	−32 (10)¥	26 (12)	−4 (11)
	English	63 (13)**	50 (13)*	91 (16)***	−40 (10)*	−44 (13)¥	36 (8)*
		69	49	64	−22	−4	1

*Magnitudes are shown in milliseconds with SE in parentheses. Significant effects after Bonferroni corrections are indicated: § p < 0.10; ¥ p < 0.05; *p < 0.01; **p < 0.001; ***p < 0.0001. GED, German–English–Dutch; CEM, Chinese–English–Malay; UCE, Uyghur–Chinese–English. A negative value for Stroop facilitation indicates that congruent condition was slower than the control condition.*

$p = 0.05$ corrected, $\eta^2 = 0.18$, and for Dutch input (95 vs. 51 ms), $t(29) = 3.38$, $p < 0.01$, $\eta^2 = 0.28$. In contrast, no effect of language input was found on Stroop interference for English output, $F(2,58) < 1$, and Dutch output, $F(2,58) < 1$.

To compare within- vs. between-language effects, we performed a 2 (type: within or between) × 3 (output language) repeated-measures ANOVA. This analysis revealed no main effect of type, $F(1,29) < 1$, indicating an equal amount of Stroop interference within-languages (73 ms, SE = 10 ms) and between-languages (74 ms, SE = 9 ms). No effect of language output or interaction between type and output language was found (all *p*s > 0.58). Equal within- and between-language interference has previously been reported with Dutch–English bilinguals (Roelofs, 2010). This can be explained in terms of the high cross-language similarity between the color word translations in terms of orthography and phonology (*red–rood–rot*, *green–groen–grun*, and *blue–blauw–blau* in English, Dutch, and German respectively). In fact, most of these translations can be considered to be cognates (same meaning and very similar orthography and phonology across languages). There is a wide literature that suggests that cognates have a special status in the multilingual lexicon because their processing differs from matched non-cognates (e.g., Dijkstra et al., 1998, 1999; Costa et al., 2000; van Hell and Dijkstra, 2002; Lemhöfer and Dijkstra, 2004; Lemhöfer et al., 2004; Hoshino and Kroll, 2008). If there is a special status for these cognates and they are activated in parallel across the three languages, within- and between-language interference should be similar.

STROOP FACILITATION (CONTROL VS. CONGRUENT CONDITION)

The size of the Stroop facilitation across input and output languages is also shown in **Table 3**. Significant Stroop facilitation was found for German output when the input language was German, $t(29) = 3.75$, $p < 0.01$, $\eta^2 = 0.33$. Furthermore, facilitation was observed for Dutch output when Dutch was the input language, $t(29) = 3.49$, $p < 0.05$, $\eta^2 = 0.30$, and German, $t(29) = 3.20$, $p < 0.05$, $\eta^2 = 0.26$.

The 3 (input language) × 3 (output language) repeated-measures ANOVA on the magnitude of Stroop facilitation effects revealed an interaction between input and output languages, $F(4,116) = 10.77$, $p < 0.0001$, $\eta^2 = 0.07$. For German output, a significant effect of language input was found, $F(2,58) = 6.12$, $p < 0.01$, $\eta^2 = 0.06$. Paired-sample *t*-tests revealed only a significant difference in facilitation magnitude between German and English input (30 ms, SE = 8 vs. 3 ms, SE = 9 ms), $t(29) = 3.14$, $p < 0.05$, $\eta^2 = 0.25$. For English output, no effect of language input was found, $F(2,58) = 1.22$, $p = 0.30$. A significant effect of language input was found for Dutch output, $F(1.62,47.0) = 27.42$, $p < 0.0001$, $\eta^2 = 0.22$. Paired-sample *t*-tests showed a significant difference in the magnitude of facilitation effects for German and English input (20 ms, SE = 6 vs. −26 ms, SE = 10 ms), $t(29) = 5.14$, $p < 0.0001$, $\eta^2 = 0.48$, and for Dutch and English input (26 ms, SE = 8 vs. −26 ms, SE = 10 ms), $t(29) = 6.38$, $p < 0.0001$, $\eta^2 = 0.58$.

In contrast to the analyses with respect to Stroop interference, the 2 (type: within or between) × 3 (output language) repeated-measures ANOVA revealed a significant main effect of type,

$F(1,29) = 40.49$, $p < 0.0001$, $\eta^2 = 0.05$, such that larger Stroop facilitation occurred within-languages (25 ms, SE = 7 ms) than between-languages (4 ms, SE = 8 ms). No effect of language output or interaction between type and output language was found (all $ps > 0.21$). Thus, Stroop facilitation was absent between-languages. Roelofs (2010) found an equal amount of within- and between-language Stroop facilitation in Dutch–English bilinguals in a Stroop task that separated color and word information in time (SOA manipulation). However, the experiment in Roelofs's study that was most comparable to the current one (color words and control conditions fully crossed), Stroop facilitation within- and between-languages was not significant at the 0-ms SOA (Experiment 4). Unfortunately, the analysis of that experiment was collapsed across within- and between-languages; thus it is unclear whether at the 0-ms SOA, within- and between-facilitation differed. Close inspection of the graphs (Figure 8 in Roelofs, 2010) shows that the within-language facilitation was in fact larger than the between-language facilitation. This suggests that the current results with GED trilinguals are very comparable to those of the Dutch–English bilinguals.

To summarize, in Experiment 1 with GED trilinguals, we found an equal amount of Stroop interference within- and between-languages, but Stroop facilitation was stronger within- than between-languages. To investigate whether similarity between the involved languages modulates these effects, a second experiment was conducted with trilinguals for whom the cross-language similarity between the color word translations was much less: two were alphabetic languages, Malay and English, and one was a logographic language, Chinese.

EXPERIMENT 2: CHINESE–ENGLISH–MALAY TRILINGUALS

METHOD

Participants

In this experiment 24 CEM trilinguals (10 males) participated. Participants were born in Malaysia and had received formal education in Mandarin, Malay, and English. They arrived in the UK between the age of 15 and 23 ($M = 21.8$ years) and were studying at the University of Nottingham (United Kingdom) at the time of testing. Most of the trilinguals could speak both Mandarin and Cantonese and rated their Cantonese proficiency higher than their Mandarin proficiency (see **Table 1**). However, seven trilinguals considered themselves more proficient in Mandarin than Cantonese. Several participants could also understand and speak other spoken Chinese dialects (e.g., Hakka). **Table 1** provides an overview of their mean age and their subjective proficiency scores for each language, as well as their first contact and years of experience with each language.

MATERIALS AND DESIGN

The design was identical to Experiment 1. The word stimuli were the English color words *red*, *green*, and *blue* (same as in Experiment 1), and their Malay (*merah*, *hijau*, *biru*) and Chinese translations. For 18 of the participants, the Chinese words were presented in Cantonese (traditional Chinese script): 紅 (hung), 綠 (luk), 藍 (laam), whereas for the six participants who rated themselves more proficient in Mandarin, the Chinese words were presented in Mandarin (simplified Chinese script): 红 (hong), 绿 (lu), 蓝 (lan). English and Malay words were presented in 32-point lowercase Courier font, and Chinese characters were presented in 32-point STHeiti font.

PROCEDURE AND APPARATUS

Same as Experiment 1.

ANALYSIS

Same as Experiment 1.

RESULTS AND DISCUSSION

The total percentage of outliers (0.71%) and the total percentage of errors (1.83%) were again very low. The mean RTs for all conditions are shown in **Table 2**. A two-way (control type: character or patch) ANOVA on the mean RTs across all language outputs showed no significant effect of control type, $F(1,23) = 1.84$, $p = 0.19$; therefore the control character was used in subsequent analyses.

STROOP INTERFERENCE (INCONGRUENT VS. CONTROL CONDITION)

Significant Stroop interference across input and output language combinations were found, $ps < 0.05$ for Bonferroni-corrected paired-sample t-tests, except for Chinese input and Malay output (see **Table 3**).

A 3 (input language) × (output language) repeated-measures ANOVA revealed an interaction between input and output language, $F(4,92) = 4.26$, $p < 0.01$, $\eta^2 = 0.04$. There was no effect of input language for Chinese output, $F(2,46) < 1$. However, for English output an effect of language input, $F(2,46) = 4.67$, $p < 0.05$, $\eta^2 = 0.07$, was found, with significant differences in Stroop interference between the Chinese and Malay input (40 ms, SE = 11 vs. 77 ms, SE = 12 ms), $t(23) = 2.90$, $p < 0.05$, $\eta^2 = 0.27$. Malay output also revealed a significant effect of language input, $F(2,46) = 11.21$, $p < 0.001$, $\eta^2 = 0.16$, with significant differences between the Chinese and English input (19 vs. 67 ms), $t(23) = 2.93$, $p < 0.05$, $\eta^2 = 0.27$, and between the Chinese and Malay input (19 ms, SE = 12 vs. 88 ms, SE = 15 ms), $t(23) = 4.44$, $p < 0.001$, $\eta^2 = 0.46$.

The 2 (type: within or between) × 3 (language output) repeated-measures ANOVA showed a significant effect of type, $F(1,23) = 6.50$, $p < 0.05$, $\eta^2 = 0.02$, such that the magnitude of interference was larger within-languages (64 ms, SE = 12 ms) than between-languages (47 ms, SE = 12 ms). Remarkably, the between-language interference is 73% of within-language interference. This percentage is very similar to the percentage of 74% reported by Francis (1999), and of 75% reported by MacLeod (1991), which were based on a review of studies with bilinguals. No main effect of output arose, $F(2,46) = 2.27$, $p = 0.11$, but there was an interesting interaction between type and output, $F(2,46) = 5.53$, $p < 0.01$, $\eta^2 = 0.02$. To follow up on this interaction, we ran a two-way (type) ANOVA for each language output. Interestingly, this analysis revealed that the Chinese and English output showed no effect of type, $F(1,23) < 1$, all $ps > 0.64$, whereas a significant effect of type was found for the Malay output, $F(1,23) = 15.20$, $p < 0.001$, $\eta^2 = 0.08$, such that the magnitude of within-language interference for Malay output was larger than between-language (Malay: 88 ms, SE = 15 ms vs. English and Chinese: 43 ms, SE = 14 ms).

Because the CEM trilinguals have two languages that share the same-script, we looked at the within- and between-language interference for same-script language pairs (Malay–English) and different script language pairs (Malay–Chinese and English–Chinese). The data revealed that for same-script languages the within- and between-language interference was similar (within: 74 ms, between: 72 ms), whereas for different script languages the between-language interference was reduced to 72% of the within-language interference for Chinese–English (within: 53 ms, between: 38 ms) and to 48% for Chinese–Malay (within: 67 ms, between: 32 ms). Thus, for the Malay–English language combination an equal amount of within- and between-language interference was found, which is consistent with the results of the GED trilinguals. Remarkably, the size of the within- and between-language interference was similar as well (GED within: 73 ms, between: 74 ms vs. CEM same-script within: 74 ms, between 72 ms). Importantly, the color words in Malay and English do not overlap in terms of orthography and phonology, except for the color word *blue* and Malay *biru*. To investigate whether the orthographic and phonological overlap of the color word blue affected the interference effects in the CEM, we analyzed the data after excluding the *biru* trials. This analysis again revealed similar within- and between-language interference (within: 65 ms, between: 66 ms) for the same-script languages (Malay and English). Thus, the similarly sized within- and between-language interference suggests that between-language interference is not stronger because the color word translations are orthographically/phonologically similar (cognates) but because they are written in the same-script. Therefore, the reduction of between-language interference in different script languages might be due to the use of script as a cue to reduce interference. This was tested in Experiment 3, involving trilinguals with three languages that differ in orthography/script and phonology.

STROOP FACILITATION (CONTROL CHARACTER VS. CONGRUENT CONDITION)

Significant facilitation was only found for English input and output, $t(23) = 3.10$, $p < 0.05$, $\eta^2 = 0.29$ (see **Table 3**). Again, an interaction was found between input and output languages in the 3 (input language) × 3 (output language) repeated-measures ANOVA on the magnitude of Stroop facilitation effects, $F(3.0,68.4) = 9.19$, $p < 0.0001$, $\eta^2 = 0.10$. For Chinese output, an effect of language input was found, $F(2,46) = 7.96$, $p < 0.01$, $\eta^2 = 0.07$. Paired-sample t-tests identified the difference in Stroop facilitation between Chinese and English input (28 ms, SE = 10 vs. −11 ms, SE = 12 ms), $t(23) = 3.89$, $p < 0.01$, $\eta^2 = 0.40$, and between Chinese and Malay input (28 ms, SE = 10 vs. 1 ms, SE = 14 ms), $t(23) = 2.65$, $p < 0.05$, $\eta^2 = 0.23$. The English output group also showed a significant effect of language input, $F(2,46) = 6.91$, $p < 0.01$, $\eta^2 = 0.12$, with significant differences between the Chinese and English input (−1 ms, SE = 7 vs. 26 ms, SE = 8 ms), $t(23) = 3.46$, $p < 0.01$, $\eta^2 = 0.34$, and between the Malay and English input (−9 ms, SE = 10 vs. 26 ms, SE = 8 ms), $t(23) = 3.19$, $p < 0.05$, $\eta^2 = 0.31$. For Malay output a significant effect of language input was also found, $F(2,46) = 6.75$, $p < 0.01$, $\eta^2 = 0.13$. This effect was due to differences in Stroop facilitation between the English and Malay input (−16 ms, SE = 12 vs. 28 ms, SE = 11 ms), $t(23) = 3.13$, $p < 0.05$, $\eta^2 = 0.30$, and between Chinese and Malay (−19 ms, SE = 11 vs. 28 ms, SE = 11 ms), $t(23) = 3.77$, $p < 0.01$, $\eta^2 = 0.38$.

The 2 (type: within or between) × 3 (output language) repeated-measures ANOVA revealed only a significant effect of type, $F(1,23) = 62.11$, $p < 0.0001$, $\eta^2 = 0.10$, such that the magnitude of facilitation was larger within-languages (27 ms, SE = 9 ms) than between-languages (−9 ms, SE = 11 ms). There was no main effect of output, $F(1.6,35.7) < 1$, and no interaction, $F(2,46) < 1$.

Interestingly, the between-language Stroop facilitation effect became negative. Thus, responses to the congruent Stroop condition (e.g., for English output: blue colored rectangle with the Chinese translation of blue written in the center) were slower than to the control condition. However, this negative Stroop facilitation effect was not significant across the different combinations of input and output languages. As discussed in the Introduction, between-language interference for Stroop facilitation has been observed before (Dalrymple-Alford, 1968; Abunuwara, 1992; MacLeod and MacDonald, 2000). Only numerically is there interference for same and different script language combinations, so it remains unclear whether script similarity modulates between-language Stroop facilitation. If script similarity does modulate the between-language facilitation, the interference effect should become larger and thus become significant when all three languages of the trilinguals differ in script. This was investigated in the next experiment, which involved trilinguals who were proficient in three languages that are written in different scripts.

EXPERIMENT 3: UYGHUR–CHINESE–ENGLISH TRILINGUALS

METHOD

Participants

Thirty-two UCE trilinguals (12 males) participated in the study. All participants were native Uyghur speakers who were born in Xinjiang, China. They received formal education in Uyghur and Mandarin and at the time of testing studied at Beijing Normal University, Beijing, China. **Table 1** presents an overview of their mean age and their subjective proficiency scores for each language, as well as their first contact and years of experience with each language.

MATERIALS AND DESIGN

The design was identical to that of the previous experiments. The only difference was that in addition to the English color words *red*, *green*, and *blue*, the word stimuli consisted of their translations in Chinese (Mandarin, simplified Chinese script): 红 (hong), 红 (lu), 绿 (lan), and Uyghur (Arabic script): قىزىل (gizil), يېشىل (yéshil), كۆك (kök). English words were presented in 32-point lowercase Courier font, Chinese characters in 32-point STHeiti font and Uyghur words in 32-point Geeza Pro font.

PROCEDURE AND APPARATUS

Same as in the previous experiments.

ANALYSIS

Same as in the previous experiments.

RESULTS AND DISCUSSION

The total percentage of errors was 1.33%, and the total percentage of outliers was 0.74%. The mean RTs for all conditions are shown in **Table 2**. An analysis on the two control conditions (character and color patch) using a two-way (control type: character or patch) repeated-measures ANOVA revealed, as in Experiment 2, no main effect of control stimulus type, $F(1,31) < 1$. Therefore, the control character was used in subsequent analyses.

STROOP INTERFERENCE (INCONGRUENT VS. CONTROL CONDITION)

Across all input and output language combinations significant Stroop interference effects were found, ps < 0.05 in all Bonferroni-corrected paired-sample t-tests (see **Table 3**).

The 3 (input language) × 3 (output language) repeated-measures ANOVA on the magnitude of Stroop interference effects revealed a trend toward a significant interaction between input and output languages, $F(4,124) = 2.18$, $p = 0.08$, $\eta^2 = 0.02$. No effect of language input was found for Uyghur output, $F(2,62) = 1.10$, $p = 0.34$, or Chinese output, $F(2,62) < 1$. In contrast, for English output a significant effect of input language was found, $F(2,62) = 4.55$, $p < 0.05$, $\eta^2 = 0.05$, with a significant difference in Stroop interference between Chinese (45 ms, SE = 12 ms) and English (91 ms, SE = 16 ms) inputs, $t(31) = 3.10$, $p < 0.05$, $\eta^2 = 0.24$.

The 2 (type: within or between) × 3 (output language) repeated-measures ANOVA revealed a main effect of type, $F(1,31) = 8.92$, $p < 0.01$, $\eta^2 = 0.02$, such that there was stronger interference for within (75 ms, SE = 14 ms) than between-languages (54 ms, SE = 14 ms), but no effect of output, $F(2,62) = 1.57$, $p = 0.22$, and no interaction, $F(12,62) = 1.96$, $p = 0.15$.

Thus, the reduction of the magnitude from within- to between-languages was 72%, which is very similar to what has been reported in the literature (MacLeod, 1991; Francis, 1999). Furthermore, the percentage is identical to the different script languages of the CEM trilinguals in Experiment 2 (72% for Chinese–English).

STROOP FACILITATION (CONTROL CHARACTER VS. CONGRUENT CONDITION)

English input lead to significant Stroop facilitation across all output languages, all ps < 0.05 (see **Table 3**), although only the English input and output combination yielded a positive effect (36 ms), while the others produced a negative effect (−40 and −44 ms). Significant negative Stroop facilitation was also found for Chinese input and Uyghur output (−32 ms), $t(31) = 3.17$, $p < 0.05$, $\eta^2 = 0.24$.

The 3 (input language) × 3 (output language) repeated-measures ANOVA revealed an interaction between input and output languages for Stroop facilitation effects, $F(4,124) = 21.01$, $p < 0.0001$, $\eta^2 = 0.16$. Uyghur output showed an effect of input language, $F(2,62) = 13.31$, $p < 0.0001$, $\eta^2 = 0.13$, with significant differences between English and Uyghur input (−40 ms, SE = 10 vs. 6 ms, SE = 7 ms), $t(31) = 4.71$, $p < 0.001$, $\eta^2 = 0.42$, and between Chinese and Uyghur (−32 ms, SE = 10 vs. 6 ms, SE = 7 ms), $t(31) = 4.09$, $p < 0.001$, $\eta^2 = 0.35$. An effect of language input was also found with Chinese output, $F(2,62) = 16.96$, $p < 0.001$, $\eta^2 = 0.18$. Significant differences in the magnitude of Stroop facilitation were found between the Chinese and English input languages (26 ms, SE = 12 vs. −44 ms, SE = 13 ms), $t(31) = 5.64$, $p < 0.0001$, $\eta^2 = 0.51$, and the English and Uyghur input (−44 ms, SE = 13 vs. 5 ms, SE = 8 ms), $t(31) = 3.64$, $p < 0.001$, $\eta^2 = 0.30$. Also English output revealed an effect of language input, $F(2,62) = 13.60$, $p < 0.0001$, $\eta^2 = 0.18$, with a significant difference between Chinese and English input (−4 ms, SE = 11 vs. 36 ms, SE = 8 ms), $t(31) = 3.99$, $p < 0.01$, $\eta^2 = 0.34$ and between Uyghur and English (−29 ms, SE = 12 vs. 36 ms. SE = 8 ms), $t(31) = 4.88$, $p < 0.0001$, $\eta^2 = 0.43$.

The 2 (type: within or between) × 3 (language output) repeated-measures ANOVA showed a main effect of type, $F(1,31) = 59.12$, $p < 0.0001$, $\eta^2 = 0.12$, such that there was a larger effect for within (22 ms, SE = 10 ms) than between language conflict (−24 ms, SE = 11 ms), and a trend toward a main effect of output, $F(2,62) = 3.08$, $p = 0.053$, $\eta^2 = 0.02$, but no interaction, $F(2,62) < 1$. Interestingly, the magnitude of within-language Stroop facilitation was positive (22 ms), while between-language Stroop facilitation was negative (−24 ms). The size of this negative Stroop facilitation was larger than for the CEM trilinguals reported in Experiment 2 (−9 vs. −24 ms). Thus, script appears to modulate the Stroop facilitation effect. This will be analyzed further in the next section and taken up in the Section "General Discussion."

ANALYSES ACROSS EXPERIMENTS 1–3

To compare the results across the three groups of trilinguals, we analyzed first the magnitude of within- and between-language Stroop interference and facilitation effects. Next, we looked across the three groups of trilinguals at the amount of within- and between-language Stroop interference and facilitation in terms of alphabetic (German, English, Dutch, Malay), Chinese, and Arabic (Uyghur) scripts to investigate the role of script.

STROOP INTERFERENCE: WITHIN- AND BETWEEN-LANGUAGE

Figure 1A summarizes the amount of within- and between-language Stroop interference when data was collapsed across input and output languages. The analysis of the data of the three groups of trilinguals revealed a similar amount of within-language Stroop interference, $F(2,255) < 1$ (GED: 73 ms, SE = 10 ms; CEM: 64 ms, SE = 12 ms; UCE: 75 ms, SE = 14 ms). In contrast, the between-language interference varied between the three trilingual groups, $F(2,513) = 7.58$, $p < 0.001$, $\eta^2 = 0.03$ (see **Figure 1B**). As reported in Experiment 1, the amount of within- and between-language interference was similar for the GED trilinguals, whereas it was reduced for the CEM (73% of the within-language interference, see Experiment 2) and UCE trilinguals (72% of the within-language interference, see Experiment 3). After discussing the Stroop facilitation we will analyze the Stroop interference in terms of script similarity to investigate whether script similarity can explain the reduction.

STROOP FACILITATION: WITHIN- AND BETWEEN-LANGUAGE

A three-way (trilingual group) ANOVA revealed that the magnitude of within-language Stroop facilitation was similar across the three groups of trilinguals, $F(2,255) < 1$ (GED: 25 ms, SE = 7 ms; CEM: 27 ms, SE = 9 ms; UCE: 22 ms, SE = 10 ms), whereas the magnitude of between-language facilitation was modulated by type of trilingual, $F(2,513) = 12.78$, $p < 0.0001$, $\eta^2 = 0.05$. No between-language Stroop facilitation was observed for the GED trilinguals (4 ms, SE = 8 ms), whereas responses to the Stroop

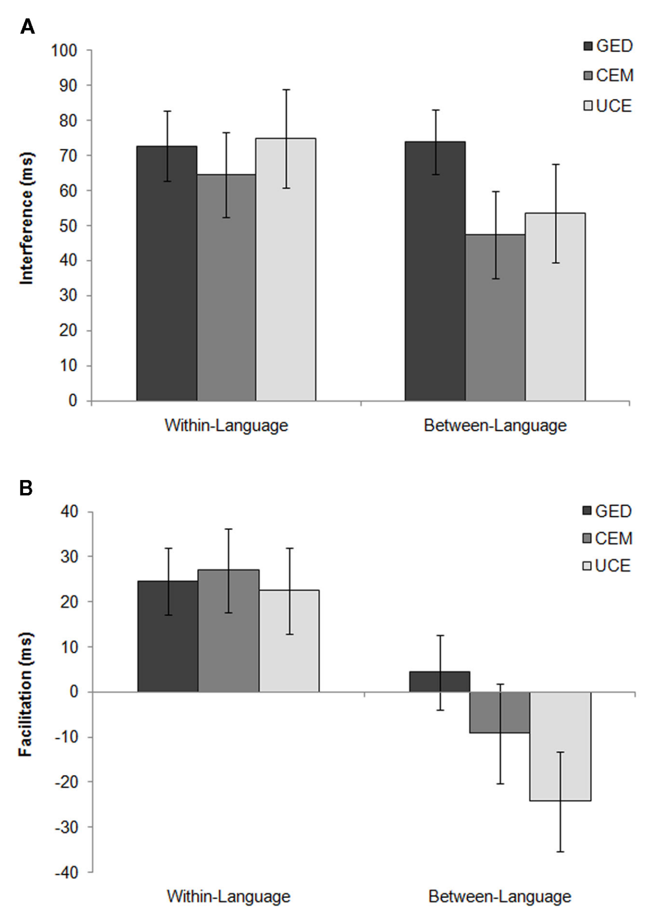


FIGURE 1 | Magnitude of (A) Stroop interference and (B) Stroop facilitation within- and between-languages for the German–English–Dutch (GED) trilinguals in Experiment 1, the Chinese–English–Malay (CEM) trilinguals in Experiment 2, and the Uyghur–Chinese–English (UCE) trilinguals in Experiment 3.

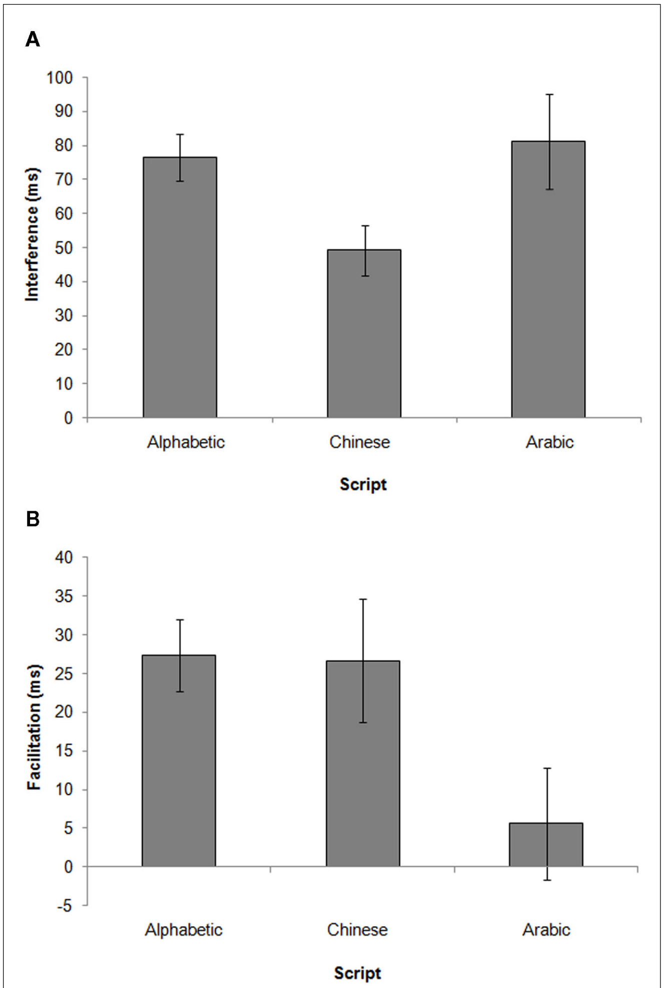


FIGURE 2 | Within-language effects broken down by script across all three trilingual groups for comparisons of (A) Stroop interference and (B) Stroop facilitation effects.

congruent condition were slower than to the control condition (negative Stroop facilitation) for the CEM (−9 ms, SE = 11 ms) and UCE trilinguals (−24 ms, SE = 11 ms). For CEM and UCE trilinguals in the congruent condition, the written word and color information matched at the conceptual level (converging conceptual information), but at the word form level a mismatch occurred between the spoken and written word (divergent phonological information) that resulted into a negative Stroop facilitation effect (cf. Roelofs, 2010). In the next section, we examine further the influence of script on within- and between-language Stroop interference and facilitation by analyzing the data across the three groups of trilinguals. This is particularly relevant in the case of the CEM trilinguals, as two of their languages share the same-script (alphabetic).

THE ROLE OF SCRIPT SIMILARITY ACROSS LANGUAGES

We explored the influence of script similarity on Stroop interference and facilitation across the languages of the three groups of trilinguals. First, the magnitude of within-language Stroop interference and facilitation across the three groups of trilinguals for alphabetic (German, English, Dutch, Malay), logographic (Chinese), and Arabic (Uyghur) scripts was calculated (see **Figure 2**). To compare the between-language effects, the cross-language similarity of each combination of languages was coded as either same or different script (same = alphabetic scripts; different = combination of alphabetic, Chinese, and Arabic scripts). Next, the between-language Stroop effects were calculated (see **Figure 3**). We will report first the analyses of the within- and between-language Stroop interference and then the analyses of the within- and between-language Stroop facilitation.

STROOP INTERFERENCE: WITHIN- AND BETWEEN-LANGUAGE

A three-way (script) ANOVA for within-language Stroop interference showed an effect of script, $F(2,255) = 4.16$, $p < 0.05$, $\eta^2 = 0.03$. Independent-sample t-tests identified a difference in the magnitude of within-language interference between the alphabetic (77 ms, SE = 7 ms) and Chinese (49 ms, SE = 6 ms) scripts, $t(108.6) = 3.05$, $p < 0.01$, $\eta^2 = 0.08$, and a trend between Chinese (49 ms, SE = 6 ms) and Arabic (81 ms, SE = 8 ms), $t\ (48.8) = 2.02$,

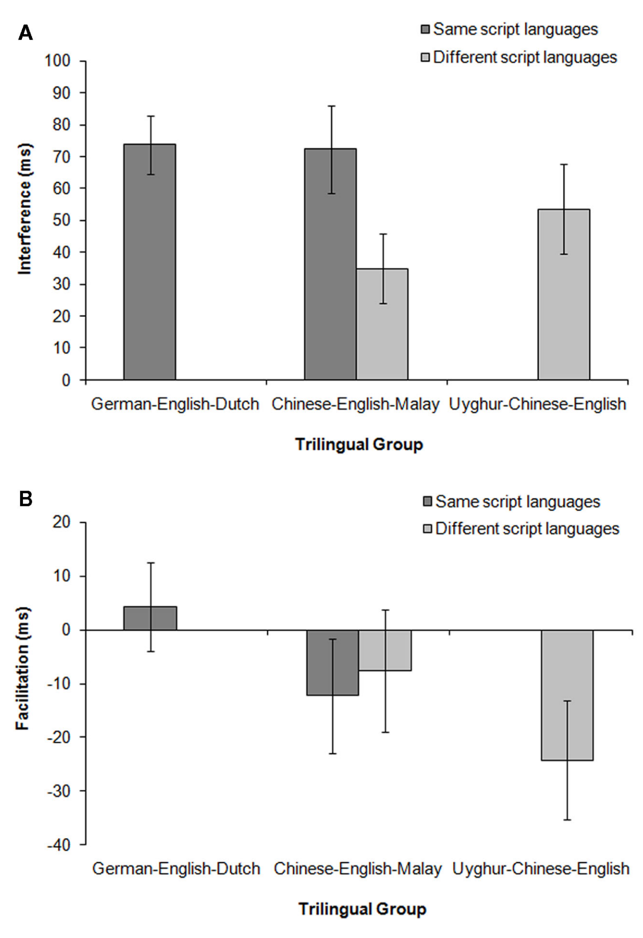


FIGURE 3 | Between-language effects broken down by script similarity across the three groups of trilinguals for comparisons of (A) Stroop interference and (B) Stroop facilitation effects.

$p < 0.05$, $\eta^2 = 0.08$, but not between alphabetic (77 ms, SE = 7 ms) and Arabic (81 ms, SE = 8 ms), t (39.3) = 0.32, $p = 0.75$ (see **Figure 2A**). This reduction of the magnitude of Stroop interference for the Chinese script relative to the other scripts (64% of the Alphabetic and 60% of the Arabic script) is interesting. However, it is unclear why this reduction occurred because in the literature either stronger Stroop interference effects have been reported in Chinese than English (Biederman and Tsao, 1979) or an equal amount of interference (Smith and Kirsner, 1982; Lee and Chan, 2000).

A 2 (script similarity: same or different) × 3 (trilingual group) ANOVA for between-language Stroop interference revealed a significant main effect of similarity, $F(1,512) = 20.79$, $p < 0.0001$, $\eta^2 = 0.04$ (see **Figure 3A**), such that the magnitude of interference of same-script languages was larger than for different script languages (74 ms, SE = 6 vs. 47 ms, SE = 8 ms) and there was a trend toward a main effect of trilingual group, $F(1,512) = 2.68$, $p = 0.07$, $\eta^2 = 0.01$. This trend of trilingual group arose from a significant difference in the different script effects of the CEM (35 ms, SE = 11 ms) and UCE groups (54 ms, SE = 14 ms), $t(263.7) = 2.36$, $p < 0.05$, $\eta^2 = 0.02$. Importantly, as discussed in the Section "Results and Discussion" of Experiment 2, there was no difference between GED same-script (German, English, Dutch: 74 ms, SE = 9 ms) and the CEM same-script languages (Malay and English: 72 ms, SE = 14 ms), $p = 0.88$ uncorrected, even though the color words translations were similar in terms of orthography and phonology for GED trilinguals and different for CEM same-script languages (Malay and English). The modulation of between-language Stroop interference by script will be discussed further in the Section "General Discussion."

STROOP FACILITATION: WITHIN- AND BETWEEN-LANGUAGE

For within-language Stroop facilitation (see **Figure 2B**) there was a trend toward an effect of script in the Stroop facilitation comparison, $F(2,252) = 2.95$, $p = 0.05$, $\eta^2 = 0.02$. Independent-sample t-tests showed a significant difference in within-language facilitation between alphabetic (27 ms, SE = 5 ms) and Arabic (6 ms, SE = 4 ms) scripts, $t(45.1) = 2.73$, $p < 0.05$, $\eta^2 = 0.14$, and a trend toward a difference between Chinese (27 ms, SE = 6 ms) and Arabic (6 ms, SE = 4 ms) scripts, $t(82.9) = 1.96$, $p = 0.053$, $\eta^2 = 0.04$, but no difference between alphabetic (27 ms, SE = 5 ms) and Chinese scripts (27 ms, SE = 6 ms), $t(75.2) = 0.08$, $p = 0.94$.

The analysis of the magnitude of between-language Stroop facilitation revealed a significant main effect of script similarity, $F(1,512) = 16.30$, $p < 0.0001$, $\eta^2 = 0.03$, because there was no Stroop facilitation effect for same-script (alphabetic) languages, whereas for different script languages the effect of between-language Stroop congruency was negative (1 ms, SE = 5 vs. −19 ms, SE = 7 ms). There was also a main effect of trilingual group, $F(2,512) = 4.73$, $p < 0.01$, $\eta^2 = 0.02$. *Post hoc* t-tests indicated that there was a trend toward a significant difference in the same-script facilitation effects between the GED and CEM groups, $t(66.9) = 2.02$, $p = 0.095$ ($p = 0.048$ uncorrected), $\eta^2 = 0.06$, such that the GED effect was more positive than the CEM effect (4 ms, SE = 8 vs. −12 ms, SE = 11 ms), and also a significant difference between the different script facilitation between CEM (−8 ms, SE = 11 ms) and UCE (−24 ms, SE = 11 ms), $t(212.7) = 2.31$, $p < 0.05$, $\eta^2 = 0.02$. The overall pattern indicates that between-language Stroop facilitation was primarily modulated by script similarity, such that increased language dissimilarity (in term of script) lead to slower responses to the between-language congruent condition relative to the control condition. However, the results also seem to indicate that cross-language similarity between the color word translations in terms of orthography/phonology seems to play a role as well (between-language facilitation difference between GED and CEM same-script is a trend), although it is clearly not as strong as the impact of script similarity. Overall, the findings could be explained by a combination of a conceptual match (converging information) at the output (e.g., blue in English) and input (blue colored rectangle and the translation of *blue* written in a different script language in the center of the rectangle) but a mismatch at the level of script (orthography) and phonology (diverging information). In the next Section "General Discussion," we will discuss this and other explanations in more detail.

GENERAL DISCUSSION

Our Stroop experiments involved three groups of trilinguals with languages that are highly similar (GED trilinguals), partly

(CEM trilinguals), or completely different (UCE trilinguals). A comparison of these trilinguals made it possible to analyze the influence of cross-language similarity on Stroop interference and facilitation in terms of orthographic/phonological overlap between the color word translations, and script. We will first consider the within- and between-language Stroop interference and facilitation effects and then focus on the WLSSE. Finally, we will discuss the implications for theories of language processing and control in bilinguals and multilinguals.

STROOP INTERFERENCE: WITHIN- AND BETWEEN-LANGUAGE

The observed magnitude of between-language Stroop interference was equal to within-language for GED trilinguals, but was reduced for CEM and UCE trilinguals (respectively 72 and 73% of the within-language interference). The size of the reduction is in line with the conclusions of Francis (1999) and MacLeod (1991). Interestingly, for CEM trilinguals, between-language interference was modulated by script similarity. For similar script (alphabetic) languages, the between-language interference effect was similar to the within-language interference effect, whereas for different script combinations between-language interference was reduced considerably. Furthermore, the size of within- and between-language interference for same-script languages (Malay and English) of CEM trilinguals was similar to that of GED trilinguals. This is theoretically important because the color word translations of the GED trilinguals are similar in terms of orthography and phonology (e.g., *rot*, *red*, and *rood*), whereas the color word translations of the same-script languages Malay and English are different in terms of their spelling and pronunciation (e.g., *merah* and *red*). Thus, unlike script similarity, the similarity between the color word translations in terms of orthography and phonology does not seem to modulate the amount of between-language Stroop interference.

Several studies in the literature have concluded that cross-language similarity modulates between-language interference (e.g., Fang et al., 1981; Brauer, 1998). However, in some studies the orthographic/phonological similarity of the color word translations was confounded with script similarity. Preston and Lambert (1969) investigated the role of color word translation similarity on Stroop interference. As discussed in the Introduction, they found that between-language Stroop interference was 68% of within-language interference in English–Hungarian bilinguals, whereas it was 95% for English–French bilinguals. In their second experiment, they manipulated cross-language similarity within German–English bilinguals using two sets of color terms (translations similar or not in terms of orthography/phonology). The results revealed an equal amount of within- and between-language interference for similar color word translations and a reduction of between-language interference for dissimilar color word translations (between-language interference was 62% of within-language interference). Thus, in contrast to our results with trilinguals, their results with bilinguals suggest that between-language interference is modulated by the orthographic/phonological similarity of the color word translations. However, note that the experimental method used by Preston and Lambert and by many others in the literature (e.g., Dyer, 1971; Fang et al., 1981; Brauer, 1998) involved cards with multiple items of the same condition (i.e., cards with 10 rows of 10 color words written in incongruent ink colors). Thus, input conditions were blocked, whereas in the current study all conditions were completely randomized. Therefore, the influence of cross-language similarity in terms of the orthography/phonology of the color word translations on between-language Stroop interference might be restricted to the specific design of Preston and Lambert.

STROOP FACILITATION: WITHIN- AND BETWEEN-LANGUAGE

The magnitude of within-language Stroop facilitation collapsed across output languages did not differ between the three groups of trilinguals. For all trilinguals, color naming latencies were faster relative to the control condition when the naming response matched the pronunciation of the written word (e.g., for English output: blue colored rectangle with the word "blue" presented in the center). In contrast, between-language Stroop facilitation was modulated by cross-language similarity. For trilinguals with languages written in the same-script (GED), responses to the congruent Stroop condition did not differ from the control condition. However, when the languages of the trilinguals were written using different scripts, the responses to the congruent Stroop condition were slower than to the control condition (negative Stroop facilitation effect). For example, when the response was in Uyghur and the stimulus consisted of a blue rectangle with the English word "blue" presented at the center, naming latencies were slower than for the control condition. Similar negative Stroop facilitation effects for between-language Stroop facilitation have been reported before with bilinguals (Dalrymple-Alford, 1968; MacLeod and MacDonald, 2000) and trilinguals (Abunuwara, 1992).

Thus, similar to between-language Stroop interference, the between-language Stroop facilitation effects are modulated by script. However, in contrast to between-language Stroop interference, orthographic and phonological similarity of the color word translations seem to have some impact on the Stroop facilitation effects in addition to script differences: for the GED trilinguals the Stroop facilitation was positive, while for the CEM and UCE trilinguals the effect was negative. In particular, for GED trilinguals and the same-script languages of the CEM trilinguals (Malay and English) the magnitude of between-language facilitation differed.

The occurrence of negative between-language Stroop facilitation effects can be explained by the *inadvertent reading hypothesis* (MacLeod and MacDonald, 2000) in terms of a covert repair processes when the congruent word is read in the incorrect language. These covert repair processes are assumed to slow down the naming process, so that naming latencies in the congruent condition are equal or slower than those in the control condition (Roelofs, 2010). However, it is unclear how the inadvertent reading hypothesis explains the modulation of between-language Stroop facilitation effects by cross-language script similarity. Furthermore, the inadvertent reading hypothesis does not explain the (positive) between-language Stroop facilitation effect observed with GED trilinguals in our study and with Dutch–English bilinguals in a Stroop task with preexposure SOAs by Roelofs (2010).

In contrast, the *converging information hypothesis* (Cohen et al., 1990; Melara and Algom, 2003; Roelofs, 2003, 2010) can account for negative between-language Stroop facilitation, as well as for the influence of cross-language similarity on between-language facilitation. This hypothesis proposes that a combination of converging information at the conceptual and phonological levels underlies between-language facilitation. Thus, the interference in the Stroop congruent condition is explained by assuming that two different phonological forms are activated. For instance, the competing phonological form of the English color word *red* and that of the translation in Chinese (红) may slow the naming response, in spite of a match at the conceptual level (Roelofs, 2010). A modulation of between-language Stroop facilitation effects by cross-language similarity can be accounted for by this hypothesis, because the converging information at the form level is influenced by the similarity of the representations at this level.

Roelofs (2010) concluded that "...Stroop facilitation and interference have a common locus within and between-languages, supporting the converging information hypothesis of Stroop facilitation." However in a recent paper, Brown (2011) showed that there is no (or very weak and inversely) correlation between Stroop interference and facilitation after correction for spurious correlations, which argues against a common locus of both effects. As Brown argued, response conflict and resolution processes are uniquely involved in Stroop interference and therefore some processes may be shared between Stroop interference and facilitation, whereas others may be unique to Stroop interference or facilitation.

Overall, our data revealed stronger Stroop interference than Stroop facilitation in all trilinguals, a finding commonly reported in the literature (MacLeod and MacDonald, 2000). However, as Brown (2011) recently pointed out, when using X's as control condition to calculate Stroop interference and facilitation, both effects are confounded by a lexicality cost, which is supported by the finding that neutral words in fact interfer with color naming (e.g., Brown et al., 1998). Thus, the Stroop facilitation effects are underestimated, whereas Stroop interference effects are exaggerated (see Brown, 2011, p. 87). As a consequence, negative Stroop facilitation might arise due to a large lexicality cost. Furthermore, the lexicality cost could vary between-languages/scripts and potentially also depend on the familiarity/proficiency with the scripts at hand. We observed a strong trend of an effect of script in the within-language Stroop facilitation of the UCE trilinguals, such that Stroop facilitation was stronger in English than Chinese and Uyghur (36 vs. 26 vs. 9 ms). This trend could reflect a larger lexicality cost in Uyghur than in Chinese and English. At the same time, the lexicality cost would affect between-language Stroop facilitation and result in the observed negative between-language Stroop facilitation.

WITHIN-LANGUAGE STROOP SUPERIORITY EFFECT

The finding of larger within- than between-language Stroop effects has been referred to as the WLSSE. We calculated the WLSSE by collapsing the data over all within and between conditions for each language output and then calculating the difference of the within-effects minus the between-effects. **Figure 4** provides an overview of the WLSSE for Stroop interference and facilitation across all output languages of the trilinguals.

As expected, based on the analyses across the experiments the magnitude of the WLSSE in all the Stroop interference and Stroop facilitation comparisons increases with increasing language differences (in terms of script) in the trilinguals. As discussed in the Introduction, Goldfarb and Tzelgov (2007) argued that the WLSSE is driven by response set effects (see also Roelofs, 2003). According to this account, within-language Stroop interference is larger than between-language due to competition at the conceptual and lexical levels in the within-language condition, whereas in the between-language condition there is only competition at the conceptual level. There is no lexical competition in the between-language condition, because the lexical items (color words in the non-target output language) do not belong to the response set. The present results with trilinguals show that WLSSE is modulated by cross-language similarity; this cannot be explained if the WLSSE is simply due to differences in response sets given that these are the same for the three groups of trilinguals.

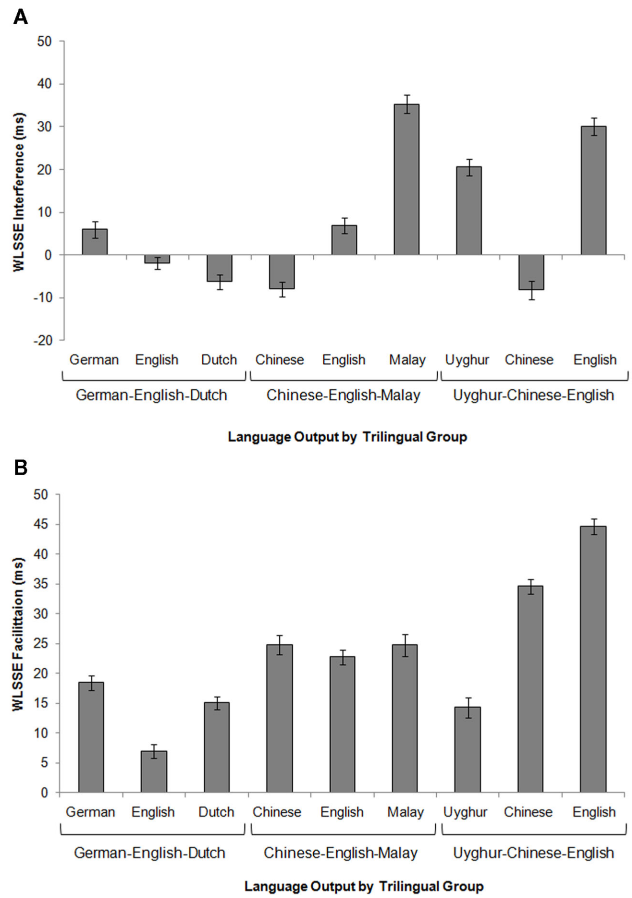


FIGURE 4 | Within-language Stroop superiority effects (WLSSE) for (A) Stroop interference and (B) Stroop facilitation effects in each trilingual group and output language.

IMPLICATIONS

Overall, our results indicate that Stroop effects in trilinguals are comparable to those of bilinguals. This is evidence that the bilingual and trilingual language systems are comparable in their organization. This conclusion is consistent with Dijkstra's (2003) theoretical analysis of language processing in bilinguals and multilinguals. An important issue is whether or not bilinguals and multilinguals can use script cues to reduce cross-language interference in the Stroop task. Our trilingual data showed that script differences between input and output languages reduced between-language Stroop interference. One potential interpretation of this finding is that trilinguals who are proficient in languages involving different scripts can use script type as a cue to reduce the impact of cross-language interference. However, this interpretation does not explain the negative between-language Stroop facilitation observed in CEM and UCE trilinguals. If script cues could be used optimally to reduce cross-language influences, between-language Stroop facilitation should be absent in different script trilinguals. Although the processes underlying Stroop interference and facilitation might be different (Tzelgov et al., 1990, 1992, 1996; Brown, 2011; see also discussion above), the different impact of script cues on Stroop interference and facilitation could be explained in terms of the level in the language processing system at which they affect processing. Script cues might affect the word production process at higher levels (thus reducing between-language Stroop interference), but might not be able to reduce interference at the lower phonological output level (thus interference occurs for between-language Stroop facilitation).

In addition to cross-language similarity, proficiency has also been shown to modulate interference effects (e.g., Preston and Lambert, 1969; Fang et al., 1981; Mägiste, 1984; Chen and Ho, 1986; Brauer, 1998; Sumiya and Healy, 2004, 2008). To date the role of proficiency and how it interacts with cross-language similarity in Stroop interference and facilitation has not been systematically investigated with trilinguals. Because the primary goal of the current study was to focus on the role of cross-language similarity in terms of script, the participants were matched as closely as possible with regards to proficiency. However, the trilinguals in the current study were not balanced across all three languages. Their L1 was clearly their native language, and their L2 and L3 were less proficient than their L1. As a consequence, the unbalanced proficiency levels across trilinguals' three languages may have modulated the observed Stroop interference and facilitation effects. Investigation of the impact of relative proficiency across languages requires further research with trilinguals having a wider variation in their language proficiency. However, taking into consideration previous research on proficiency, we suggest that the findings of the present study are mainly modulated by language similarity.

CONCLUSION

To conclude, the present study provides evidence that cross-language similarity across color words modulates both between-language Stroop interference and facilitation in trilinguals. In particular, cross-language similarity in terms of the scripts in which the languages of the trilinguals are written modulates between-language Stroop effects. Furthermore, cross-language similarity in terms of the degree of orthographic and phonological overlap between the color word translations seems to have some impact on the Stroop facilitation effects in addition to script similarity. Overall, the observed patterns of Stroop interference and facilitation can be explained by diverging and converging color and word information across languages.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grant Number 31170970) and the Fundamental Research Funds for the Central Universities to Taomei Guo. We would like to thank Helen Lui, Lonneke Bucken, Jingjing Guo, Shangjie Lin, and Zhimin Zeng for their help with the data collection.

REFERENCES

Abunuwara, E. (1992). The structure of the trilingual lexicon. *Eur. J. Cogn. Psychol.* 4, 311–322.

Biederman, I., and Tsao, Y.-C. (1979). On processing Chinese ideographs and English words: some implications from Stroop-test results. *Cogn. Psychol.* 11, 125–132.

Brauer, M. (1998). "Stroop interference in bilinguals: the role of script similarity between two languages," in *Foreign Language Learning: Psycholinguistic Studies on Training and Retention,* eds A. F. Healy and L. E. Jr. Bourne (Mahwah, NJ: Erlbaum), 317–337.

Brown, T. L. (2011). The relationship between Stroop interference and facilitation effects: statistical artifacts, baselines, and a reassessment. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 85–99.

Brown, T. L., Gore, C. L., and Pearson, T. (1998). Visual half-field Stroop effects with spatial separation of words and color targets. *Brain. Lang.* 63, 122–142.

Chen, H.-C., and Ho, C. (1986). Development of Stroop interference in Chinese-English bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 12, 397–401.

Cohen, J., Dunbar, K., and McClelland, J. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97, 332–361.

Colomé, A. (2001). Lexical activation in Bilinguals' speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736.

Costa, A., Albareda, B., and Santesteban, M. (2008). Assessing the presence of lexical competition across languages: evidence from the Stroop task. *Biling. (Camb. Engl.)* 11, 121–131.

Costa, A., Caramazza, A., and Sebastian Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1283–1296.

Dalrymple-Alford, E. C. (1968). Interlingual interference in a color- naming task. *Psychon. Sci.* 10, 215–216.

de Groot, A. M. B., Delmaar, P., and Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: support for nonselective access to bilingual memory. *Q. J. Exp. Psychol. A* 53, 397–428.

Dijkstra, T. (2003). "Lexical processing in bilinguals and multilinguals: the word selection problem," in *The Multilingual Lexicon,* eds J. Cenoz, B. Hufeisen, and U. Jessner (Dordrecht: Kluwer Academic Publishers), 11–26.

Dijkstra, T., Grainger, J., and van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: the neglected role of phonology. *J. Mem. Lang.* 41, 496–518.

Dijkstra, T., and van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. (Camb. Engl.)* 5, 175–197.

Dijkstra, T., van Jaarsveld, H., and ten Brinke, S. (1998). Interlingual homograph recognition: effects of task demands and language intermixing. *Biling. (Camb. Engl.)* 1, 51–66.

Dyer, F. N. (1971). Color-naming interference in monolinguals and bilinguals. *J. Verbal Learn. Verbal Behav.* 10, 297–302.

Fang, S.-P., Tzeng, O. J. L., and Alva, L. (1981). Intralingual vs. interlingual Stroop effects in two types of writing systems. *Mem. Cognit.* 9, 609–617.

Forster, K. I., and Forster, J. C. (2003). DMDX: a Windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124.

Francis, W. S. (1999). Cognitive integration of language and memory in bilingual memory: semantic representations. *Psychol. Bull.* 125, 193–122.

Goldfarb, L., and Tzelgov, J. (2007). The cause of the within-language Stroop superiority effect and its implications. *Q. J. Exp. Psychol.* 60, 179–185.

Green, D. (1998). Mental control of the bilingual lexico-semantic system. *Biling. (Camb. Engl.)* 1, 67–81.

Guo, T., and Peng, D. (2006). ERP evidence for parallel activation of two languages in bilingual speech production. *Neuroreport* 17, 1757–1760.

Hermans, D., Bongaerts, T., de Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language? *Biling. (Camb. Engl.)* 1, 213–229.

Hoshino, N., and Kroll, J. F. (2008). Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106, 501–511.

Jared, D., and Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *J. Mem. Lang.* 44, 2–31.

Kane, M. J., and Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J. Exp. Psychol. Gen.* 132, 47–70.

Lee, T. M. C., and Chan, C. C. H. (2000). Stroop interference in Chinese and English. *J. Clin. Exp. Neuropsychol.* 22, 465–471.

Lee, W. L., Wee, G. C., Tzeng, O. J. L., and Hung, D. L. (1992). "A study of interlingual and intralingual Stroop effect in three different scripts: logographic, syllabary, and alphabet," in *Cognitive Processing in Bilinguals,* ed. R. J. Harris (Amsterdam: Elsevier Science Publishers), 427–442.

Lemhöfer, K., and Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. *Mem. Cognit.* 32, 533–550.

Lemhöfer, K., Dijkstra, T., and Michel, M. C. (2004). Three languages, one ECHO: cognate effects in trilingual word recognition. *Lang. Cogn. Process.* 19, 585–611.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203.

MacLeod, C. M., and MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends Cogn. Sci. (Regul. Ed.)* 4, 383–391.

Mägiste, E. (1984). Stroop task and dichotic translation: the development of interference patterns in bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 10, 304–315.

Melara, R. D., and Algom, D. (2003). Driven by information: a tectonic theory of Stroop effects. *Psychol. Rev.* 110, 422–471.

Preston, M. S., and Lambert, W. E. (1969). Interlingual interference in a bilingual version of the Stroop color-word task. *J. Verbal Learn. Verbal Behav.* 8, 295–301.

Protopapas, A. (2007). CheckVocal: a program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behav. Res. Methods* 39, 859–862.

Roelofs, A. (2003). Goal-referenced selection of verbal action: modeling attentional control in the Stroop task. *Psychol. Rev.* 110, 88–125.

Roelofs, A. (2010). Attention and facilitation: converging information versus inadvertent reading in Stroop task performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 411–422.

Smith, M. C., and Kirsner, K. (1982). Language and orthography as irrelevant features in colour-word and picture-word Stroop interference. *Q. J. Exp. Psychol. A* 34, 153–170.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Sumiya, H., and Healy, A. F. (2004). Phonology in the bilingual Stroop effect. *Mem. Cognit.* 32, 752–758.

Sumiya, H., and Healy, A. F. (2008). The Stroop effect in English-Japanese Bilinguals: the effect of phonological similarity. *Exp. Psychol.* 55, 93–101.

Tzelgov, J., Henik, A., and Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Mem. Cognit.* 20, 727–635.

Tzelgov, J., Henik, A., and Leiser, D. (1990). Controlling Stroop interference: evidence from a bilingual task. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 760–771.

Tzelgov, J., Henik, A., Sneg, R., and Baruch, O. (1996). Unintentional word reading via the phonological route: the Stroop effect with cross-script homophones. *J. Exp. Psychol. Learn. Mem. Cogn.* 22, 336–349.

van Hell, J. G., and de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: effects of concreteness and cognate status in word association. *Biling. (Camb. Engl.)* 1, 193–211.

van Hell, J. G., and Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychon. Bull. Rev.* 9, 780–789.

van Heuven, W. J. B., Dijkstra, A., and Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *J. Mem. Lang.* 39, 458–483.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 May 2011; accepted: 27 November 2011; published online: 13 December 2011.

*Citation: van Heuven WJB, Conklin K, Coderre EL, Guo T and Dijkstra T (2011) The influence of cross-language similarity on within- and between-language Stroop effects in trilinguals. Front. Psychology **2**:374. doi: 10.3389/fpsyg.2011.00374*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 02 February 2012
doi: 10.3389/fpsyg.2012.00009

Do Spanish–English bilinguals have their fingers in two pies – or is it their toes? An electrophysiological investigation of semantic access in bilinguals

***Noriko Hoshino*[1]* *and Guillaume Thierry*[2]

[1] *Department of English, Kobe City University of Foreign Studies, Kobe, Japan*
[2] *School of Psychology, Economic and Social Research Council Centre for Research on Bilingualism in Theory and Practice, Bangor University, Bangor, UK*

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
Natasha Tokowicz, University of Pittsburgh, USA
Pedro Macizo, University of Granada, Spain
Patricia E. Román, Max Planck Society, Germany

****Correspondence:***
Noriko Hoshino, Department of English, Kobe City University of Foreign Studies, 9-1 Gakuen Higashi-machi, Nishi-ku, Kobe, Hyogo-ken 651-2187, Japan.
e-mail: noriko.hoshino@inst.kobe-cufs.ac.jp

We examined the time course of cross-language activation during word recognition in the context of semantic priming with interlingual homographs. Spanish–English bilinguals were presented pairs of English words visually one word at a time and judged whether the two words were related in meaning while recording event-related potentials. Interlingual homographs (e.g., "pie": "Pie" in Spanish is a foot.) appeared in the target position and were preceded by primes that were either related to the English meaning (e.g., "apple"), related to the Spanish meaning of interlingual homographs (e.g., "toe") or totally unrelated (e.g., "floor"/"bed"). Spanish–English bilinguals showed semantic priming not only when interlingual homographs were related to the English meaning but also to the Spanish meaning of the prime. These priming effects were detectable in the mean amplitude of the N400 (350–500 ms) even when the target word was related to the prime in Spanish and the context of the experiment was English. However, the relatedness effect was found in the window of a late positive component (LPC; 550–700 ms) only for stimulus pairs related in English. To verify that the observed pattern of the results was due to participants' bilingualism, we also tested a group of English monolinguals. The monolinguals showed a semantic priming effect for the N400 and LPC time windows only when interlingual homographs were related to the English meaning. These results suggest that both languages are activated in the classical time frame of semantic activation indexed by N400 modulations, but that semantic activation in the non-target language failed to be explicitly processed.

Keywords: interlingual homographs, semantic priming, ERPs

INTRODUCTION

Previous research suggests that even when bilinguals read in one language (the target language), the other language (the non-target language) is also active and influences the process of reading in the target language to some extent (e.g., Dijkstra et al., 1999; Jared and Kroll, 2001). Some evidence for non-selective language activation comes from studies with interlingual homographs. Interlingual homographs are words whose spellings are similar but whose meanings are different across two languages (e.g., "pie" in English is food and "pie" in Spanish is a foot). Past studies have shown that the meaning of interlingual homographs in a non-target language is also activated (e.g., Beauvillain and Grainger, 1987; Dijkstra et al., 1998, 1999; Elston-Güttler, 2000; De Bruijn et al., 2001; Lemhöfer and Dijkstra, 2004; Kerkhofs et al., 2006; Paulmann et al., 2006; Macizo et al., 2010; Martín et al., 2010). However, cognitive control mechanisms effectively allow the meaning in a target language to be selected at some point in processing.

In a seminal study with interlingual homographs by Dijkstra et al. (1998), Dutch–English bilinguals performed a lexical decision task. Dijkstra and his colleagues found that Dutch–English bilinguals were slower responding to interlingual homographs (e.g., "boom" in English is a sudden increase and "boom" in Dutch is a tree) than to English control words (e.g., "riot") when some Dutch words were included in the task. This inhibitory effect was also modulated by the relative frequency of interlingual homograph meaning across languages. A greater inhibition was observed when the frequency of the meaning of the homograph in Dutch was high than when it was low. These results suggest that not only the reading of interlingual homographs in one language but also its reading in the other language is active and participates in the competition for selection.

If the meaning of interlingual homographs in a non-target language is also activated, how do bilinguals select the meaning of words presented in a target language? Using a negative priming paradigm, Macizo et al. (2010) have recently shown that the non-target meaning of interlingual homographs is activated but inhibited subsequently. In their study, Spanish–English bilinguals made semantic judgments on pairs of English words. The bilinguals responded more slowly to pairs involving homographs (e.g., *pie–toe*; "pie" meaning "foot" in Spanish) than to pairs without homographs. Critically, participants were slower when the pair included the translation of the Spanish meaning of the homograph (e.g., *foot–hand*) after a trial involving an homograph (e.g., *pie–toe*), but not after a trial involving a control word (e.g., *log–toe*).

These results suggest that the Spanish (non-target) meaning of homographs is activated initially and inhibited subsequently. Furthermore, Martín et al. (2010) found this inhibitory effect when the subsequent trial was initiated 500 ms after responding to pairs involving homographs, but not 750 ms. This finding suggests that the reactive inhibition decays over time.

However, it remains undetermined when the activation of the non-target language is present. In the study by Macizo et al. (2010) and Martín et al. (2010), the inhibition of the non-target meaning was assessed during the presentation of a subsequent pair but online temporal monitoring of inhibition processes was not available. In the present study, we investigated the time window in which the non-target language is accessed in late Spanish–English bilinguals reading English. We used a classical semantic priming paradigm and we recorded event-related potential (ERPs). The study of Kerkhofs et al.'s (2006) investigating semantic priming with interlingual homographs in Dutch–English bilinguals is directly relevant here. Kerkhofs et al. (2006) showed that the primes that were semantically related to the English meaning of homographs facilitated the recognition of English homograph targets as indexed by N400 modulations. The N400 is thought to reflect lexical-semantic processing (e.g., Kutas and Hillyard, 1980 for semantics; Thierry and Wu, 2007 for form). Critically, Kerkhofs et al. (2006) found that the mean amplitude of the N400 was also modulated by the frequency of the non-target Dutch meaning of homographs, suggesting that the non-target representations of interlingual homographs also affect N400 modulation. Here, unlike Kerkhofs et al. (2006), we included primes that were semantically related to the non-target meaning of homographs as well as primes that were semantically related to the target meaning of homographs to assess the degree of non-target language activation, instead of manipulating the relative frequency of homograph meanings.

We asked Spanish–English bilinguals and English monolinguals to perform a semantic relatedness task on words presented in pairs within a go/no-go paradigm in which all the critical trials belonged to the no-go condition to prevent contamination of the ERPs by response-specific decision processes (e.g., Midgley et al., 2009; Thierry et al., 2009; Hoshino et al., 2010). This designed enabled us to compare the level of priming found for words related in L1 and words related in the non-target L2 indexed by N400 modulations. Furthermore, we examined the modulation of a late positive component (LPC), associated with explicit/controlled processing and stimulus re-evaluation (e.g., Thierry et al., 1998, 2003; Martin et al., 2009, 2010). If the non-target meaning of interlingual homographs continues to be active and/or is explicitly processed at a post-lexical processing stage, we should observe an LPC modulation in the non-target language as well as an N400 modulation. Specifically, we predicted that bilingual participants would access their first language (L1) during second language (L2) word reading whilst monolingual English participants would show no homograph priming effect (since they did not know Spanish). We were interested in measuring the magnitude of semantic priming in the window of the N400 and assessing the explicit or implicit nature of L1 activation. We expected priming to be weaker and more short-lived in the L1 Spanish than the L2 English, given that the L1 Spanish was not the target language. Therefore, we predicted that semantic priming in the non-target language Spanish would be measurable in the window of the N400, but might not result in modulation in the LPC range.

MATERIALS AND METHODS

PARTICIPANTS

Fourteen Spanish–English bilinguals and 14 English monolinguals participated in the study. All of the Spanish–English bilinguals spoke English as an L2 and were dominant in their L1. The characteristics of the two groups are provided in **Table 1**. Participant gave informed consent to take part in the study, which was approved by the ethics committee of Bangor University, Wales.

MATERIALS

Semantic judgment task

The critical stimuli consisted of 72 interlingual homographs (e.g., *pie*) and two sets of 72 non-cognates (e.g., *apple*). The mean length of the homographs was 5.1 letters (SD = 1.7) and the mean lexical frequency was 75.3 (SD = 144.4) based on the CELEX database (Baayen et al., 1995). The two sets of non-cognates were matched on length [$M = 5.4$, SD = 1.7; $M = 5.5$, SD = 1.7; $t(72) = -0.73$, $p > 0.10$] and lexical frequency [$M = 121.3$, SD = 227.5; $M = 76.5$, SD = 108.4; $t(72) = 1.61$, $p > 0.10$]. Four types of word pairs were formed where the first word was a non-cognate prime and the second word an interlingual homograph target. Primes and targets were either related (e.g., *apple–pie*) or unrelated (e.g., *rug–pie*) in English, and either related (e.g., *toe–pie*) or unrelated (e.g., *stove–pie*) when considering the target's meaning in Spanish. The Edinburgh Associative Norms were used to generate prime words. Because each target followed a prime in the semantic judgment task, the target was entered as a response, not as a stimulus in the Norms. The associative strength of related pairs for the English reading was 0.14, whereas that of related pairs for the Spanish

Table 1 | Characteristics of English monolinguals and Spanish–English bilinguals.

Measure	Monolinguals ($n = 14$)	Bilinguals ($n = 14$)
Age (years)	21.3 (5.4)	27.0 (5.5)
L1 self-rating (10 pt scale)	9.3 (0.8)	9.7 (0.5)
Reading	9.3 (0.9)	9.6 (0.6)
Writing	9.0 (1.1)	9.4 (0.9)
Speaking	9.3 (0.9)	9.8 (0.4)
Listening	9.4 (0.8)	9.9 (0.3)*
L2 self-rating (10 pt scale)	N/A	7.1 (1.7)
Reading	N/A	7.6 (1.5)
Writing	N/A	7.1 (1.9)
Speaking	N/A	6.8 (2.0)
Listening	N/A	7.1 (1.6)
Daily L1 usage (%)	N/A	44.3 (26.2)
Daily L2 usage (%)	N/A	55.7 (26.2)
Age of L2 acquisition (years)	N/A	9.7 (4.4)
Length of immersion (months)	N/A	43.7 (62.8)

*SDs are in parentheses. *$p < 0.05$.*

reading was 0.22 ($p > 0.10$). Therefore, if anything, the associative strength was greater for the Spanish reading. One set of non-cognates was used for related and unrelated primes in English and the other set for related and unrelated primes in Spanish. That is, the list of critical stimuli consisted of 72 English related trials, 72 English unrelated trials, 72 Spanish related trials, and 72 Spanish unrelated trials. It is important to note that the primes for English related were identical to those for English unrelated but appeared in the context of a different target (e.g., *rug–carpet, rug–pie*). Likewise, the other set of 72 non-cognates comprised the primes for Spanish related and unrelated. This design allowed us to minimize the possibility that observed differences in ERPs across conditions would be due to differences in terms of stimulus physical features or lexical properties rather than to the relationship between prime and target. Note that we used the same number of related and unrelated trials to avoid spurious P3/P600 contamination due to imbalanced probability between experimental conditions. If we had used only one set of unrelated pairs for two sets of related ones, we may have elicited P3/P600 effects just by virtue of their relatively lower probability (33% instead of 50%).

In addition to the critical stimuli, 216 non-cognates were selected to generate filler pairs: 144 non-cognates (two sets of 72 non-cognates) as primes and 72 as targets. Similar to the construction of the critical stimuli, each target word was paired with four different primes – two of them related and the other two unrelated. None of the filler items were the same as the experimental words. A total of 288 filler pairs were created. Half of the filler pairs served as probes as in a go/no-go semantic verification task in which participants were asked to press a yes/no button when the target was presented in red. Each filler target was presented twice in red and twice in black. Therefore, in the present study, all the critical pairs belonged to no-go trials.

Participants completed four blocks of 72 critical trials (18 trials per condition in each block) and 72 filler trials (18 go trials with yes responses, 18 go trials with no responses, and 36 no-go trials). Thus, the total number of critical trials was 288 trials (72 per condition). Each target and prime appeared only once in each block. The order of blocks was counterbalanced across participants and the presentation of items was randomized within each block.

Language history questionnaire

A questionnaire was designed to obtain information about participants' language learning experiences (see Tokowicz et al., 2004, for a similar instrument).

PROCEDURE

Participants were given the semantic judgment task, followed by the language history questionnaire.

Semantic judgment task

Participants were informed that a series of English word pairs would be presented and their task would be to judge whether the two words were related in meaning. They were asked to press a left button if the second word of a pair was red and related to the previous word and to press a right button if the word was red and unrelated. If the second word of a pair was black, they were asked not to press any button. On each trial, a prime word was presented for 300 ms, followed by a 300 ms interstimulus interval and then a target word. The target word remained on the screen for 1500 ms or until the participants pressed a button. At the end of the trial, the screen remained blank for 700 ms. Twelve practice trials preceded the experimental trials.

ELECTROPHYSIOLOGICAL RECORDING

The EEG was continuously recorded at a rate of 1 kHz from 64 Ag/AgCl electrodes placed according to the extended 10–20 convention. The 64 electrodes were referenced to Cz online. Eye blinks were monitored through additional electrodes attached above and below the right eye. Impedances were maintained below 5 kΩ. EEG signals were filtered online with a bandpass of 0.01–200 Hz and re-filtered off-line with a 30 Hz low pass zero phase shift digital filter.

DATA ANALYSES

Eye-blink artifacts were mathematically corrected using the algorithm provided in Scan 4.4 (Neuroscan, Inc.). Trials with uncorrected eye-blink artifacts and other artifacts such as horizontal eye movements and muscle movements were manually dismissed. The number of accepted trials was on average 66 per experimental condition and did not differ significantly between experimental conditions ($p > 0.10$). ERPs were then computed by averaging EEG epochs from −100 to 800 ms after the onset of the stimulus. Baseline correction was applied in relation to 100 ms of pre-stimulus activity and individual averages were re-referenced to the average of the left and right mastoid electrodes. ERPs time-locked to the target word were visually inspected and two expected ERP components (N400 and LPC) were identified. They were both maximal over centroparietal electrodes as predicted by the literature (e.g., Kutas and Hillyard, 1980; Martin et al., 2009, 2012). Peak latencies were measured at sites of maximal amplitude and mean ERP amplitudes were measured in a region of interest around the site of maximal amplitude taking into account mean global field power variations (Picton et al., 2000; Luck, 2005; centroparietal region: C1, C2, Cz, CP1, CP2, CPz, P1, P2, and Pz). The time window for each component was defined *a priori* based on previous research: 350–500 ms for the N400 (e.g., Kuipers and Thierry, 2010; Hoshino and Thierry, 2011) and 500–650 ms for the LPC (e.g., Martin et al., 2009, 2010). For statistical analyses, a four-way ANOVA was performed on mean amplitudes for each component with relatedness (related and unrelated), homograph language status (Spanish and English meaning), and electrode site (C1, C2, Cz, CP1, CP2, CPz, P1, P2, and Pz) as within-participants variables and language group (bilinguals and monolinguals) as a between-participants variable. The Greenhouse–Geisser correction was applied to variables with more than one degree of freedom in the numerator (Greenhouse and Geisser, 1959). Reported degrees of freedom are uncorrected but p-values are corrected. Significant interactions were further examined with simple effects tests. We do not report main effects of electrode site and report interactions between electrode site and other independent variables only when they are significant.

RESULTS

Event-related potentials displayed a classical P1–N1–P2 pattern in all conditions and in both participant groups. The P1, N1, and P2 did not differ in amplitude or latency between groups and conditions. An N2 modulation was visible just before the N400 time window. There was no main effect of homographs language status, relatedness, or group on the amplitude of the N2. There was a marginal interaction between all three factors [$F(1, 26) = 3.68$, $p = 0.07$]. There was no effect of any factor or interaction on N2 peak latencies and this was also the case for N400 and LPC.

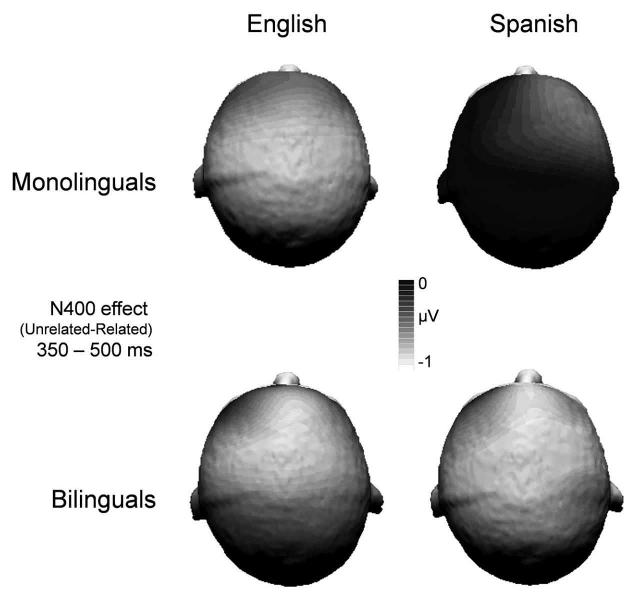


FIGURE 1 | Topographic maps for the N400 time window (350–500 ms) by language group, resulting from the subtraction of English related from English unrelated and Spanish related from Spanish unrelated.

N400 (350–500 ms)

The N400 was maximal over the centroparietal region (see **Figure 1**). There were significant main effects of relatedness and homograph language status on N400 mean amplitude [$F(1, 26) = 33.79, p < 0.001$; $F(1,26) = 4.94, p < 0.05$, respectively]. The main effect of language group did not emerge ($F < 1$). The interactions between relation and homograph language status and between homograph language status and language group were significant [$F(1, 26) = 15.48$, $p < 0.01$; $F(1, 26) = 4.64$, $p < 0.05$, respectively]. There was no interaction between relatedness and language group ($F < 1$). Most critically, the three-way interaction of relatedness, homograph language status, and language group was significant [$F(1, 26) = 9.77$, $p < 0.01$].

To further investigate the three-way interaction, simple effects tests were performed. As can be seen in **Figure 2**, monolinguals showed the interaction between relatedness and homograph language status [$F(1, 13) = 26.74$, $p < 0.001$], whereas bilinguals did not ($F < 1$). Specifically, monolinguals showed that the mean amplitude of the related condition was significantly reduced, relative to the mean amplitude of the unrelated condition for the English homograph meaning [$F(1, 13) = 39.27$, $p < 0.001$], but not for the Spanish homograph meaning ($F < 1$). In contrast,

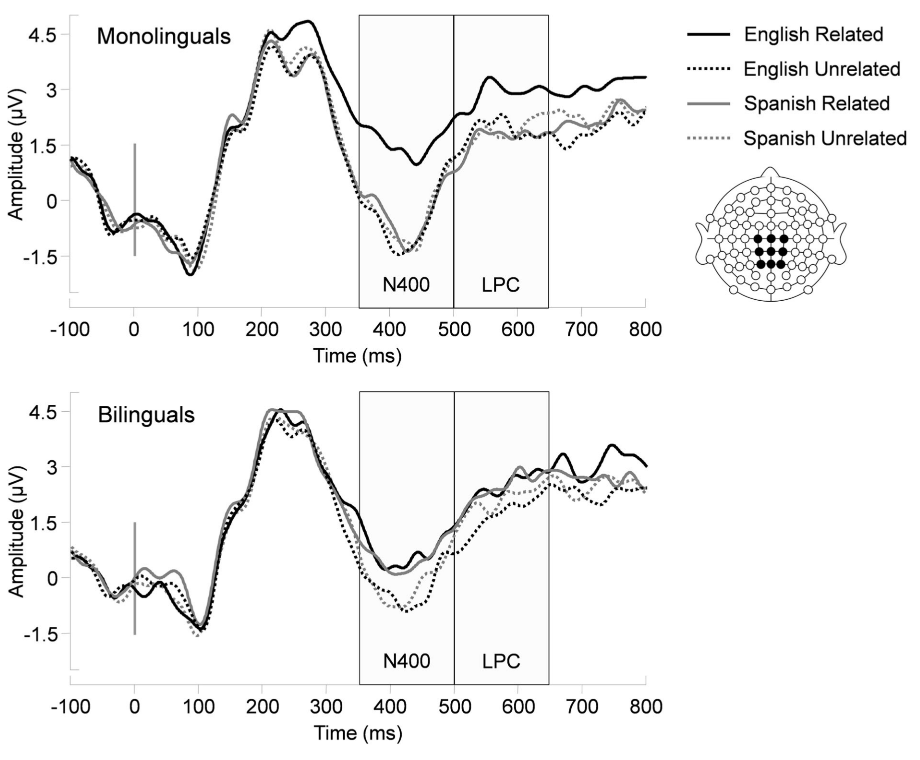


FIGURE 2 | ERPs measured over the centroparietal electrodes where N400 and LPC were maximal (a liner derivation of C1, Cz, C2, CP1, CPz, CP2, P1, Pz, P2) as a function of relatedness, homograph language, and language group.

bilinguals showed that the mean amplitude of the related condition was less negative than that of the unrelated condition for both the English homograph meaning [$F(1, 13) = 7.15$, $p < 0.05$] and the Spanish homograph meaning [$F(1, 13) = 12.65$, $p < 0.01$].

LPC (500–650 ms)

The LPC was maximal over the centroparietal region. The main effect of relatedness was significant [$F(1, 26) = 11.07$, $p < 0.01$]. The main effects of homograph language status and language group did not emerge ($F\text{s} < 1$). The interaction between relation and homograph language status was significant [$F(1, 26) = 5.56$, $p < 0.05$]. Specifically, the mean amplitude was more positive in the related than unrelated condition in English [$F(1, 27) = 11.96$, $p < 0.01$]. However, there was no difference between the related and unrelated conditions in Spanish ($F < 1$). The interactions between homograph language status and language group and between relation and language group were not significant [$F(1, 26) = 3.32$, $p = 0.08$; $F < 1$, respectively]. The three-way interaction of relatedness, homograph language status, and language group was not reliable [$F(1, 26) = 1.36$, $p > 0.10$].

DISCUSSION

The goal of the present study was to examine the time window in which the non-target language is accessed in late Spanish–English bilinguals reading English. The time course of the activation of homograph meanings in the non-target language was tracked using ERPs. We found that monolinguals displayed semantic priming only when primes were related to the English meaning of homograph targets. This semantic priming effect was observed in the N400 time window (350–500 ms) but also affected ERP amplitudes in the LPC time window (500–650 ms). In contrast, Spanish–English bilinguals showed semantic priming effects in the N400 time window (350–500 ms) when primes were related either to the English or the Spanish meaning of homograph targets. However, the LPC effect was found for stimulus pairs related on the basis of the English meaning of homographs only. These results suggest that both languages are activated in the classical time frame of semantic activation indexed by N400 modulations, but that semantic activation in the non-target language failed to be explicitly processed by bilingual participants.

The time course of activation of non-target homograph meaning in the present study is compatible with a previous ERP studies on processing of interlingual homographs out of context (Kerkhofs et al., 2006; Paulmann et al., 2006). In the study by Kerkhofs et al. (2006), the frequency effect in the non-target language, which was an index of non-target homograph meaning activation, was observed from 400 to 550 ms but was not found from 550 to 650 ms. However, the present result contrasts with previous research on the processing of interlingual homographs in sentence context (Elston-Güttler et al., 2005). In the study by Elston-Güttler et al. (2005), German–English bilinguals activated the non-target meaning of homographs from 150 to 250 ms and from 300 to 500 ms only when their L1 German was activated through viewing a film in German prior to the experiment and only for the first half of the experiment. This result suggests that a sentence context can suppress non-target language activation before the stage of post-lexical processing.

The absence of explicit processing in the case of word pairs related via the non-target L1 is compatible with the inhibitory control (IC) model proposed by Green (1998). Indeed, if inhibition occurs, it is likely to intervene after automatic aspects of lexico-semantic processing. According to the IC model, both L1 and L2 lexico-semantic representations are activated and the selection of a target representation is dependent on the language which is more active at a given time. A task schema is hypothesized to be responsible for controlling the level of activation of lexico-semantic representations in the non-target language. When bilinguals intend to perform a task in one language, the task schema inhibits the activation of lexico-semantic representations in the non-target language. In the present study, Spanish–English bilinguals activated lexico-semantic representations with English and Spanish language tags while reading a series of pairs of English words, which was reflected in the N400 time window (350–500 ms)[1]. Although half of the trials included interlingual homographs in English and Spanish, the instructions were given exclusively in English, the experiment was introduced as consisting of English words, and precaution was taken not to mention the inclusion of interlingual homographs. Therefore, participants did not need to be aware of the presence of word targets related to prime via Spanish, and indeed they were not given the lack of an LPC variation in that condition.

Similarly, the bilingual interactive activation (BIA)+ model (Dijkstra and van Heuven, 2002) is compatible with the absence of a modulation in the LPC range for interlingual homographs related via the non-target language L1. The BIA+ model assumes that the word identification system can be influenced by the linguistic context, but not by the non-linguistic context (e.g., task instruction, task demands, etc.), and that the influence of non-linguistic context on word recognition is post-lexical. In the present study, semantic priming effects were found for both meanings of interlingual homographs in the target and non-target languages in the N400 time window, which is supposed to reflect lexico-semantic processing. However, the semantic priming effect in the non-target language did not result in an LPC modulation, which is likely to reflect post-lexical re-evaluation processes associated with explicit access.

One concern about the design of the present study was the repetition of stimuli. As described in the Section "Materials and Methods," we used the same set of targets for all the experimental conditions in order to minimize the possibility that observed differences in ERPs across conditions would be due to differences in terms of stimulus physical features or lexical properties rather than to the relationship between prime and target. Four repetitions were involved in this design. To assess the effect of repetitions, we reanalyzed the bilingual data by including experimental block as an additional within-participants variable. A 2 (relatedness: related and unrelated) × 2 (homograph language status: Spanish and English meaning) × 2 (block: first half and second half) × 9 (electrode site: C1, C2, Cz, CP1, CP2, CPz, P1, P2, and Pz) ANOVA showed that the variable "block" did not

[1]We note that there was a marginal three-way interaction on N2 mean amplitudes as well, suggesting the lexical-semantic integration may have started slightly earlier in English monolinguals than Spanish–English bilinguals.

interact with any other variable (all $ps > 0.10$). This result suggests that the observed pattern of effects was not affected by repetitions.

In conclusion, the present ERP study suggests that bilinguals activate semantic representations of interlingual homographs in the non-target language as well as in the target language initially but process these representations at an explicit level only for meaning in the target language. In future research, it will be critical to investigate the role of L2 proficiency, the nature of the target language (e.g., L1 rather than L2), and the linguistic and non-linguistic contextual parameters in the modulation of cross-languages activation at lexical and post-lexical processing stages (cf., Wu and Thierry, 2010). If the target language were the L1, the activation of the non-target language might not be present because of the relatively weaker representations in L2. Resolving these issues will provide implications for models of bilingual language processing.

ACKNOWLEDGMENTS

This research was supported by the Economic and Social Research Council U.K. (RES-E024556-1). We thank Jan-Rouke Kuipers and Clara Martin for technical assistance.

REFERENCES

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database [CD-ROM]*. Philadelphia, PA: University of Pennsylvania.

Beauvillain, C., and Grainger, J. (1987). Accessing interlexical homographs: some limitations of a language-selective access. *J. Mem. Lang.* 26, 658–672.

De Bruijn, E., Dijkstra, A., Chwilla, D. J., and Schriefers, H. J. (2001). Language context effects on interlingual homograph recognition: evidence from event-related potentials and response times in semantic priming. *Biling. Lang. Cogn.* 4, 155–168.

Dijkstra, A., Grainger, J., and van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: the neglected role of phonology. *J. Mem. Lang.* 41, 496–518.

Dijkstra, A., Van Jaarsveld, H., and Ten Brinke, S. (1998). Interlingual homograph recognition: effects of task demands and language intermixing. *Biling. Lang. Cogn.* 1, 51–66.

Dijkstra, T., and van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. *Biling. Lang. Cogn.* 5, 175–197.

Elston-Güttler, K. E. (2000). *An Inquiry into Cross-Language Lexical-Conceptual Relationships and Their Effect on L2 Lexical Processing.* Unpublished doctoral dissertation, University of Cambridge, Cambridge.

Elston-Güttler, K. E., Gunter, T. C., and Kotz, S. A. (2005). Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. *Brain Res. Cogn. Brain Res.* 25, 57–70.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrica* 24, 95–112.

Hoshino, N., Midgley, K. J., Holcomb, P. J., and Grainger, J. (2010). An ERP investigation of masked cross-script translation priming. *Brain Res.* 1344, 159–172.

Hoshino, N., and Thierry, G. (2011). Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain Res.* 1371, 100–109.

Jared, D., and Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *J. Mem. Lang.* 44, 2–31.

Kerkhofs, R., Dijkstra, A., Chwilla, D., and De Bruijn, E. (2006). Testing a model for bilingual semantic priming with interlingual homographs: RT and ERP effects. *Brain Res.* 1068, 170–183.

Kuipers, J. R., and Thierry, G. (2010). Event-related brain potentials reveal the time-course of language change detection in early bilinguals. *Neuroimage* 50, 1633–1638.

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.

Lemhöfer, K., and Dijkstra, T. (2004). Recognizing cognates and interlexical homographs: effects of code similarity in language specific and generalized lexical decision. *Mem. Cognit.* 32, 533–550.

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique.* Cambridge, MA: MIT Press.

Macizo, P., Bajo, T., and Martín, M. C. (2010). Inhibitory processes in bilingual language comprehension: evidence from Spanish-English interlexical homographs. *J. Mem. Lang.* 63, 232–244.

Martin, C. D., Costa, A., Dering, B., Hoshino, N., Wu, Y. J., and Thierry, G. (2012). Effects of speed of word processing on semantic access: the case of bilingualism. *Brain Lang.* 120, 61–65.

Martin, C. D., Dering, B., Thomas, E. M., and Thierry, G. (2009). Brain potentials reveal semantic priming in both the 'active' and the 'non-attended' language of early bilinguals. *Neuroimage* 47, 326–333.

Martin, C. D., Thierry, G., and Démonet, J. F. (2010). ERP characterization of sustained attention effects in visual lexical categorization. *PLoS ONE* 5, e9892. doi:10.1371/journal.pone.0009892

Martín, M. C., Macizo, P., and Bajo, T. (2010). Time course of inhibitory processes in bilingual language processing. *Br. J. Psychol.* 101, 679–693.

Midgley, K. J., Holcomb, P. J., and Grainger, J. (2009). Masked repetition and translation priming in second language learners: a window on the time-course of form and meaning activation using ERPs. *Psychophysiology* 46, 551–565.

Paulmann, S., Elston-Güttler, K. E., Gunter, T. C., and Kotz, S. A. (2006). Is bilingual lexical access influenced by language context? *Neuroreport* 17, 727–731.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R. Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., and Taylor, M. J.(2000). Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology* 37, 127–152.

Thierry, G., Athanasopoulos, A., Wiggett, A., Dering, B., and Kuipers, J. R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4567–4570.

Thierry, G., Cardebat, D., and Démonet, J.-F. (2003). Electrophysiological comparison of grammatical processing and semantic processing of single spoken nouns. *Brain Res. Cogn. Brain Res.* 17, 535–547.

Thierry, G., Doyon, B., and Démonet, J. F. (1998). ERP mapping in phonological and lexical semantic monitoring tasks: a study complementing previous PET results. *Neuroimage* 8, 391–408.

Thierry, G., and Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign language comprehension. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12530–12535.

Tokowicz, N., Michael, E., and Kroll, J. F. (2004). The roles of study abroad experience and working memory capacity in the types of errors made during translation. *Biling. Lang. Cogn.* 7, 255–272.

Wu, Y. J., and Thierry, G. (2010). Investigating bilingual processing: the neglected role of language processing contexts. *Front. Psychol.* 1:178. doi:10.3389/fpsyg.2010.00178

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 May 2011; accepted: 09 January 2012; published online: 02 February 2012.

*Citation: Hoshino N and 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. Front. Psychology **3**:9. doi: 10.3389/fpsyg.2012.00009*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY
ORIGINAL RESEARCH ARTICLE
published: 20 September 2011
doi: 10.3389/fpsyg.2011.00219

Error-related activity and correlates of grammatical plasticity

Doug J. Davidson[1]* and Peter Indefrey[2]

[1] Basque Centre on Cognition, Brain and Language, Donostia, Basque Country, Spain
[2] Institut für Sprache und Information, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

Edited by:
Judith F. Kroll, Penn State University, USA

Reviewed by:
Natasha Tokowicz, University of Pittsburgh, USA
Susan C. Bobb, Georg-August-Universität Göttingen, Germany
Janet McDonald, Louisiana State University, USA

***Correspondence:**
Doug J. Davidson, Basque Centre on Cognition, Brain and Language, Mikeletegi Pasealekua 69, 2, 20009 Donostia, Basque Country, Spain.
e-mail: ddavidson@bcbl.eu

Cognitive control involves not only the ability to manage competing task demands, but also the ability to adapt task performance during learning. This study investigated how violation-, response-, and feedback-related electrophysiological (EEG) activity changes over time during language learning. Twenty-two Dutch learners of German classified short prepositional phrases presented serially as text. The phrases were initially presented without feedback during a pre-test phase, and then with feedback in a training phase on two separate days spaced 1 week apart. The stimuli included grammatically correct phrases, as well as grammatical violations of gender and declension. Without feedback, participants' classification was near chance and did not improve over trials. During training with feedback, behavioral classification improved and violation responses appeared to both types of violation in the form of a P600. Feedback-related negative and positive components were also present from the first day of training. The results show changes in the electrophysiological responses in concert with improving behavioral discrimination, suggesting that the activity is related to grammar learning.

Keywords: error-related activity, language learning, plasticity, bilingualism, morphosyntax

1. INTRODUCTION

Grammatical learning has been subject to extensive debate in linguistics, psychology, and neuroscience. One reason for this is the widely discussed sensitive period hypothesis (Lenneberg, 1967; Johnson and Newport, 1989, 1991; Newport et al., 2001), which maintains that the adult-onset learning of grammatical features or rules is less effective than child-onset learning. A second reason is the practical relevance of adult grammar learning for second language (L2) learning and bilingualism, both of which have a profound impact on group social organization and work productivity (Knudsen et al., 2006). However, the dynamic patterns of change in grammar learning are much less discussed. Learning dynamics could be important in these debates because grammatical features are learned over multiple, embedded time scales: The learning that occurs in the span of days or hours of a single lecture is embedded within the years or months that make up a second language course. In principle, learning must include brain activity at even shorter time scales, such as the time span of individual sentences or words, because the learning events that make up the longer time scales of the lecture or the course consist of individual sentences or words. For these reasons, a potentially effective method to study the cortical mechanisms involved in grammatical learning would be to relate the cortical activity present at the shorter temporal scales of individual sentences to activity at longer temporal scales.

The event-related potential (ERP) is an effective tool to understand this process because unlike other measures of physiology such as fMRI or PET, it has sufficient temporal resolution to separate responses to individual words within a sentence, as well as to a classification response, if it is used in a learning task, or the response to the feedback that might occur after the classification response. In studies of first language comprehension, ERP grammatical violation responses have been observed to a variety of morphosyntactic errors within a sentence (Hagoort et al., 1993; Osterhout and Holcomb, 1995; for review see Kutas et al., 2006). These responses have also been observed in adult language learners who have obtained a relatively high level of second language proficiency, but usually not in those who have not (Weber-Fox and Neville, 1996; Hahne, 2001; Hahne and Friederici, 2001; Friederici et al., 2002; Rossi et al., 2006; see Kotz, 2009 for a recent review). These L2 studies indicate that at a time scale that is sufficiently long to attain (behavioral) L2 grammatical competence, the electrophysiological responses to L2 violations have also changed to resemble those of the native L1 response. However, many of these studies have compared cross-sections of learner groups (Weber-Fox and Neville, 1996; Hahne and Friederici, 2001; Rossi et al., 2006), or relatively long time scales of learning in longitudinal designs (Osterhout et al., 2006, 2008; cf. Mueller et al., 2005). Strikingly, the work of Osterhout and collegues has shown a sequence of ERP violation effects such that grammatical violations elicit an N400 effect early in learning, but in contrast a P600 more like the native response later on. Morgan-Short et al. (2010) have recently compared explicit and implicit training conditions for an artificial grammar-learning task, observing either an N400-like pattern or P600-like pattern depending on proficiency levels and the type of instruction. It is, however, less clear which types of EEG activity are present at shorter time scales when learners are acquiring knowledge, or perhaps more importantly, how this activity is related to behavioral change.

One way this could be done is to examine explicit learning. This might occur, for instance, when learners judge sentences as grammatical or not (e.g., providing a classification response), in the case where feedback is provided to indicate whether the classification is correct or not (e.g., comprehending a feedback signal). The feedback would allow learners to establish, over a number of trials, which features of the sentences are relevant to obtain a correct classification, so that over trials more of the behavioral responses would be correct. In the electrophysiological literature, there is a second type of ERP effect observed in choice–response tasks related to this process of explicit learning which can be viewed as error-related activity (Falkenstein et al., 1991; Gehring et al., 1993). The response effect, termed the error-related negativity (N_e) is obtained by subtracting the average ERP for correct choices from error choices in a time window of −150 ms before and +150 ms after participants make a behavioral response, appearing on centro-frontal electrodes. An error-related positivity (P_e) is observed after the N_e, in the time window from +150 to 300 ms post-response on a similar set of electrodes (Overbeek et al., 2005). There are also error-related ERP effects seen in response to feedback that informs subjects that their previous responses were incorrect. In this paper we will refer to these as feedback-N_e and feedback-P_e, so that the terminology is symmetrical to the behavioral response terminology. It should be noted, however, that the positive component is often described as a P300 effect in the literature. Obtained by a similar subtraction of correct feedback-related ERPs from error feedback-related ERPs, the time windows for the feedback-N_e (+150 to +300 ms post-feedback) and feedback-P_e (+300 to +500 ms post-feedback) reflect the average electrophysiological response to the feedback stimulus.

There is an extensive body of research in cognitive neuroscience that shows that brain areas in medial frontal cortex are involved in cognitive control, including performance monitoring, response errors, conflict, and uncertainty about correct responses, as well as responses to feedback indicating that performance is correct or incorrect in learning tasks (Ridderinkhof et al., 2004a,b; Overbeek et al., 2005). This work seeks to provide a unified account of brain activity in situations of response conflict as well as during adaptive adjustments to the environment. Earlier work by Holroyd and Coles (2002) provided a theoretical account of the changing relationship between the feedback-N_e and response-N_e in learning tasks. In their account, the initial large-amplitude feedback-N_e (but not response-N_e) reflects a learning process in which the internal state of learners is modified to predict the likely outcome of behavioral choices. Later in training, this modification is reflected in a larger response-N_e (and not feedback-N_e) at the point in time where the response choice is made. It is in this sense that the components can be seen as correlates of cognitive control: Learners adjust their internal state according to the constraints of the experimental task. The interpretation of the response- or feedback-P_e is less clear in the literature. It has been argued that the response P_e is a type of P300 that is related to error awareness (Leuthold and Sommer, 1999; Frank et al., 2007), which predicts that it could be a mediating factor in learning-related improvement.

This study therefore investigated how violation-, response-, and feedback-related activity correlates with behavioral grammatical learning of declension, following a previous study investigating the same topic (Davidson and Indefrey, 2009). In particular, we focus on the dynamics of these ERP components to better understand how events at shorter time scales of individual learning sessions are related to the stability of grammatical knowledge at the time scale of weeks or months.

Several recent studies of second language usage or artificial grammar learning have also employed error-related activity. Sebastian-Gallés et al. (2006) showed that error-related negativity was reduced or absent in Spanish-dominant early Spanish–Catalan bilinguals making a difficult non-word decision. Although this study employed multiple ERP effects (N400, N_e) to examine the lexical decision making process, it did not directly concern learning processes linked to these components. Also studying already-proficient learners, Ganushchak and Schiller (2009) showed that an N_e effect present on error trials of a phonome-monitoring task was larger with increased time pressure in Dutch–German bilinguals. Finally, Opitz et al. (2011) have shown using an artificial grammar-learning task with visually presented strings that a feedback-related negativity (N_e in the present paper) remained constant over learning, while the amplitude of a feedback-related positivity decreased with learning. Like the present study, this study employed both classification and feedback, although the task was not that of L2 language learning. The present study, like Davidson and Indefrey (2009) attempts to use multiple ERP components present in a learning task (P600, response-N_e/P_e, feedback-N_e/P_e) to try to understand how brain activity that is linked to discrimination or learning changes over time.

The experiment reported here investigates how Dutch learners acquire German morphosyntactic distinctions related to gender and declension within short prepositional phrases (see the example in **Table 1**). The task for the learners was the same as in Davidson and Indefrey (2009). In brief, they were asked to judge the correctness of prepositional phrases in which we manipulated whether or not the adjective carried syntactic feature information and whether or not there was gender agreement between the head noun and the preceding determiner or adjective. In German, the expression of case, number, and gender features on an adjective depends on the preceding elements of the noun phrase. This dependency is considered to be a syntactic dependency (Zwicky, 1986) and, following Schlenker (1999), it can be described as a rule according to which syntactic features are to be expressed only on the first inflectable element of a noun phrase. If the adjective

Table 1 | Example phrases illustrating the experimental conditions.

Correct, 3 word	mit kleinem$_{[+Dat, -F, -Pl]}$ Kind$_{[-F, -M]}$ with small child *"with a small child"*
Correct, 4 word	mit dem$_{[+Dat, -F, -Pl]}$ kleinen$_{[]}$ Kind$_{[-F, -M]}$ with the small child *"with the small child"*
Declension violation, 3 word	mit *kleinen$_{[]}$ Kind$_{[-F, -M]}$
Declension violation, 4 word	mit dem$_{[+Dat, -F, -Pl]}$ *kleinem$_{[+Dat, -F, -Pl]}$ Kind$_{[-F, -M]}$
Gender violation, 3 word	mit kleinem$_{[+Dat, -F, -Pl]}$ *Frau$_{[+F]}$
Gender violation, 4 word	mit dem$_{[+Dat, -F, -Pl]}$ kleinen$_{[]}$ *Frau$_{[+F]}$

Dat, dative; F, feminine; M, masculine; Pl, plural.

is the first inflectable element of a noun phrase, it takes on a suffix of the "strong" declension paradigm. The suffix *−em* in "mit kleinem$_{[+Dat, -F, -Pl]}$ Kind$_{[-F, -M]}$" ("with a small child"), for example, specifies dative case, non-feminine gender, and singular. By contrast, if the adjective is preceded by a definite determiner that expresses the feature information, the adjective has a suffix from the weak declension paradigm that is compatible with the feature specification of the determiner but does not express the features itself ["mit dem$_{[+Dat, -F, -Pl]}$ kleinen$_{[]}$ Kind$_{[-F, -M]}$" (also "with the small child")]. According to previous linguistic analyses of German adjective declension the weak *−en* suffix can be seen as a default or "elsewhere" form (Bierwisch, 1967; Zwicky, 1986; Blevins, 1995, 2003; Cahill and Gazdar, 1997; Wunderlich, 1997; Schlenker, 1999; Clahsen et al., 2001; Penke et al., 2004). For our stimuli, we used a subset of the full German paradigm involving dative case singular noun phrases, in order to restrict the learning problem. Please see Davidson and Indefrey (2009) for further details, as well as the primary linguistic work (Zwicky, 1986; Schlenker, 1999) for a description of the full paradigm. Examples for correct noun phrases and the declension and gender violation conditions are given in **Table 1**.

In Davidson and Indefrey (2009), both native German speakers and Dutch L2 learners of German responded to declension violations with P600 effects, but for gender violations, only native speakers showed P600 effects. In that study, after an initial pre-test phase in which no instructions or feedback was provided, we provided explicit instructions for classifying these phrases, and feedback immediately after the classification response. In the present study, we again used a pre-test phase without explicit instructions or feedback, and in the training phase provided the feedback, but with some delay, after the classification response. We also changed the procedure by not providing instructions about the grammatical rules. These two changes to the EEG experiment (slightly delayed feedback and no explicit instructions) were designed to reveal changes in the different aspects of the error-related activity over time. The separation of the behavioral response and the feedback was designed to examine differences in the dynamical behavior of the response-related and feedback-related activity, to see whether the predicted changes in activity derived from Holroyd and Coles' (2002) account apply in this task. Also, without explicit instructions, it was hypothesized that participants would take longer to reach the comparable levels of proficiency. This was based on the assumption that the previously used instructions had informed participants about which aspects of the phrases to attend and remember during the task. Without instruction, there should be a slower evolution of the changes in behavior, and allow us to see more clearly how the ERP responses are related to behavior.

1.1. SUMMARY AND HYPOTHESES

Based on our previous work, we hypothesized that changes in violation- and error-related responses will occur in conjunction with morphosyntactic learning, to the extent that this learning is revealed by classification performance. Our linking assumption is that ensemble electrophysiological activity can be recorded with EEG in language learners related to the following: (*i*) recognizing grammatical constraints, (*ii*) making correct and incorrect choices, and (*iii*) processing feedback signals. With respect to recognition of grammatical constraints, averaging the single trials of EEG and comparing violation ERPs to their controls should reveal a P600 violation response in the learners. Our assumption is that the synchronized ensemble activity giving rise to the violation ERP reflects the recognition or repair of grammatical violations. A prediction from this is that the P600 amplitude to the violations will be greater after learning than before learning to the extent that the learners can employ the knowledge they have acquired in real time. With respect to the electrophysiological response to feedback signals, it is assumed that comparing the ERP to feedback indicating an incorrect choice to the ERP to feedback indicating a correct choice will reveal difference components such as the N_e (and possibly the P_e), because this has been observed in previous EEG work with two-alternative forced-choice responses. The prediction is that the feedback N_e or P_e effect will be present early during learning but decrease in amplitude over learning trials, as predicted by the Holroyd and Coles (2002) account. With respect to the behavioral classification, participants' discrimination should improve over trials. In concert with this, a response-related N_e could appear, as this is also predicted by the Holroyd and Coles (2002) account. Finally, individual learner variation in the violation- and/or error-related ERP magnitudes should be statistically related to variation in grammatical classification, to the extent that there is a simple and direct (linear) relationship between the activity and the classification performance (see also van der Helden et al., 2010).

The present experiment also included additional behavioral tasks, including an *n*-back test of working memory and a computerized version of the Wisconsin card sort task. These behavioral tasks were used in an attempt to measure components of individual variation which might be related to the learning task. It was hypothesized that differences in working memory ability, for example, might be related to participants' ability to remember the outcome of previous trials while processing the phrases. The card-sorting task was hypothesized to relate to participants' tendency to change classification rules in response to feedback.

2. METHOD

2.1. PARTICIPANTS

Twenty-two native Dutch speakers (16 female, all right-handed, average age $M = 23.1$ years, $SD = 3.1$ years, range 19–29 years) were recruited with posted advertisements from Radboud University in Nijmegen, The Netherlands, a city near a border with Germany. The advertisements described a generic EEG experiment, and did not refer to language learning or to German instruction. As shown in **Table 2**, most participants had previous coursework in German during high school, and their average self-rated proficiency (5-point scale, $max = 5$ indicating high proficiency) was near the midpoint of the scale, or below the midpoint. This was true for most language skill components: speaking ($M = 2.1$), listening ($M = 2.8$), writing ($M = 1.7$), reading ($M = 2.8$), grammar ($M = 1.4$), and expression ($M = 2.0$). Also, recent exposure (self-reported number of hours in the last 3 months) was relatively low. Before completing the EEG tasks, all participants completed a European Reference Frame multiple choice assessment of German (maximum 30 possible correct) prepared by the Goethe Institute[1].

[1] www.goethe.de

Table 2 | Participant variables related to knowledge of German or valence toward German.

Variable	Mean	SD	Range (min–max)
Self-rated proficiency (average)	2.1	1.1	1.3–3.0
Age of initial German language education (years)	13.5	5.0	12–23
Duration of German language education (years)	3.2	2.4	0–6
German language proficiency, global test	14.7	3.5	9–22
German language proficiency, specific	3.7	1.3	1–6
Comfortable–uncomfortable using German	1.9	1.4	1–3
Like–dislike using German	3.5	1.8	1–5
Important–non-important to use German	2.5	2.5	1–5
Easy-difficult to use German	3.1	1.5	2–4
Recent exposure to German	0.5	1.0	0–1

Six of the questions on the test concerned morphosyntactic properties, the average score for this subset ($max = 6$) was slightly higher than chance. The Goethe Test scores indicate relatively low levels of German proficiency, and in addition, the test scores were not significantly correlated with self-rated proficiency, $r = 0.276$.

2.2. DESIGN AND PROCEDURE

The design of the experiment (see **Table 1**) was similar to that of Davidson and Indefrey (2009). In the pre-test, training, and follow-up phases, there were three repeated measures factors: prepositional phrase Grammaticality (violation, control), number of words in the phrase (3 or 4, corresponding to strong and weak forms of the adjective, respectively), and sentence Type (declension, gender).

The procedure consisted of two experimental EEG sessions and a third behavioral-only follow-up. The first EEG session included several behavioral measures, a pre-test and the first part of the training phase. The second EEG session, approximately 1 week after the first session ($M = 8.0$ days, SD = 1.9, $min = 5$, $max = 13$), included behavioral measures and the second and third parts of the training phase. All participants completed both EEG sessions. Approximately three and a half months after the second EEG session ($M = 109.6$ days, SD = 17.6, $min = 83$, $max = 148$), 15 participants returned for a behavioral follow-up (the others did not respond to the follow-up request, or were not available).

During both EEG sessions, the main experimental task was the classification task. This task consisted of a series of trials in which phrases were presented on a CRT monitor at 300 ms/word with an ISI of 600 ms between words. The words were presented in 26 point white Arial characters on a black background in a dimly lit room. Each trial began with a yellow fixation cross for 600 ms, followed by the first word of the phrase. The last word of the phrase was followed by a white fixation cross, which remained on screen until 1 s after participants responded.

The pre-test was conducted to assess participants' grammatical knowledge of and performance on the materials used in the experiment at the start of the experiment. During this pre-test (as well as in the behavioral-only follow-up), participants classified phrases without feedback on their response choices. Participants classified phrases as acceptable or not acceptable by pressing one of two keys with the index or ring fingers of their right hand.

The learning phase on the first day followed the pre-test after a short (5 min) break. During the learning phase, the classification task was presented just as in the pre-test, but with feedback after each response. Specifically, 1.0 s after participants' response choice, a small green (red) square was presented on the center of the screen for 0.25 s, indicating that their classification was correct (incorrect). Note that the feedback did not indicate the source of the participants' correct response or error, and it did not indicate the correct version of the presented phrase, only whether the classification was correct or not. After the feedback, the next trial began 1 s after the feedback signal.

In addition to the classification task, several other behavioral measures were provided during the application of the electrodes before the classification task in both EEG sessions. These included an *n*-back test of working memory (Owen et al., 2005), and a computerized version of the Wisconsin card-sorting task. These tasks were included to provide different measures of individual differences in general and language-related performance. EEG was not recorded during their administration.

2.3. MATERIALS

Four common German adjectives (*klein, groß, alt, neu*; respectively "small," "large," "old," "new") and 40 common German nouns were chosen to serve as stimulus materials, as well as the preposition *mit* (with) and the determiners *dem* and *der*; corresponding to dative case neuter and feminine forms of the definite determiner ("the"). Dutch has a two-gender system with a neuter gender corresponding to the neuter gender of most German cognate nouns and a so-called common gender corresponding to the masculine or feminine gender of most German cognate nouns. The nouns were chosen so that they had the corresponding gender of the Dutch translation (neuter: *Fenster, Haus, Pferd, Gleis, Schaf, Buch, Glas, Bett, Messer, Institut, Museum, Hemd, Hotel, Gebäude, Bild, Dorf, Büro, Schloss, Schiff, Auge*; corresponding to (respectively) "window," "house," "horse," "track," "sheep," "book," "glass," "bed," "knife," "institute," "museum," "shirt," "hotel," "building," "picture," "village," "office," "castle," "ship," "eye"; and feminine: *Tür, Schule, Kuh, Katze, Straße, Geschichte, Tasse, Couch, Gabel, Universität, Ausstellung, Hose, Garage, Wohnung, Zeichnung, Stadt, Bank, Kirche, Bahn, Nase*; corresponding to "door," "school," "cow," "cat," "street," "story," "cup," "couch," "fork," "university," "exhibition," "pants," "garage," "house," "drawing," "city," "bank," "church," "train," "nose"). To ensure that determiner and adjective forms unambiguously predicted the gender of the following nouns, we did not use masculine nouns, because the masculine forms of the determiner and adjectives are identical to neuter forms in German. The critical words (CW) included the adjective and noun for the various conditions (see **Table 1**). The phrases were created by pairing each noun to two of the adjectives in one set of stimuli and to the remaining two adjectives in a second set of stimuli. In each phase (Pre-test, Training 1–3), there were 240 stimuli. These consisted of 40 phrase stimuli presented for each violation and each control condition for the three- and the four-word versions of the phrases, for the gender and the declension contrasts. Thus, on the first day of the experiment, there were 240 items presented in the pre-test, involving six repetitions of a particular noun and three repetitions of a particular article–noun pair. A distinct set of 240 items was presented in Training 1. In the second day (1 week later), the same items were presented again, 240 in Training 2, and 240 in Training 3. The number of trials did not vary between sessions. Additional practice items (10) were presented before the Pre-test to insure that participants understood the task.

2.4. APPARATUS

EEG was recorded from 64 electrodes using battery-powered BrainVision BRAINAMP Series amplifiers (Brain Products GmbH, München, Germany). Signals were sampled at 500 Hz, with a low-pass filter at 200 Hz and a high-pass filter with a time constant of 159 s during acquisition. Electrodes were applied to an equivalent inter-electrode distance Easy-Cap (Brain Products; see **Figure 1** for the electrode arrangement). Impedance levels were kept below 10 kΩ at the electrode–skin interface, with input impedance at the amplifiers at 10 MΩ (see Ferree et al., 2001). The data were recorded with respect to a left mastoid reference, and later re-referenced to an average reference including all electrodes before analysis. An additional electrode was placed below the left eye to record activity related to vertical eye movements referenced to an electrode above the eye. Lateral eye movement activity was recorded as the difference between channels near the left and right canthus.

2.5. DATA ANALYSIS

Fixed and random effects for the behavioral measures, in addition to several covariates, were modeled using a general linear mixed-effects model approach (Bagiella et al., 2000; Pinheiro and Bates, 2000; Friston et al., 2005; Baayen et al., 2008). Beta weights for the regression (b) and a t-test for the parameter values (t) are provided in the text in order to show the magnitude of the effects. Uncertainty in the parameter estimates was evaluated using highest posterior density (*HPD*) intervals, which can be treated as 95% confidence intervals for the regression parameters. The outcome measure in the regression analysis was discrimination performance (average d'), calculated as the (z-transformed) average proportion of correctly rejected violation stimuli minus the (z-transformed) average proportion of control stimuli incorrectly classified as a violation. This use of d' was intended as a measure of the ability of participants to reject violations, while at the same time correcting for false alarms.

In addition to the EEG measures, the covariates were taken as six loadings (accounting for 80% of the variance, identified by plotting variance captured by the loadings) from a principle components analysis (PCA) of a set of variables (see **Table 2**), including average self-rated proficiency (5-point scale for speaking, listening, writing, reading, grammar, and expression), age of initial German language education, German language proficiency as measured by the Goethe Institute Test (Goethe-Institut, 2005; both the global test, and specific items concerning case and gender), number of errors on the Wisconsin card sort task ($M = 21.7$ errors, range 3–64, out of 136 trials), average RT following an error on the Wisconsin card sort task, the slope of each participant's error curve on the n-back test (average proportion correct for $n = 1,2$, and 3 was 0.86, 0.76, and 0.68, respectively), several 5-point scale measures of valence toward learning German (comfortable–uncomfortable using German, like–dislike, important–non-important, easy–difficult), as well as an indicator of recent exposure to German (number of hours in the last 3 months). The questions for the valence toward German scales asked participants to rate whether, e.g., they liked to use German. The selection of these variables was based on a preliminary inspection of the individual difference measures, which suggested that several variables were correlated with each other. Those measures which appeared to have substantive variability across subjects were entered into the PCA analysis in order to identify a collection of linearly independent factors for the regression analysis. Together, the loadings on these principle components were hypothesized to reflect individual variation in general task performance, German proficiency, attitude toward learning German, and recent German exposure. The principle components PC1 to PC6 were most highly loaded, respectively, on the following single factors: comfortable–uncomfortable, Goethe Institute specific test, like–dislike, Wisconsin card sort number of errors, n-back task slope, and Goethe Institute global test. Note that only in the case of the n-back loading (PC5) was a single factor clearly related to the loading. In all other cases, multiple factors were related.

For the ERP analyses, trials of the recorded EEG data containing eye movement, muscle, and other noise artifacts were excluded using Matlab-based preprocessing functions[2] (Oostenveld et al., 2011), filtered with a low-pass filter (two-pass

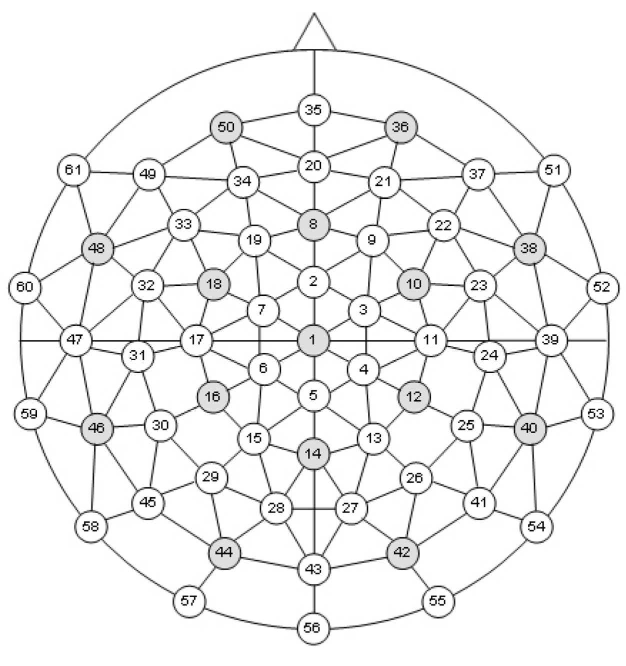


FIGURE 1 | Electrode array with locations in gray indicating approximate 10–20 locations.

[2] www.ru.nl/donders/fieldtrip/

6th-order Butterworth finite impulse response) with square-root half-maximum attenuation at 20 Hz, re-referenced to an average reference (please see Nunez and Srinivasan, 2006 for discussion of different re-referencing schemes), and segmented into 1 s epochs consisting of 100 ms before the onset of the CW and 900 ms following the CW. The resulting epochs were baselined with respect to the 100-ms baseline interval before CW onset and averaged according to experimental condition. Only trials with correct responses were included in the violation–control ERP contrasts, and only those participants with at least 10 observations in both violation and control conditions (two participants were excluded on this basis). The time interval for the P600 effect was defined as the range from 500 to 900 ms. On average, the numbers of non-error observations contributing to the average ERPs for the gender (declension) violation and control conditions were: $M = 17.5$, 17.8, (18.7, 18.8) for the four-word versions, and $M = 18.1$, 18.0, (18.7, 19.1) for the three-word versions. Response-locked data were averaged to quantify activity related to correct and incorrect responses in two time windows based on inspection of the grand average response-locked waveforms: −150 to 150 ms (N_e) and 150 to 300 ms (P_e). In both cases, the baseline interval was −300 to −150 ms. The response-locked epochs were baselined with respect to the interval from −400 to −200 ms before response onset. Feedback trials were time-locked to the feedback stimulus. The feedback N_e was defined over the interval from 100 to 300 ms, while the feedback-P_e was defined over 300–500 ms. For the feedback N_e/P_e ERPs, there were $M = 114.2$ and $M = 70.1$ trials for correct and incorrect for Training 1; $M = 132.0$ and $M = 66.8$ for Training 2, and $M = 149.0$ and $M = 45.8$ for Training 3. For the response-N_e/P_e ERPs, there were $M = 122.2$ and $M = 122.8$ for correct and incorrect in the pre-test; $M = 144.6$ and $M = 98.4$ for Training 1; $M = 344.3$ and $M = 159.1$ for Training 2 and 3 combined (see Section 3). The statistical significance of observed differences in the electrophysiological data was assessed using a clustering and randomization test (Maris, 2004; Maris and Oostenveld, 2007). As it is used here, for the average potential in a time window, the clustering and randomization test first computes a contrast statistic between conditions which is thresholded and clustered for observations in adjacent electrodes. A cluster-level statistic (sum of t-statistic) is computed for the samples in this joint set. The maximum is taken from this set, and a p-value is calculated using Monte Carlo resampling in the randomization test. The contrast values for the ERP measures were taken as the average effect over the electrodes within the significant clusters.

3. RESULTS

3.1. GRAMMATICAL CLASSIFICATION

Figure 2 shows the average classification (d') for each phase. During the pre-test, classification was near chance ($b = 0.0834$, $t = 0.5661$, HPD = −0.2057, 0.3713), but improved during the first block of the training for the trials that included the declension violations and controls (TRN1, $b = 0.6434$, $t = 2.2781$, HPD = 0.0891, 1.2116). There was no statistical evidence for improvement with the gender trials in the first block of training (TRN1, $b = 0.2068$, $t = 1.0267$, HPD = −0.2057, 0.3713, interval includes zero). Over the entire three training blocks (TRN1–3), classification improved for the gender trials ($b = 0.5098$, $t = 4.112$, HPD = 0.2693, 0.7647), but improved more for the declension trials ($b = 0.3851$, $t = 2.210$, HPD = 0.0290, 0.7309). In the follow-up phase, classification was better than during the pre-test ($b = 0.7225$, $t = 2.4194$, HPD = 0.1386, 1.3277), but lower than during the final block of training ($b = -0.8490$, $t = -2.382$, HPD = −1.5828, −0.1236). This did not depend on whether the follow-up trials were part of the declension or gender contrast ($b = -0.1084$, $t = -0.213$, HPD = −1.7538, 0.3112, interval includes zero). Note that during the pre-test and follow-up phases, there was no feedback on performance.

3.2. VIOLATION-RELATED EVOKED ACTIVITY

Figure 3 shows the isovoltage topographical distribution of the average difference between violation and control conditions for the gender contrast in the P600 time window (500–900 ms) at the CW noun in the four-word version of the phrases. A P600 effect was present on posterior electrodes in Training 3, but not in Training 1, or the Pre-test (see **Table 3**). The P600 effect was marginally significant in Training 2 for this condition. **Figure 4** shows the corresponding distribution for the declension contrast at the CW adjective, in the four-word version. A P600 effect on posterior electrodes was present in all Training sessions, but not the Pre-test (**Table 3**). The trace plots of the violation and control conditions at electrode Cz/1 for both the gender and declension contrasts indicate that the difference was primarily in the 500- to 900-ms time window. There were no significant differences in the three-word versions of the phrases, in any phase (see **Figures 5** and **6**).

3.3. ERROR-RELATED EVOKED ACTIVITY

3.3.1. *Feedback-related activity*

Figure 7 shows the feedback-locked average isovoltage contours for the difference between error- and correct-response trials in the feedback-N_e and -P_e time windows. During Training 1 and 2 there was an N_e effect, as indicated by the negative difference over centro-frontal electrodes and a positive difference on peripheral electrodes (see **Table 4**). In Training 3 the amplitude of this difference was reduced and the statistical effect was no longer present. In the P_e time window, there was a significant positive difference in Training 1, and in Training 2 and 3 this difference became larger, on a similar set of electrodes as in Training 1. Finally, a correlation analysis of the ERP effects did not reveal any statistically significant relationships (positive or negative) between the P600 and error-related components.

3.3.2. *Response-related activity*

The difference between error- and correct-response trials in the N_e and P_e time windows in the time interval near the classification response showed little evidence of a response-related N_e (time window −150 to 150 ms) or response-related P_e (time window 150–300 ms). There were no significant N_e or P_e effects during the Pre-test or any of the Training sessions. When Training 2 and 3 were pooled to increase statistical power, there was some evidence of a small centro-frontal negativity (−0.44 μV, Sum-$t = -24.68$, $p = 0.019$, on eight electrodes: 2–3, 6–8, 10–12, and 17), shown in **Figure 8**. However, this negativity was sustained into the later P_e time window (−0.54 μV, Sum-$t = -14.53$, $p = 0.037$, on five electrodes: 2–3, 10–11, and 22).

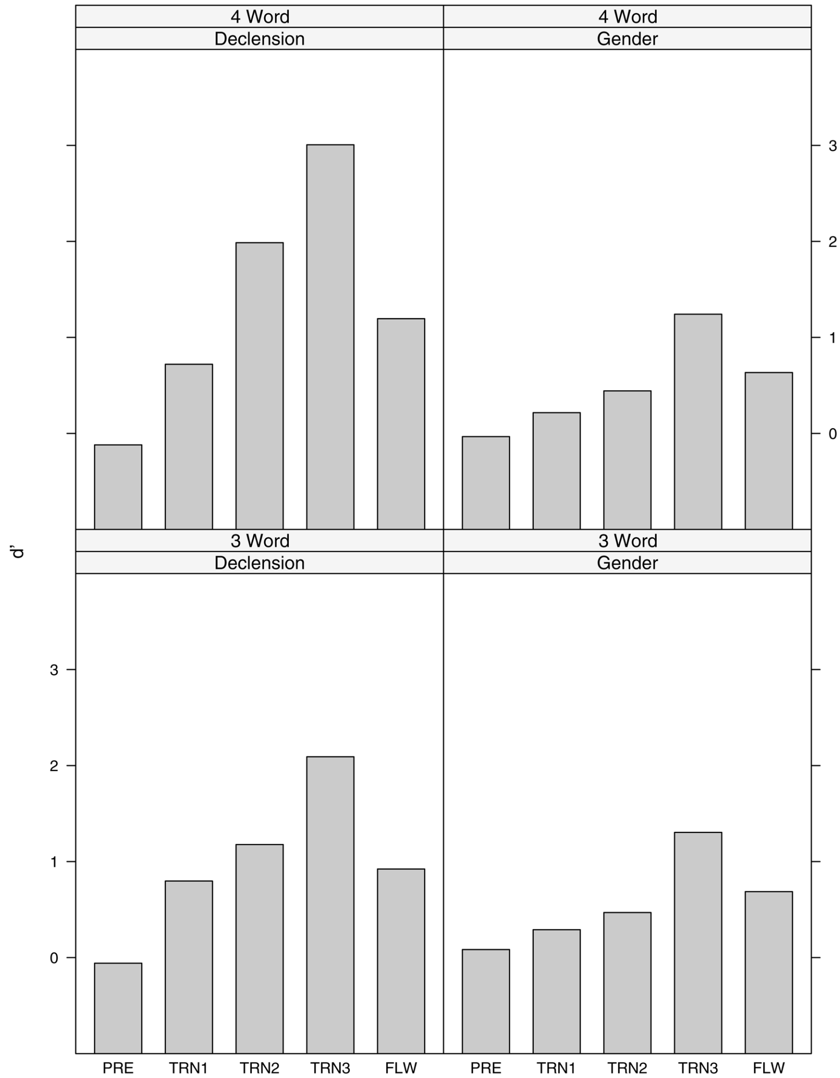


FIGURE 2 | Average classification performance (d') over pre-test, training (1–3), and follow-up blocks for the three- and four-word phrases, for the declension and gender contrasts.

3.4. RELATIONSHIP BETWEEN PERFORMANCE AND ERP ACTIVITY

To investigate the relationship between EEG activity and the relatively short-term behavioral changes during training, the change in discrimination performance (average d') was modeled as a function of Training session (1–3), the ERP difference wave amplitude for the violation- and feedback-error components in sessions 1–3, as well as the individual subject loadings of the six principle components summarizing the individual difference measures (see Section 2.5). Recall that the principle components were most highly loaded, respectively, on the factors (1) comfort using German, (2) Goethe Institute specific test, (3) whether participants like using German, (4) Wisconsin card sort number of errors, (5) *n*-back task slope, and (6) Goethe Institute global test. Only in the case of the *n*-back loading (PC5) was a single factor clearly related to the loading. In all other cases, multiple factors were related. The violation components included the P600 effect amplitudes for the (four-word) gender and declension violation effects, and the error-related components included the feedback N_e and P_e effects, in all cases for each of the Training sessions 1–3. The interaction of session with each of the ERP effects and each of the principle components were included as predictors for the classification response. Each of the predictors was scaled to a mean of zero and unit SD.

The regression indicated a significant interaction of feedback-P_e effect magnitude with session such that participants with a larger P_e effect improved more over the sessions (e.g., the increasing slope in **Figure 9**, $b = 0.1498$, $t = 2.6149$, HPD = 0.0453, 0.2793). Please note that the first training session took place 1 week before sessions 2 and 3. None of the other ERP measures predicted performance alone, or in interaction with session. In particular, N_e effect magnitude did not predict performance, in contrast to the results reported in Davidson and Indefrey (2009).

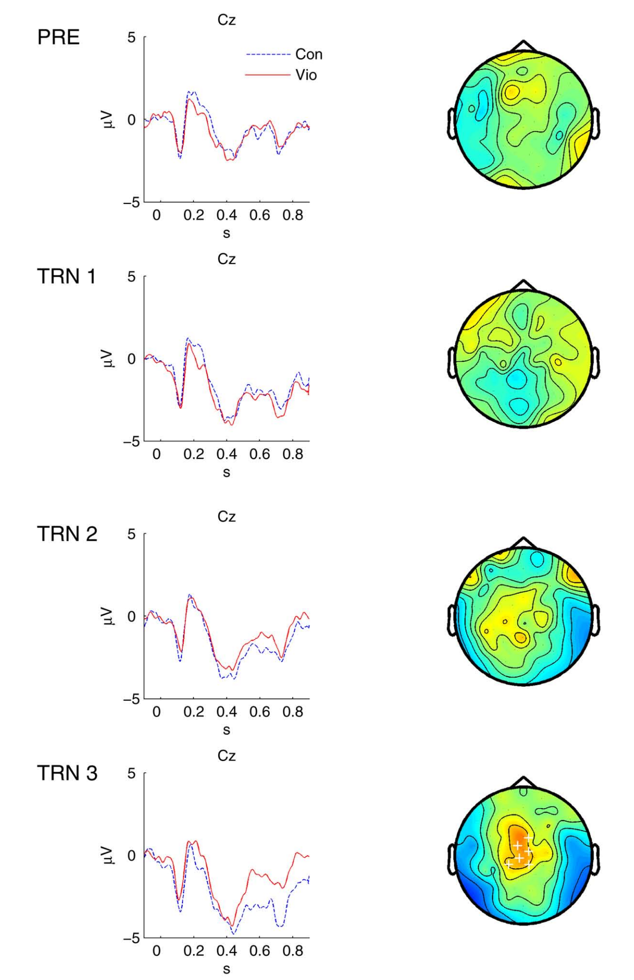


FIGURE 3 | Average event-related potentials for gender violation contrast (4 word). Trace plots indicate average voltage (in μV) as a function of time at Cz/1; control = Con (blue, dotted line), violation = Vio (red, solid line). Symbols plotted on the topographical plots indicate membership within a statistically significant negative ("−") or positive ("+") cluster of electrodes. Topographical plots are viewed from the top (left is on the left-hand side).

There were two other interactions of session with principle components of the auxiliary measures: Session*PC3, $b = -0.1387$, $t = -2.2594$, HPD = −0.2796, −0.0081; and Session*PC5, $b = 0.1744$, $t = 2.5839$, HPD = 0.0345, 0.3353. PC3 was positively related to valence toward German (the scale like–dislike learning German) as well as the number of languages reported to be known by the subjects, but negatively related to hours of recent exposure to German. PC5 was related most strongly to n-back performance, and it was not strongly related to other variables.

To investigate discrimination forgetting (or performance decline), the difference in discrimination from the last phase of training (Training 3) to the follow-up phase was modeled as a function of the EEG and principle component measures that were significant predictors in the Training analysis. In this analysis, P_e magnitude positively predicted discrimination in Training 3 ($b = 0.6233$, $t = 4.7330$, HPD = 0.3553, 0.9017), as was

Table 3 | Statistics for violation-related EEG activity.

Phase	Type	Ave	Sum-*t*	*p*	Electrodes
Training 1	Decl	+1.53	38.02	0.0013	4, 6, 12–18, 25–29 (14)
	Decl	−1.68	−26.50	0.0131	45–46, 53–58 (8)
	Gen	+0.94	–	–	–
	Gen	−0.76	–	–	–
Training 2	Decl	+1.93	36.10	0.0020	1, 3–6, 10, 12–17, 28 (12)
	Decl	−1.43	−24.79	0.0116	33, 46–49, 58, 61, 63 (8)
	Gen	+0.87	9.10	0.0609	14–17 (4)
	Gen	−1.31	−8.71	0.0662	39–40, 53–54 (4)
Training 3	Decl	+2.09	16.32	0.0218	3–5, 11–13, 25 (7)
	Decl	−2.27	−40.11	0.0001	41, 48, 53–60, 63 (11)
	Gen	+1.56	17.26	0.0307	1–2, 5–7, 16 (6)
	Gen	−1.40	−13.02	0.0593	39–40, 52–54 (5)

*Average difference (Ave, in μV) calculated as the average within the P600 interval (500–900 ms) after the CW onset for the two violation contrasts (Type). The summary statistics (Sum-t) and p-values (p) for the clustering and randomization tests are provided, along with the approximate electrode locations according to **Figure 1**, as well as the number of electrodes in the cluster.*

shown in the previous analysis, and there was a negative relationship between P_e magnitude and discrimination in the follow-up ($b = -0.4852$, $t = -3.1530$, HPD = −0.9551, −0.0316). This result indicates that P_e magnitude predicted which participants gained during training, but also which participants lost discrimination 3 months later. In this analysis, the principle components which were significant predictors in the Training sessions, were not predictors for the follow-up loss. In summary, the P_e effect amplitude was a consistent predictor of discrimination gain (and loss, over the longer term), even when adjusting for individual differences in a variety of performance tasks.

4. DISCUSSION

We expected violation- and error-related ERP effects to appear in concert with the learners' discrimination improvement. The experiment reported here provided evidence for several of these effects: As participants' grammatical discrimination improved over time, P600 responses to grammatical violations emerged, error-related activity was observed in response to feedback, and the amplitude of one of the feedback-related responses was related to improved grammatical discrimination. Finally, the magnitude of the error-related activity predicted later retention of the discrimination ability.

The observed P600 responses to violations of the syntactically determined declension of German adjectives replicated findings of an earlier study (Davidson and Indefrey, 2009) in which participants had been provided with explicit instruction on the rules of German adjective declension. In the present study, no explicit instruction was given and the learners had to rely on positive and negative feedback for the learning of adjective declension rules. The behavioral data showed a gradual increase from grammaticality judgment performance near-chance level in the pre-test to high performance in the range of a native speaker control group participating in the previous study (the range of the hit rate – false alarm rate measure in native speakers was approximately 0.6–0.9,

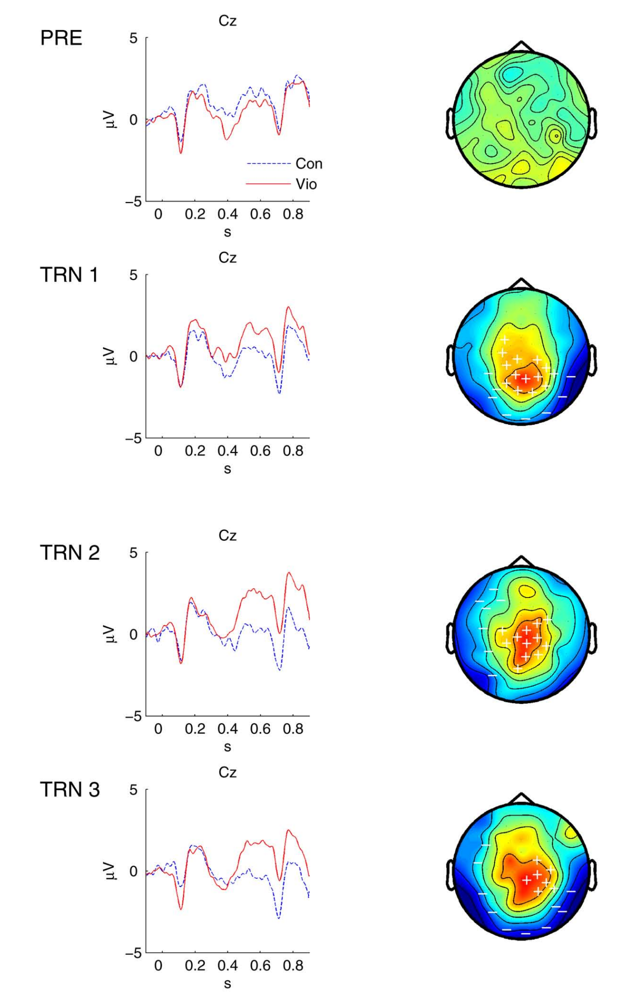


FIGURE 4 | Average event-related potentials for declension violation contrast (4 word). Trace plots indicate average voltage (in μV) as a function of time at Cz/1; control = Con, violation = Vio. Symbols plotted on the topographical plots indicate membership within a statistically significant negative ("−") or positive ("+") cluster of electrodes. Topographical plots are viewed from the top (left is on the left-hand side).

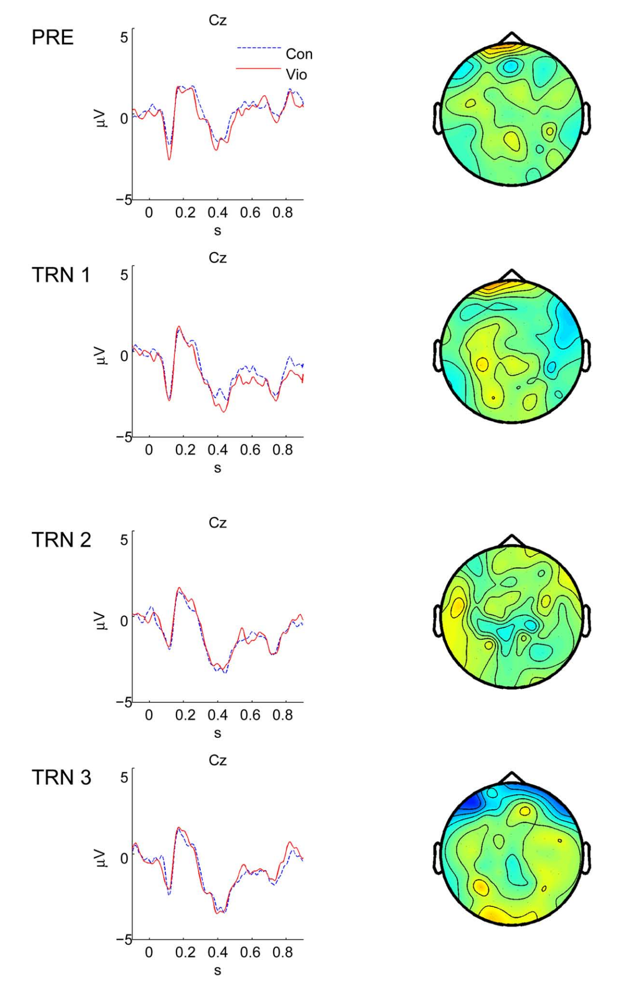


FIGURE 5 | Average event-related potentials for gender violation contrast (3 word). Trace plots indicate average voltage (in μV) as a function of time at Cz/1; control = Con (blue, dotted line), violation = Vio (red, solid line). There were no statistically significant clusters. Topographical plots are viewed from the top (left is on the left-hand side).

see Figure 2 of Davidson and Indefrey, 2009). In the previous study with explicit instruction, there was a steep performance increase in the first training session compared to the pre-test. Despite the absence of explicit instruction in the present study, classification performance also improved rapidly and there was a significant P600 responses to declension violations already in the first training session. This gained in strength over the following training sessions. These findings indicate that the changes in the neural responses are related to the acquired grammatical knowledge *per se* (i.e., adjective declension rules) rather than how this knowledge is acquired, i.e., by receiving explicit rule instruction or by finding the rules themselves.

Like Davidson and Indefrey (2009), we only found declension violation responses in the four word prepositional phrases (e.g., *mit dem *kleinem Kind*) where the violating adjective carries a strong inflectional suffix redundantly specifying case, number, and gender information but not in three word prepositional phrases (e.g., *mit *kleinen Kind*) where the violating adjective carries a default suffix that does not specify syntactic feature information. Although both types of incorrect prepositional phrases violate the syntactic rule according to which case, number, and gender features must be specified on the first (and only on the first) inflectable element of the noun phrase, the neural violation response thus seems to hinge on the presence of (incorrect) positively specified syntactic features. Hierarchical feature specification analyses of the German adjective paradigm predict a differential response to strong forms positively specifying syntactic features and weak default forms and hence are supported by our data (see Clahsen et al., 2001; Penke et al., 2004 for corresponding psycholinguistic evidence).

In contrast to the experiment reported in Davidson and Indefrey (2009), which observed P600 responses to grammatical gender violations in native speakers but not in Dutch learners of German,

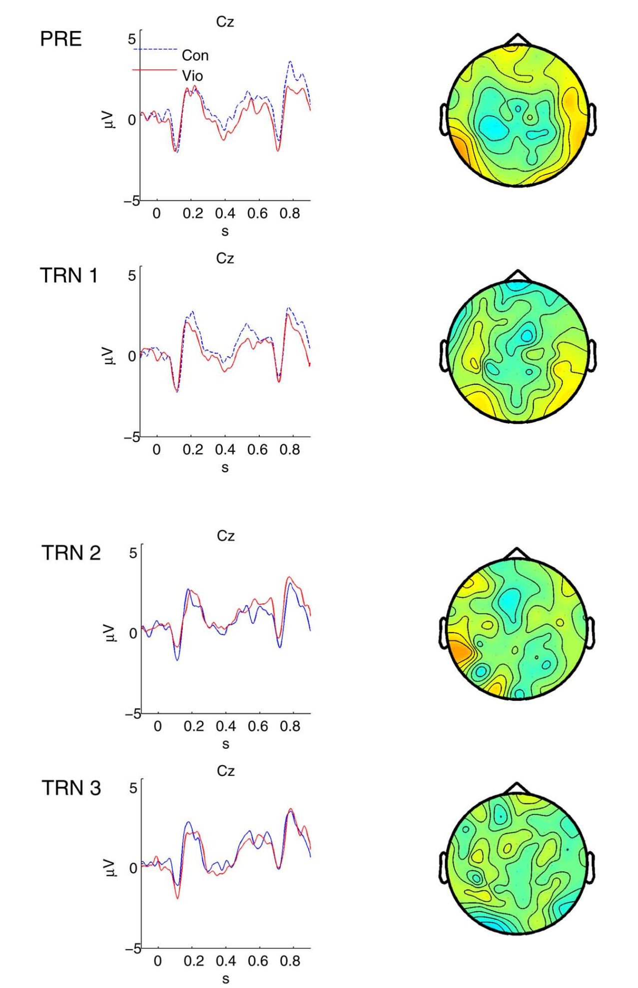


FIGURE 6 | Average event-related potentials for declension violation contrast (3 word). Trace plots indicate average voltage (in µV) as a function of time at Cz/1; control = Con, violation = Vio. There were no statistically significant clusters. Topographical plots are viewed from the top (left is on the left-hand side).

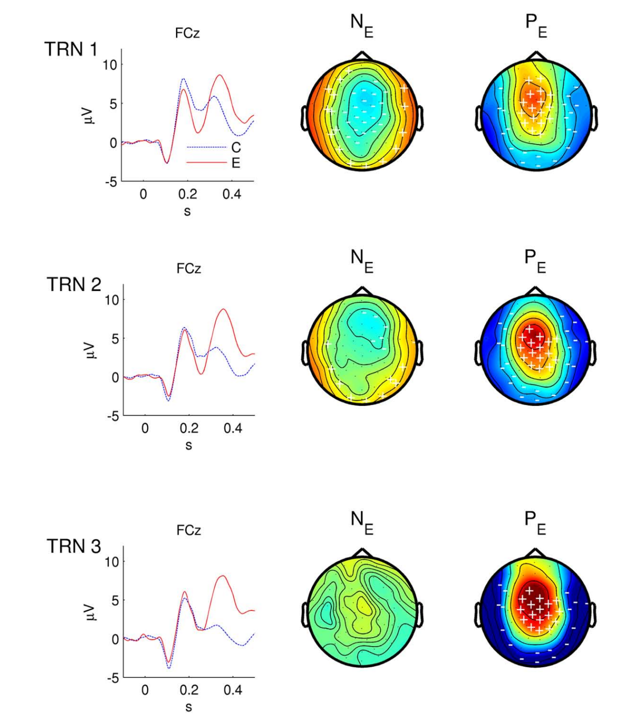


FIGURE 7 | Average event-related potentials for feedback-related activity. Trace plots indicate average voltage (in µV) over time at FCz/8; correct classification = C (blue, dotted line), error classification = E (red, solid line). Symbols plotted on the topographical plots indicate membership within a statistically significant negative cluster ("−") or positive cluster ("+") of electrodes.

in the present study we found significant P600 gender violation responses in the learners in the last training session. The observation that P600 violation effects can be observed to gender violations is in line with other studies reporting P600 responses to gender violations in a second language (Sabourin, 2003; Tokowicz and MacWhinney, 2005; Sabourin et al., 2006). One possible reason for the absence of a significant gender violation response in Davidson and Indefrey (2009) could be the fact that there was only one training session, providing no opportunity for a relatively late emergence of a response as in the present study. As indicated in Davidson and Indefrey (2009), learning to retrieve and apply grammatical gender information may be a more difficult learning task than learning to apply a declension rule because the gender category must be associatively linked to every single noun. Even though we chose the nouns such that their Dutch translation equivalents also fell into two different gender classes, the Dutch participants may have been insecure about the German gender (masculine or feminine) of the nouns whose Dutch translation equivalents belong to the common gender. A suggested by an anonymous reviewer, the Dutch participants might also have had some residual knowledge of German nominative determiner forms and were initially confused by the dative forms used in our experiment. Taken together, there is reason to assume that learning the gender class of German nouns may require more training, teaching, exposure, or usage. In Davidson and Indefrey (2009), there was one training phase with feedback, whereas in the present experiment there were three phases with feedback. See Blom et al. (2008) and Sabourin and Stowe (2008) for recent discussions of L2 gender learning and factors which contribute to learning variability.

Another important factor in the rate of grammatical learning might be the size of the set of lexical items used in the learning task. In the present experiment, a relatively small number of adjectives and nouns were used, in order to reduce the chance that participants would not know the meanings of the words. However, as suggested by an anonymous reviewer, the diversity of the item set may play an important role in modulating the appearance of the violation effects. At one extreme, if only a few carrier phrases are used, then the rate of learning might be relatively fast because the repetition of items would highlight morphosyntactic patterns (and/or reduce the lexical knowledge burden on learners). At the other extreme, if each trial had different carrier items (at both

Table 4 | Statistics for feedback error-related EEG activity (Rsp) in the three training phases.

Phase	Rsp	Win	Ave	Sum-*t*	*p*	Electrodes
Training 1	N_e	100–300	+1.17	62.33	0.0073	46–64 (18)
	N_e	100–300	−0.85	−72.20	0.0018	1–17, 19, 21–22, 28–29 (22)
	P_e	300–500	+1.49	49.37	0.0002	1–9, 14–20, 34 (17)
	P_e	300–500	−1.44	−73.56	0.0001	23–24, 38–40, 43–44, 46–47, 51–64 (22)
Training 2	N_e	100–300	+0.84	29.24	0.0144	41, 53–60, 63 (10)
	N_e	100–300	−0.79	−26.58	0.0185	3, 8–11, 20–22 (8)
	P_e	300–500	+2.18	86.35	0.0001	1–10, 13–19 (18)
	P_e	300–500	−1.86	−87.22	0.0001	37–40, 43–49, 51–64 (24)
Training 3	N_e	100–300	+0.09	–	–	–
	N_e	100–300	−0.29	–	–	–
	P_e	300–500	+3.79	105.22	0.0001	1–19 (19)
	P_e	300–500	−3.12	−114.02	0.0001	38–39, 42–43, 47–48, 52–61, 63–64 (18)

*Average difference (Ave, in μV) calculated as the mean value within a time interval (Win, in ms) after the feedback onset. The summary statistics (Sum-t) and p-values (p) for the clustering and randomization tests, along with the approximate electrode locations are listed according to **Figure 1,** along with the number of electrodes in the cluster.*

adjective and noun positions), then it may take a substantially higher number of trials for participants to successfully apply their grammatical knowledge, especially if it is likely that they do not know the meanings of all the words. This factor, the diversity of the item set, would be expected to have more of an impact for the gender contrast than the declension contrast, for the reasons outlined earlier. Also, in future studies it would be better to test for generalization by including new items that were not seen in the training set in an explicit test of generalization. There are potentially other explanations of the declension–gender difference seen here, but perhaps it would be advisable to find ERP evidence using a wider variety of items before elaborating different predictions.

In addition to using a small lexicon, we also simplified the German determiner and adjective declension system by only using dative case and avoiding syncretism. Our results suggest that the full system most likely would not have been learned in the available number of sessions. These simplifications, however, do not mean that our participants merely learned to associate particular items or item combinations with particular responses. Firstly, just because a relatively small set of adjectives were repeated many times in all possible forms, the identification of a particular adjective stem did not provide any cue as to its appropriate ending. Secondly, as can be seen in **Table 1**, due to the presence of both gender and declension violations our stimulus set contained an equal number of trials in which a specific adjective form with the same preceding context (e.g., mit kleinem) required a correct response (as in mit kleinem Kind) and an incorrect response (as in mit kleinem *Frau). Given that our feedback did not distinguish between incorrect responses due to declension or gender errors, participants could not learn to base their declension class decision on a particular combination of word forms such as mit kleinem. For the same reason, correct gender agreement could not be learned based on simple associations between particular determiner–noun or adjective–noun combinations and constant response requirements. Taken together, these properties of our stimulus set mean that in order to respond correctly at the observed performance level our participants had to learn the grammatical rules of the reduced system.

There have been relatively few previous EEG studies of practice-related improvement with similar tasks. A natural question is whether the P600 responses that were observed in the present study might be more generally related to practice-related improvement. In a non-linguistic domain, both Romero et al. (2008) and Pauli et al. (1994) found that practice on tasks requiring mathematical knowledge reduces the amplitude of frontal positive potentials. Both studies found that a non-selective frontal P300 component was attenuated with practice, while Romero et al. also found that a later posterior P500 component selective for correct equations than incorrect equations became larger after practice. In contrast, Pauli et al. (1994) found that the posterior positive potential was not attenuated with practice. Romero et al. attributed the difference to the fact that their experiment used a task involving the verification of alphabet–arithmetic equations, which were likely to be unfamiliar to participants before practice, while the Pauli et al. task involved producing the answers to ordinary multiplication equations, which were likely to be known before practice. If the positive responses are general task-related effects, then the results from the equation-processing experiments would suggest that late positive components seen in the present experiment should have either decreased as a function of practice (Pauli et al., 1994), or become larger for correct-string as compared to incorrect-string stimuli (Romero et al., 2008).

However, these predictions are not in agreement with the results for the P600 component. In the present experiment (also in Davidson and Indefrey, 2009), the P600 violation effect was absent initially, and only appeared after practice. This pattern after practice was similar to the native speaker control group in Davidson and Indefrey (2009). Together, these findings suggest that the emergence of the grammatical violation effect seen in the present study is likely to be related to grammatical processing, rather than

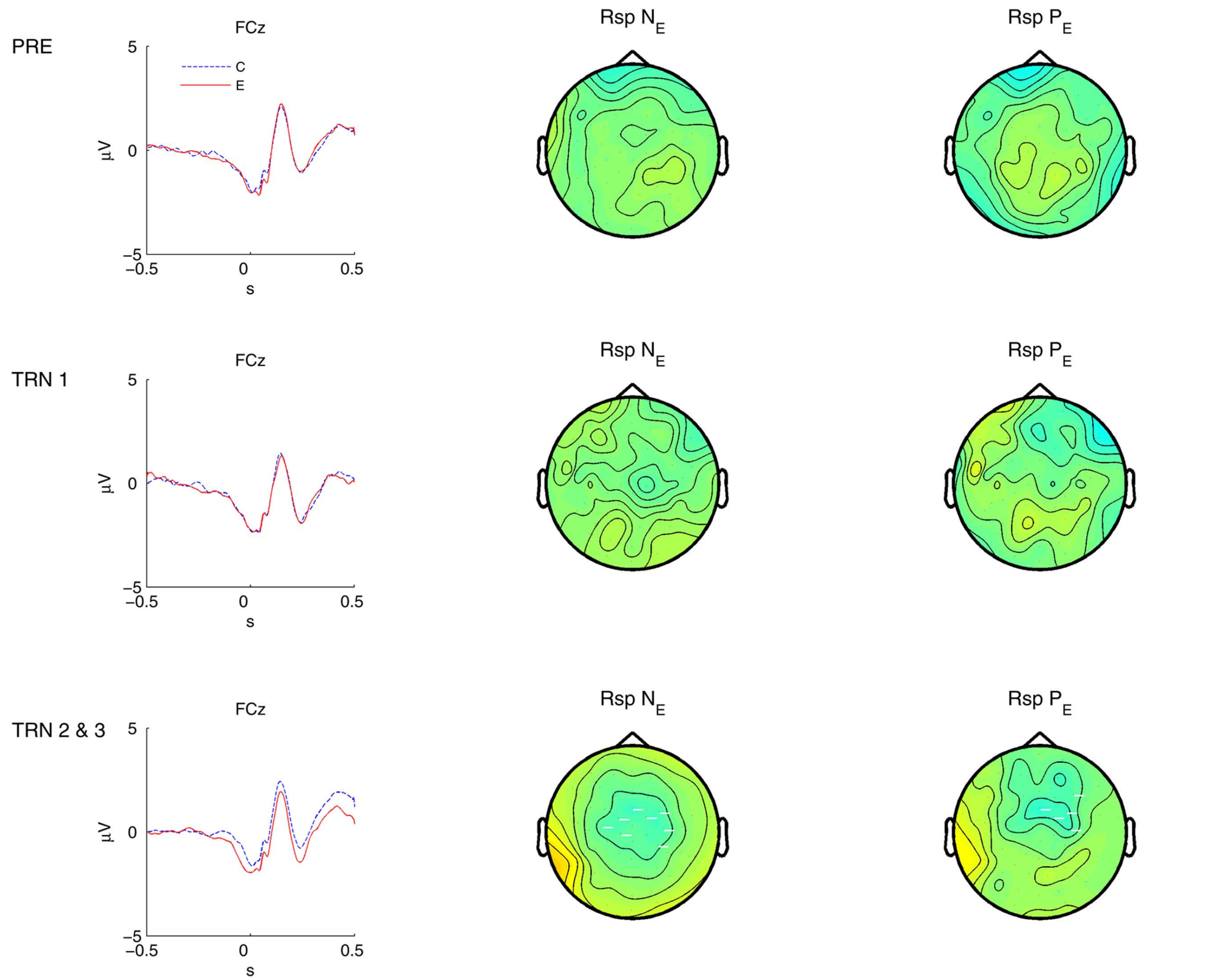


FIGURE 8 | Average event-related potentials for response-related activity. Trace plots indicate average voltage (in μV) over time at FCz/8; correct classification = C (blue, dotted line), error classification = E (red, solid line). Symbols plotted on the topographical plots indicate membership within a statistically significant negative cluster ("−") of electrodes.

a general task-related P300 effect. The changing pattern of the P600 responses seen here is consistent with the results of Davidson and Indefrey (2009), with the exception that a P600 response was observed for gender violations in the present experiment and not the previous study (see Davidson and Indefrey, 2009, as well as Sabourin, 2003; Sabourin et al., 2006).

One of the main aims of the present work was to disentangle the contributions of response- and feedback-related activity as the results in Davidson and Indefrey (2009) were not able to distinguish these, and to test differential predictions for the two error responses, suggesting that the magnitude of the feedback-N_e decreases over time, in concert with an increase in the magnitude of the response-N_e (Holroyd and Coles, 2002). In the present study we found a clear feedback negativity whereas the amplitude of the response-related N_e activity, although statistically significant, was weak and unreliable, suggesting that the error negativity observed in Davidson and Indefrey (2009) was likely due to feedback-related activity.

With respect to changes of feedback and response negativities as a function of time, during learning, our data confirmed the first prediction: The feedback-N_e indeed decreased in magnitude from the first to the last training session. This finding supports Holroyd and Coles (2002) hypothesis that the feedback-N_e reflects a learning process in which the internal state of learners is modified. More specifically, this modification likely involved the learners representation of the declension regularities on which they based their grammaticality decisions. The learners, starting with no or very little knowledge as indicated by their near-chance performance in the pre-test, had to extract the relevant grammatical knowledge from the information provided by the feedback. This means that initially this feedback must have had a relatively large impact on the learners representations as indicated by a performance increase to a high level. Even though at higher performance levels negative feedback arguably constituted a stronger conflict with the participants expectation (they knew the probability of having made a correct decision was higher than at the beginning), the amplitude

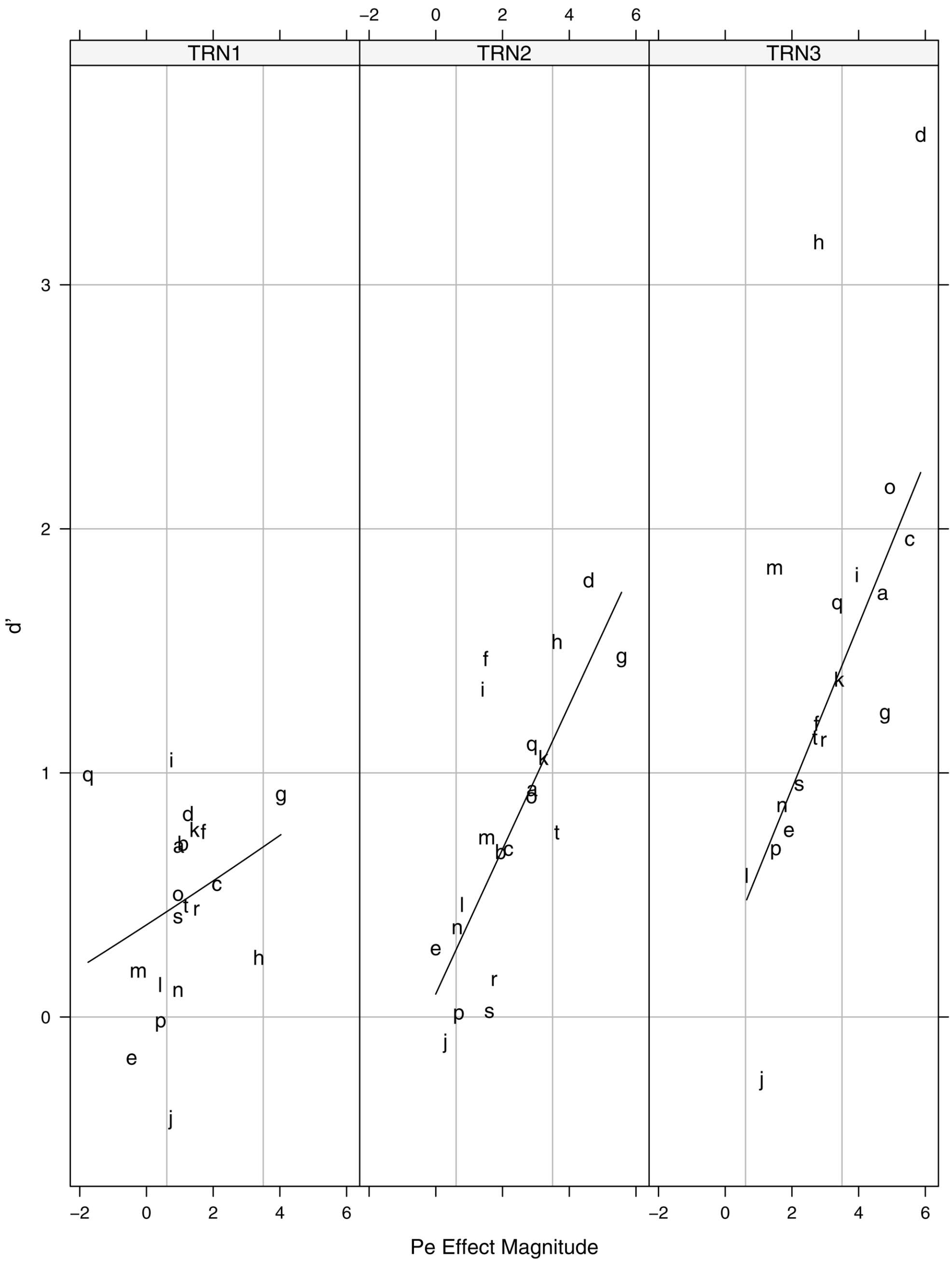


FIGURE 9 | Relationship between feedback-P_e and discrimination performance during Training blocks 1–3 (subjects indicated by letters).

of the ERP response to negative feedback decreased. In line with Holroyd and Coles (2002) prediction this may be interpreted as showing that in spite of negative feedback the learners were less prone to change their internal representations at later stages of training.

Our data do not allow any conclusion with respect to a possible response negativity. A plausible explanation for the weak response-N_e might be the task parameters. Unlike the speeded response time tasks used in previous studies of the response-N_e, participants in our study were not under substantial time pressure to provide their responses. This may have contributed to the relative weakness of the response-N_e effect, and future work investigating the response-N_e in grammatical learning might impose a shorter response deadline to boost the effect.

The second component of the feedback response, the feedback positivity (P_e), contrary to the feedback negativity became larger as a function of training and individual variation in the feedback-P_e amplitude was related to discrimination performance.

As mentioned in the Introduction, the response P_e has been suggested to be a type of P300 reflecting error awareness (Leuthold and Sommer, 1999; Frank et al., 2007). An extension of this functional characterization of the response P_e to the feedback-P_e would be in accordance with our data as the awareness of a conflict between the participants expectation and the actual feedback quite plausibly increased with the participants performance level. The better their performance level, the larger the perceived conflict would be and hence the corresponding feedback positivity.

The changing P600 and N_e/P_e ERP responses in this experiment suggest that the presentation of a series of phrases (along with the feedback) affects the behavioral classification of phrases presented at a later point in time. The relationship between the feedback activity and the violation-related activity appears to be complex, however, for at least two reasons. First, the error-related activity was both changing and multi-phasic. Over the course of training from the first to the second day, the feedback-N_e amplitude decreased while the feedback-P_e amplitude increased. Second, while the P600 amplitude increased with training, it was not itself statistically related to the N_e/P_e activity, possibly due to too much variability in the responses. While this pattern of activity precludes a simple account of the relationship between feedback and discrimination improvement, the results do provide additional support for the claim that feedback-related activity (N_e and/or P_e) can be related to grammatical learning under some circumstances. Nevertheless, given these findings, future work might employ experimental designs which are better optimized to estimate the P600 and N_e/P_e relationship, perhaps by focusing on a single type of violation with more trials and more training. The regression results also indicate that valence toward learning German or languages generally (like–dislike scale, number of languages known), as well as working memory span may be important modulating variables. Given the sample size of the present experiment, perhaps behavioral experiments on learning, which can be run with substantially larger sample sizes, could better elucidate whether these factors strongly modulate learning.

While error-related activity was related to discrimination improvement like the previous study, one notable exception was that the present study did not show a direct relationship between N_e amplitude and discrimination performance. The main differences in design were the absence of explicit instruction in the present study, and the temporal separation of the classification response from the feedback. In addition, feedback was presented on both days of training in the present study, but only during the first day of the previous study. It was hypothesized that slower learning would slow the evolution of the error-related activity over a longer time scale, but the absence of the instruction may have altered the task dynamics in such a way to make the experiments less than fully comparable. As expected, performance on the present experiment improved more slowly than in Davidson and Indefrey (2009), most likely because participants in the present study had to determine how to classify the phrases by trial and error, rather than by relying on their memory for the explicitly provided rules.

It may be that in this task, the feedback-N_e reflects recognition that the current hypothesis about grammatical classification needs to be changed. With little initial knowledge (in the present experiment), the large N_e may reflect a new, updated hypothesis, but this new hypothesis may not have been correct. Participants who updated to a better hypothesis early would in fact have shown smaller N_e subsequently. Those who changed several times before they got it right might have shown in total larger N_e responses. This would explain the present data well, but in turn raises the question about the relationship found in the previous study. In the previous study, the instruction may have made it more likely that the initial change in hypothesis was effective, because it could have strengthened the memory of the instructions. Although speculative, this account of the N_e/P_e contribution to the effects observed here could be investigated in future experiments by including a variable that would affect participants' ability to apply rules, or the number of rules to be applied. The results also suggest independent roles for the feedback-N_e and feedback-P_e effects, which have not been extensively investigated previously (see also Overbeek et al., 2005).

Finally, the results suggest that future models of grammatical plasticity should include not only an account of the learning of grammatical knowledge, but also an account of how grammatical knowledge is lost, or otherwise made unavailable after a period of disuse. The present results, along with several other recent findings (Mueller et al., 2005; Osterhout et al., 2005, 2006) suggest that the learning of grammatical knowledge can occur quite rapidly in adults, at least when acquired explicitly. The learning of this knowledge does not imply that it is stable, however. Without maintenance or usage to reinforce learning, adult grammar knowledge appears to be vulnerable to decay or interference. Future work might investigate whether the dominant factor(s) determining the effects of the hypothesized sensitive period in adult grammar learning are more related to retention than learning.

5. CONCLUSION

The experiment described here has shown that there are several electrophysiological correlates of learning in grammar-learning tasks with feedback. The results showed that these ERP measures are *dynamic*, in the sense that they can change within the span of one or two experimental sessions, at least with Dutch participants learning German as studied here. The results were largely congruent with the pattern of data reported in Davidson and Indefrey (2009), despite the absence of instructions in the present experiment. The response- and feedback-ERP results can be taken as evidence that cognitive control mechanisms function during explicit learning to help modify the knowledge state of second language learners, and/or enable the memory of this knowledge so that it can be put to use during real-time language comprehension.

ACKNOWLEDGMENTS

This work was supported by the Dutch science foundation Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) and the German science foundation Max Planck Gesellschaft (MPG). Esther Meeuwissen and Daniel von Rhein recruited participants, helped create the stimuli, and recorded the EEG with assistance from the first author.

REFERENCES

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412.

Bagiella, E., Sloan, R. P., and Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology* 37, 13–20.

Bierwisch, M. (1967). "Syntactic features in morphology: general problems of so-called pronominal inflection in German," in *To Honor Roman Jakobson* (The Hague: Mouton), 239–270.

Blevins, J. P. (1995). Syncretism and paradigmatic opposition. *Linguist. Philos.* 18, 113–152.

Blevins, J. P. (2003). Stems and paradigms. *Language*, 79, 737–767.

Blom, E., Polisenská, D., and Unsworth, S. (2008). The acquisition of grammatical gender in Dutch. *Second Lang. Res.* 24, 259–265.

Cahill, L., and Gazdar, G. (1997). The inflectional phonology of German adjectives, determiners, and pronouns. *Linguistics* 35, 211–245.

Clahsen, H., Sonnenstuhl, I., Hadler, M., and Eisenbeiss, S. (2001). Morphological paradigms in language processing and language disorders. *Trans. Philol. Soc.* 99, 247–277.

Davidson, D. J., and Indefrey, P. (2009). An ERP study on changes of violation and error responses during morphosyntactic learning. *J. Cogn. Neurosci.* 21, 433–446.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components 2. Error processing in choice reaction tasks. *Electroencephelogr. Clin. Neurophysiol.* 78, 447–455.

Ferree, T. C., Luu, P., Russell, G. S., and Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clin. Neurophysiol.* 112, 536–544.

Frank, M. J., D'Lauro, C., and Curran, T. (2007). Cross-task individual differences in error processing: neural, electrophysiological, and genetic components. *Cogn. Affect. Behav. Neurosci.* 7, 297–308.

Friederici, A. D., Steinhauer, K., and Pfeifer, E. (2002). Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 529–534.

Friston, K. J., Stephan, K. E., Lund, T. E., Morcom, A., and Kiebel, S. (2005). Mixed-effects and fMRI studies. *Neuroimage* 24, 244–252.

Ganushchak, L. Y., and Schiller, N. O. (2009). Speaking one's second language under time pressure: an ERP study on verbal self-monitoring in German–Dutch bilinguals. *Psychophysiology* 46, 410–419.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error-detection and compensation. *Psychol. Sci.* 30, 306–315.

Goethe-Institut. (2005). *Placement Test German.* Available at: http://www.goethe.de [retrieved March 5, 2005].

Hagoort, P., Brown, C., and Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP-measure of syntactic processing. *Lang. Cogn. Process.* 8, 439–483.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *J. Psycholinguist. Res.* 30, 251–266.

Hahne, A., and Friederici, A. D. (2001). Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Biling. (Camb. Engl.)* 4, 123–141.

Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.

Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* 20, 60–99.

Johnson, J. S., and Newport, E. L. (1991). Critical period effects on universal properties of language: the status of subjacency in the acquisition of a second language. *Cognition* 39, 215–258.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., and Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10155–10162.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain Lang.* 109, 68–74.

Kutas, M., Van Petten, C., and Kluender, R. (2006). "Psycholinguistics electrified II: 1995–2005," in *Handbook of Psycholinguistics*, eds M. Traxler and M. Gernsbacher (Amsterdam: Academic Press), 659–724.

Lenneberg, E. H. (1967). *Biological Foundations of Language.* New York: Wiley.

Leuthold, H., and Sommer, W. (1999). ERP correlates of error processing in spatial S-R compatibility tasks. *Clin. Neurophysiol.* 110, 342–357.

Maris, E. (2004). Randomization tests for ERP topographies and whole spatio-temporal data matrices. *Psychophysiology* 41, 142–151.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190.

Morgan-Short, K., Sanz, C., Steinhauer, K., and Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: an event-related potential study. *Lang. Learn.* 60, 154–193.

Mueller, J. L., Hahne, A., Fujii, Y., and Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *J. Cogn. Neurosci.* 17, 1229–1244.

Newport, E. L., Bavelier, D., and Neville, H. J. (2001). "Critical thinking about critical periods: perspectives on a critical period for language acquisition," in *Language, Brain, and Development*, ed. E. Dupoux (Cambridge, MA: MIT Press), 481–502.

Nunez, P. L., and Srinivasan, R. (2006). *Electric Fields of the Brain: The Neurophysics of EEG.* New York, NY: Oxford University Press.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011.

Opitz, B., Ferdinand, N. K., and Mecklinger, A. (2011). Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Front. Hum. Neurosci.* 5:8. doi: 10.3389/fnhum.2011.00008

Osterhout, L., and Holcomb, P. J. (1995). Event-related brain potentials elicited by syntactic anomaly. *J. Mem. Lang.* 31, 785–806.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., and Inoue, K. (2005). "Sentences in the brain: real-time reflections of sentence comprehension and language learning," in *The On-line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond*, eds M. Carreiras and J. C. Clifton (New York, NY: Psychology Press), 271–308.

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., and Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Lang. Learn.* 56, 199–230.

Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., and Hirschensohn, J. (2008). Second language learning and changes in the brain. *J. Neurolinguistics* 21, 509–521.

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, K. R. (2005). Dissociable components of error processing: on the functional significance of the vis-á-vis the ERN/N_e. *J. Psychophysiol.* 19, 319–329.

Owen, A. M., McMillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59.

Pauli, P., Lutzenberger, W., Rau, H., Birbaumer, H., Rickard, T. C., Yaroush, R. A., and Bourne, L. E. (1994). Brain potentials during mental arithmetric: effects of extensive practice and problem difficulty. *Brain Res. Cogn. Brain Res.* 2, 21–29.

Penke, M., Janssen, U., and Eisenbeiss, S. (2004). Psycholinguistic evidence for the underspecification of morphosyntactic features. *Brain Lang.* 90, 423–433.

Pinheiro, J. C., and Bates, D. M. (2000). *Mixed-effects Models in S and S-PLUS.* New York: Springer-Verlag.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004a). The role of the medial frontal cortex in cognitive control. *Science* 306, 443.

Ridderinkhof, K. R., Van Den Wildenberg, W. P. M., Segalowitz, S. J., and Carter, C. S. (2004b). Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* 56, 129–140.

Romero, S. G., McFarland, D. J., Faust, R., Farrell, L., and Cacace, A. T. (2008). Electrophysiological markers of skill-related neuroplasticity. *Biol. Psychol.* 78, 221–230.

Rossi, S., Gugler, M. F., Friederici, A., and Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: evidence from event-related potentials. *J. Cogn. Neurosci.* 18, 2030–2048.

Sabourin, L., and Stowe, L. A. (2008). Second language processing: when are first and second languages processed similarly? *Second Lang. Res.* 24, 397–430.

Sabourin, L., Stowe, L. A., and de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Lang. Res.* 22, 1–19.

Sabourin, L. L. (2003). *Grammatical Gender and Second Language Processing: an ERP Study*. Unpublished PhD, Rijksuniversiteit Groningen, Groningen.

Schlenker, P. (1999). La flexion de lB'adjectif en allemand: la morphologie de haut en bas [The inflection of adjectives in German: Morphology from top to bottom]. *Recherches linguistiques de Vincennes*, 28, 115–132.

Sebastian-Gallés, N., Rodríguez-Fornells, A., de Diego-Balaguer, R., and Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *J. Cogn. Neurosci.* 18, 1277–1291.

Tokowicz, N., and MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potentials study. *Stud. Second Lang. Acquisition* 27, 173–204.

van der Helden, J., Boksem, M. A. S., and Blom, J. H. G. (2010). The importance of failure: feedback-related negativity predicts motor learning efficiency. *Cereb. Cortex* 20, 1596–1603.

Weber-Fox, C., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behvioral evidence in bilingual speakers. *J. Cogn. Neurosci.* 8, 231–256.

Wunderlich, D. (1997). "Der unterspezifizierte Artikel. [The underspecified article]," in *Sprache im Fokus*, M. Schwarz, K.-H. Ramers and C. Düjrscheid (Tübingen: Niemeyer), 47–55.

Zwicky, A. M. (1986). German adjective agreement in GPSG. *Linguistics* 24, 957–990.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 March 2011; paper pending published: 16 May 2011; accepted: 21 August 2011; published online: 20 September 2011.

*Citation: Davidson DJ and Indefrey P (2011) Error-related activity and correlates of grammatical plasticity. Front. Psychology **2**:219. doi: 10.3389/fpsyg.2011.00219*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 25 October 2011
doi: 10.3389/fpsyg.2011.00284

Grammatical gender inhibition in bilinguals

Luis Morales*, Daniela Paolieri and Teresa Bajo

Department of Experimental Psychology and Behavioural Physiology, University of Granada, Granada, Spain

Edited by:
Judith F. Kroll, Penn State University, USA

Reviewed by:
Laurie Ann Stowe, University of Groningen, Netherlands
Wido La Heij, Leiden University, Netherlands
Jared Linck, University of Maryland Center for Advanced Study of Language, USA

****Correspondence:***
Luis Morales, Department of Experimental Psychology and Behavioural Physiology, University of Granada, Campus de Cartuja s/n, 18071, Granada, Spain.
e-mail: luismorales@ugr.es

Inhibitory control processes have been recently considered to be involved in interference resolution in bilinguals at the phonological level. In this study we explored if interference resolution is also carried out by this inhibitory mechanism at the grammatical level. Thirty-two bilinguals (Italian-L1 and Spanish-L2) participated. All of them completed two tasks. In the first one they had to name pictures in L2. We manipulated gender congruency between the two languages and the number of presentations of the pictures (1 and 5). Results showed a gender congruency effect with slower naming latencies in the incongruent condition. In the second task, participants were presented with the pictures practiced during the first naming task, but now they were asked to produce the L1 article. Results showed a grammatical gender congruency effect in L1 that increased for those words practiced five times in L2. Our conclusion is that an inhibitory mechanism was involved in the suppression of the native language during a picture naming task. Furthermore, this inhibitory process was also involved in suppressing grammatical gender when it was a source of competition between the languages.

Keywords: grammatical gender, inhibition, bilinguals

INTRODUCTION

One important question in bilingual language processing is how people who speak two or more languages are able to control their linguistic production. People immersed in a context of second language acquisition often make mistakes and access native language words even when the alternative language is needed (Kroll and Stewart, 1994; Colomè, 2001). Furthermore, sometimes they report forgetting some words in their native language, when it is infrequently practiced (Seliger and Vago, 1991; De Bot, 1999). One approach to understanding these detrimental effects has been to propose that they are produced by a process that is similar to that producing forgetting during memory retrieval. Levy et al. (2007) have shown that retrieving some information from memory can produce forgetting of associated competing information (Anderson et al., 1994). They have suggested that forgetting of first-language lexical representation when immersed in a second language context may be due to an attentional inhibitory mechanism that suppresses unwanted memories to facilitate retrieval of the relevant information (Levy et al., 2007). In general, any situation that requires memory retrieval in the presence of competition will entail inhibition of the competing information (Anderson et al., 1994, 2000; Anderson and Spellman, 1995; Anderson and McCulloch, 1999; Anderson and Green, 2001). An indirect consequence of this process is that the inhibited information will be less accessible and harder to retrieve at a later moment. Two important properties of inhibition as a memory selection mechanism is that (1) inhibition depends on the presence of competition (Anderson et al., 1994); and (2) it is specific to the dimension of the memory representation that is competing for selection, meaning that inhibitory effects should depend on the degree to which the memory trace tapped by the final test matches the memory trace that was competing during the encoding phase. For example, if competition is lexical in nature (i.e., words starting with the same beginning), inhibition will specifically act upon the lexical representation that will be less accessible in a later test. But, to capture lexical inhibition, a lexical test (i.e., lexical decision) would be needed (Tulving and Thomson, 1973; Morris et al., 1977; Bajo et al., 2006).

Hence, similar to what occurs in memory, first-language forgetting may arise, at least in part, from the suppression or inhibition of native language. For that to occur, interference and competition between the two languages of the bilingual are required (Kroll and Stewart, 1994). The aim of the experiments reported in this paper is to show that in similar vein grammatical gender can also be inhibited. In order to provide a context for this claim, we will first discuss the evidence regarding language co-activation and between-language competition at the lexical and grammatical level and next we will go back to the evidence regarding inhibitory control in language selection.

LANGUAGE CO-ACTIVATION AND COMPETITION

Numerous studies have provided evidence that linguistic properties of the non-intended language affect the production of the intended language at the semantic and the phonological levels (Hermans et al., 1998; Costa and Caramazza, 1999; Costa et al., 1999, 2000; Colomè, 2001; Dijkstra, 2005; Macizo and Bajo, 2006; but see Costa et al., 2006, for a critical discussion). For instance, in a series of picture–word tasks, Costa et al. (1999) reported lexical connections between the two systems of bilingual Catalan–Spanish speakers. They found interference effects when participants had to name pictures presented with semantically related words for both same- and different-language conditions, relative to when they were presented with semantically unrelated words. On the other hand, Colomè (2001) used a phoneme monitoring task on words self-elicited from pictures to demonstrate that the language that a bilingual is not using is nevertheless activated. When

Spanish–Catalan bilinguals had to decide if a specific phoneme was present in the Catalan name of the picture, participants took longer to reject the phoneme when it was part of the Spanish word relative to a control condition.

Given the evidence of interaction between the semantic and phonological features of the lexical systems in bilinguals, we can also expect between-language competition at the level of grammatical gender. Grammatical gender has been proposed to be a property of the nouns that is stored at one representational level different from conceptual and phonological information (Caramazza and Miozzo, 1997; Levelt et al., 1999). However, how grammatical gender interacts during lexical selection in bilinguals is more controversial (Costa et al., 2003; Salamoura and Williams, 2007; Bordag and Pechmann, 2008; Lemhöfer et al., 2008; Paolieri et al., 2010a), probably due to the different characteristics of the gender systems in different languages. For instance, Costa et al. (2003) assume a complete autonomy of the gender systems of the two languages of the bilinguals. In a series of picture naming experiments manipulating grammatical gender congruency in different pairs of languages, the authors found similar naming times for same- and different-gender pictures. Costa et al. (2003) concluded that the grammatical gender of the words in the non-response language does not affect lexical processing in the response language. In contrast, Bordag and Pechmann (2007) and Lemhöfer et al. (2008) reported L1 and L2 interactions at the grammatical level of representation in Czech–German and German–Dutch bilinguals, respectively. Furthermore, they observed between-language gender interaction even when they controlled for the influence of phonological form (e.g., noun termination) in both production and comprehension tasks (Bordag and Pechmann, 2007); and even when using a lexical decision task where the cognate status of the words was manipulated (Lemhöfer et al., 2008).

Within the context of bilingualism, the effect of gender congruency has been also found in both bare noun production and noun phrase production with German–Dutch and Italian–Spanish speakers (Lemhöfer et al., 2008; Paolieri et al., 2010a)[1]. Paolieri et al. (2010a) observed robust gender congruency effects with Italian–Spanish bilinguals. In this study participants had to name pictures in L2 or to translate words from L1 to L2, producing either bare noun or noun phrases. In all conditions, participants showed shorter response latencies when the nouns of the two languages shared grammatical gender than when their grammatical gender was different. Thus, independently of the type of task (L2 picture naming or forward translation) and on the type of response (bare noun or noun phrase), their results speak in favor of grammatical gender interactions between the two languages of the bilinguals. These results contradict the notion that grammatical gender is only selected when producing gender-marked utterances (Caramazza and Miozzo, 1997; Levelt et al., 1999), and support the idea that the selection of one lexical node involves obligatory access to syntactic features (Cubelli et al., 2005; Paolieri et al., 2010a,b). And more importantly, they suggest that the two lexical–grammatical representations of the words are activated in the bilinguals mind and compete whenever lexical selection is needed.

INHIBITORY CONTROL IN LANGUAGE SELECTION

Given that most studies point to a non-selective activation of languages during bilingual production, the question concerns how the system handles such unintended activation. For example, the model proposed by Costa and collaborators (Costa et al., 1999; Costa, 2005) assumes that although the lexical candidates in both languages are active simultaneously, the intention to speak only one of them restricts selection to the target language. In this way, co-activation does not lead to interference and competition during the planning of the utterance.

However, another possibility is that both lexical representations also compete for selection, and that such selection is managed by inhibitory processes acting on the lexicon. One version of inhibitory model (Inhibitory Control Model; Green, 1998) claims that inhibitory control is triggered whenever active lexical representations from the two languages compete for selection. This inhibitory mechanism is in charge of suppressing the non-target representations; as a consequence between-language interference is reduced and selection of the appropriate entries is facilitated. The role of inhibitory processes on selection is not restricted to the bilingual field, but it is shared with other cognitive areas such as visual attention, memory and language comprehension and production. For example, popular explanations of the inhibition of return effect (e.g., Tipper et al., 2003) have suggested that already-sampled spatial locations are inhibited to facilitate visual search. Similarly, some memory theories assume that inhibition of competing representation facilitate retrieval of target memories (Anderson et al., 1994), and many theories of language production assume that lexical selection is achieved by means of inhibitory connections at the level of lexical representations (e.g., Berg and Schade, 1992; Cutting and Ferreira, 1999). Hence, research in different cognitive domains has suggested that both lexical and perceptual representations can be inhibited.

Most of the evidence regarding inhibitory language control comes from results of the language switching tasks (Meuter and Allport, 1999; but see Abutalebi and Green, 2008; for a review). In these studies participants are required to name digits or pictures in L1 or L2 in an unpredictable manner. Hence, there are trials in which the response language is the same as that in the preceding trial (non-switch trials) and trials in which the response language differs from the preceding trial (switch trials). When bilinguals perform this task they are slower in switching trials relative to non-switch trials (switching cost), but the most interesting pattern is that switching from L2 to L1 produces a larger cost than switching from L1 to L2. This asymmetrical cost has been interpreted as evidence of inhibition by assuming that naming in L2 requires inhibition of the more dominant L1, so that when bilinguals switch back into the L1 naming, additional time would be required to overcome the strong inhibition of L1 representations.

[1]The selection of grammatical gender in bare noun production is a controversial topic in monolingual language production. A reliable effect of grammatical gender congruency in bare noun production has been found in Italian and Spanish, two Romance languages with a similar morphological system. In contrast, with Italian and Spanish, the gender congruity effect has never been observed in Dutch (La Heij et al., 1998; Starreveld and La Heij, 2004), where no inflection has to be selected for the production of bare nouns. To explain the different pattern of results in Italian, Spanish, and Dutch, Cubelli et al., 2005; see also Paolieri et al., 2011) assumed that the gender congruency effect in the picture–word paradigm depends on the specific, formal, morphosyntactic properties of individual languages.

Similarly, Linck et al. (2009) provided support for the inhibitory account in a study in which they compared L2 learners immersed in a L2 context with L2 learners without immersion experience. In a very simple task, they showed that relative to classroom learners, the immersed learners produced significantly more examples in L2, but more interestingly, they produced significantly fewer examples in their L1, indicating that L2 immersion increases the amount of inhibition on L1 so that L1 become less available for the immersed bilinguals. Note that inhibition in the language switching and verbal fluency tasks are global in nature and directed to the non-appropriate language. In this sense, these tasks do not tap into specific memory representations since the lexical and conceptual units change from one trial to next and therefore is the language what it needs to be inhibited.

Evidence for representation specific inhibition comes from several recent lines of research. For example, Macizo et al. (2010) and Martín et al. (2010) asked Spanish–English bilinguals to perform a relatedness judgment task including interlexical homographs (e.g., "*pie*," meaning "foot" in Spanish). Pairs of English words were presented and the participants had to decide whether or not they were related. Results indicated that participants were slower to respond to homographs presented along with words related to the irrelevant Spanish meaning of the homograph relative to control words (e.g., "*pie-toe*" vs. "*log-toe*"). Moreover, after responding to homographs, the participants responded more slowly when the following trial required activation of the irrelevant homograph meaning (e.g., "*foot-hand*" preceded by "*pie-toe*"). These results suggest that bilinguals activated both of their languages (homograph interference) and that they inhibited the irrelevant homograph meaning in order to overcome interference and perform the task.

Similarly, Levy et al. (2007) have also demonstrated that inhibition is responsible for the suppression of native language at the phonological level. In their study, native English speakers had to name pictures in Spanish-L2 for 1, 5, or 10 times (e.g., *culebra*; snake). Afterward, the accessibility to the same words in the native language was measured using an independent probe (Anderson and Spellman, 1995) rhyme test (e.g., *break-s*_____). Results showed that words named in Spanish (L2) 5 or 10 times led to decreased recall of the corresponding English (L1) names than those named in L1 or named in L2 only once. Moreover, in Experiment 2 they were able to isolate the specific inhibitory effect to phonology since presenting semantic cues (e.g., *venom-s*_____) did not produce the forgetting effect of naming repeatedly pictures in L2. Thus, repeatedly naming L2-words inhibited the phonology of their English (L1) names, but facilitated concept accessibility. The authors conclude that phonological first-language attrition arises from inhibition of the phonological native language representations during second language use. This experiment illustrates the importance of inhibitory mechanism in overcoming interference during second language acquisition.

Hence, although there is much evidence showing that the two language systems of the bilingual interact at the semantic, phonological, and grammatical levels (Costa and Caramazza, 1999; Costa et al., 2000; Colomè, 2001; Paolieri et al., 2010a), and that inhibitory mechanisms are triggered to reduce the interference due to co-activation at the semantic (Macizo et al., 2010; Martín et al., 2010) and phonological level (Levy et al., 2007; but see Finkbeiner et al., 2006, for a critical discussion), there is no evidence showing that inhibitory processes can also act at the lexical/grammatical (gender) level of representation.

The aim of this study is to confirm that the two lexical systems of the bilingual are activated and interact at the grammatical gender level, and more interestingly, to investigate whether inhibitory mechanisms are responsible for resolving between-language competition at this representational level. Similarly to Levy's study (Levy et al., 2007), we asked Italian native speakers to name pictures in Spanish-L2 by producing bare nouns. In this first picture naming phase, we manipulated the gender congruency of the nouns between the two languages (grammatical gender congruent vs. grammatical gender incongruent) and the number of presentations of each picture (one or five times), in order to create more or less L1 inhibition. Note that picture naming involves the activation of the grammatical properties of the language (Cubelli et al., 2005; Paolieri et al., 2010a,b), as long as these grammatical properties of the two languages are activated and are incongruent (Paolieri et al., 2010a), they will compete for selection and the inappropriate grammatical feature would be inhibited. Hence, words with incongruent gender in the two languages would produce competition that in turn would trigger inhibition. In addition, the higher the number of naming trials in L2, the greater the inhibition that would act upon the particular L1 incongruent grammatical property.

In the second phase, participants had to complete an article production task in Italian-L1 for the same pictures practiced in L2 during the first task. This task was selected because it specifically captures gender access since participants are asked to produce only the definite article. We expected that trials containing incongruent gender stimulus would show slower response times when producing the article in L1; and more importantly, that this difference would be larger for words practiced more times in the previous L2 naming task. For this later task, new pictures (never presented during the naming phase of the experiment) were added as a baseline to observe the effect of previous naming (see Levy et al., 2007, for a similar procedure). Given that participants had to produce the definite article in their native language, we did not expect gender effects for these new items.

MATERIALS AND METHODS

PARTICIPANTS

Thirty-two Italian–Spanish bilinguals voluntarily participated in the experiment. L2 proficiency was assessed at the end of the session through a subjective questionnaire (see **Table A1** in Appendix for a description of the sample of participants). They all had normal or corrected-to-normal vision.

DESIGN AND MATERIALS

The experiment consisted of two main phases: (1) Picture naming task in L2 (Spanish) by producing bare nouns, and (2) Retrieval of L1 article corresponding to the presented pictures. This design was created in order to produce the inhibition of Italian-L1 gender by naming gender congruent and incongruent items in Spanish-L2 during the first part of the experiment, and then measure access to the specific representations of these lexical entries during the Italian-L1 task (see Levy et al., 2007, for a similar procedure).

Grammatical gender (Congruent vs. Incongruent) and Number of presentations of each picture (1 vs. 5) were manipulated within subjects during the Spanish (L2) naming task. Seventy-two pictures were chosen from the sets of Lotto et al. (2001), half with the same gender in Italian and Spanish (e.g., $Sciarpa_{FEM}$ and $Bufanda_{FEM}$, in Italian and Spanish respectively –scarf–) and half with different gender (e.g., $Letto_{MAS}$ and $Cama_{FEM}$, in Italian and Spanish respectively –bed–). At the same time, half of the words were masculine and half were feminine in gender. This set of stimulus consisted of 48 experimental pictures to be used both in the first and second task, and 24 additional control items to be included as baseline for the second task (a complete list of the stimulus materials is provided in **Table A2** in Appendix). Cognate words were not included as experimental material. Gender Congruent and Incongruent words did not differ (all ts < 1) for frequency (Alameda and Cuetos, 1995 for Spanish, and Bertinetto et al., 2005 for Italian), number of letters, number of syllables, and phonological/orthographic overlap. The last one was calculated computing the percentage of number of letters shared by the words in the two languages.

For the picture naming task in L2 (task 1), half of the pictures were presented once and half five times. Two different pseudo-random lists including 48 experimental items were created. Lists were constrained as follows: (1) No more than three congruent or incongruent stimuli could appear consecutively; (2) the lag between repetitions of a particular picture had to be of at least three trials; (3) no semantic or phonological relationship could exist between pictures in consecutive trials. Finally, each list included a total of 144 trials. Repetitions of each picture and list were counterbalanced across participants.

Regarding the article retrieval task in L1 (task 2), one randomized list was created and divided in two blocks counterbalanced across participants. The list consisted of a total of 72 trials (48 experimental pictures named in L2 during the previous task plus 24 new control pictures).

PROCEDURE

Participants were tested individually. The experimenter was seated behind the participant to record errors and responses. The stimuli were presented using E-Prime experimental software, 1.1 version (Schneider et al., 2002). The whole experiment lasted about 40 min. Before starting, participants completed a familiarization phase with the complete set of pictures. A trial in the familiarization phase consisted of the presentation of each picture with its translation in both languages (e.g., "*Il letto – La cama*," for the picture of a bed). The participants had to indicate to the experimenter if they knew the words in Spanish (L2). Then, the experimental tasks were administered in the following order: (1) Picture naming task in L2 and (2) article naming task in L1.

Task 1: picture naming task in L2

The objective of this task was to produce inhibition of the nouns in Italian-L1. Participants had to name pictures in Spanish-L2, and they were instructed to name them as quickly and accurately as possible using the bare noun (i.e., without using the define article "*el*" or "*la*" in Spanish). A trial consisted of the following events: A fixation point (+) presented at the center on screen for 750 ms; presentation of the picture until the participants' response or for a maximum of 4000 ms; and a blank interval for 750 ms before the next trial. A practice block of 12 trials was administrated before starting the task. Naming latencies were measured from the onset of the stimuli until the beginning of the response. Naming errors and equipment failures were registered.

Task 2: article production task in L1

The objective of this task was to measure the speed of access to the grammatical gender information of those nouns practiced during the previous task in L2. For that, the participants had to retrieve and name the definite article corresponding to the presented pictures (the same practiced in L2 during the previous task plus the new control items). Each trial consisted of the following sequence of events: A fixation point (+) for 750 ms; the presentation of the picture that remained on the screen until response or for a maximum of 4000 ms; and a blank interval for 750 ms.

Finally, an L2 subjective questionnaire was administered.

RESULTS

TASK 1: PICTURE NAMING IN L2

Several types of responses were excluded: (1) Naming latencies below 300 ms and exceeding 2500 ms, (2) naming errors and verbal dysfluencies, (3) Spanish words unknown by the participant. Overall, 24.11% of the trials were excluded from the analyses [70% of that percentage was due to non-responses, and these trials were not included in the analyses of the second task (see below)]. An ANOVA introducing Grammatical Gender (Congruent vs. Incongruent) and Number of Presentations for each picture (1 vs. 5) revealed a main effect of Grammatical Gender [F (1, 31) = 4.367, MSE = 2.954, $p = 0.004$], with congruent items 20 ms faster than incongruent ones [884 ms (SD = 176) and 904 ms (SD = 177), respectively]. The main effect of Number of Presentations was also significant [F (1, 31) = 183.474, MSE = 11.525, $p = 0.0001$], revealing faster naming latencies with pictures practiced more times [1022 ms (SD = 131) and 756 ms (SD = 110), for pictures practiced once or five times, respectively]. Finally, the interaction between the variables was not significant [F (1, 31) = 0.855, MSE = 4.565, $p = 0.362$].

TASK 2: ARTICLE PRODUCTION TASK IN L1

Naming errors (8.37% of the trials), verbal dysfluencies, response times below 300 ms and exceeding 2500 ms, and naming latencies for those pictures that were never successfully named during the previous task in L2 were eliminated from the analysis (overall, 31.34% of the trials). Naming errors included cases where the participants produced the wrong name of the picture as well as the wrong article (unfortunately our coding system did not permit to separate the two types of naming errors). An analysis of these combined errors comparing the Congruent and Incongruent conditions showed that incongruent nouns produced significantly more errors than congruent ones (109 and 64, respectively) [F (1, 31) = 12.1304, MSE = 2.6084, $p = 0.001$]. Regarding the latencies for the article production, an ANOVA including Grammatical Gender (Congruent vs. Incongruent) × Number of Presentations (0, 1, and 5) showed a main effect of Grammatical Gender [F (1, 31) = 19.684, MSE = 19.429, $p = 0.0001$], Number of Presentations [F (2, 62) = 10.021, MSE = 14.039, p = 0.0001], and the interaction of Gender × Number of Presentations [F (2, 62) = 11.554, MSE = 11.163, p = 0.0001].

In order to understand this interaction, we compared, first, congruency effects for each level of repetition. Planned comparisons yielded significant differences between congruent and incongruent items practiced once in L2, with slower RT in the incongruent condition [1089 ms (SD = 203) and 1200 ms (SD = 195); F (1, 31) = 17.224, p = 0.0002]. This difference was also significant when the pictures were practiced five times [1091 ms (SD = 196) for congruent and 1257 ms (SD = 213) for incongruent; F (1, 31) = 27.074, p = 0.0001], but not when the pictures were practiced zero times [1241 ms (SD = 220) for congruent and 1231 ms (SD = 193) for incongruent; F (1, 31) = 0.100, p = 0.752]. Note that non-repeated pictures were never named in L2, and therefore they were never subject to interference. Because article production for these new pictures was performed in L1, it is not surprising that congruency effect were not present. However, when the pictures were named in L2 and they were incongruent, the more times the pictures were named in L2, the harder to find the appropriate article in L1. That is, the gender congruency effect became larger with repetitions because incongruent articles were harder to retrieve.

When we compared 1 vs. 5 L2 naming for congruent and incongruent items, repetition effects were only present for the incongruent condition[2]. The results of these comparisons indicated that for incongruent nouns significant differences between pictures practiced one and five times in L2 were obtained, with slower RT for the pictures practiced five times [1200 ms (SD = 195) and 1257 ms (SD = 213); F (1,31) = 4.896, p = 0.03]. In contrast, this difference was not significant for congruent items [1089 ms (SD = 203) for one repetition and 1091 ms (SD = 196) for five repetitions; F (1, 31) = 0.006, p = 0.93]. This pattern indicates that the congruency effect was driven by an increased interference in the incongruent condition with more repetitions, and not by facilitation in the congruent condition across repetitions.

DISCUSSION

The aim of this study was to demonstrate that not only the two gender systems of a bilingual are functionally connected, but also that this co-activation can cause competition processes that are resolved by inhibitory mechanisms at the grammatical level of representation. In the first phase of the experiment, we found

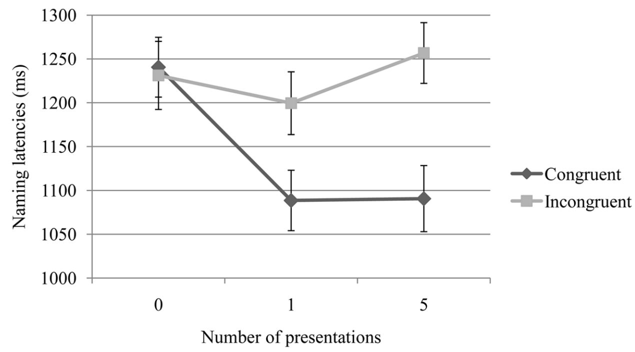


FIGURE 1 | Response latencies (in milliseconds) producing the L1 definite article for those pictures presented zero, one, or five times in the previous L2 picture naming task.

that participants took more time naming pictures with incongruent Italian–Spanish gender. Furthermore, this effect was observed through a bare noun picture naming task in which explicit access to the grammatical gender information of each word is not mandatory. This result supports the notion that grammatical gender selection is not constrained to noun phrase production tasks, in which explicit access to the gender representation is required (i.e., when participants are asked to name pictures using the gender-marked definite article; Cubelli et al., 2005; Paolieri et al., 2010a,b), and that grammatical gender is a lexical property that is automatically activated and interacts across the bilinguals lexical systems (Bordag and Pechmann, 2007; Lemhöfer et al., 2008; Paolieri et al., 2010a). Although we do not have a monolinguals control condition in the experiment to show that the effect is really due to between-language activation in bilinguals and to possible differences between gender congruent and incongruent words, Paolieri et al. (2010b) tested Italian monolingual participants with similar materials and showed that this effect was not present in monolinguals. In summary, results from task 1 suggest that grammatical gender is an intrinsic part of the lexical representation, and it is always available when a noun is retrieved (Paolieri et al., 2010a,b). Therefore, gender effects should be observed in all tasks requiring lexical access, whether producing a noun phrase with explicit gender markers or the bare noun along.

In our study, between-language gender incongruency introduced competition so that when there was no agreement between Italian–Spanish gender for the corresponding object, naming latencies were slower than when there was agreement between them. This between-language competition at the grammatical level seems to have triggered inhibitory processes. Then, the interference created by gender incongruency was resolved by inhibiting grammatical gender representation of the Italian-L1 words in order to facilitate the correct naming of each picture in Spanish-L2. Because of this inhibition, later retrieval of L1 grammatical information (retrieving the appropriate article) of incongruent words took longer relative to the retrieval of the appropriate article for gender congruent pictures.

According to the IC model (Green, 1998); bilinguals trigger inhibitory control mechanisms to select the desired

[2]It could be argued that the proper comparison to claim inhibitory effects is the comparison between zero and five repetitions. In fact, RIF effects in standard memory procedures with categorical materials come from comparing practiced items from practiced categories to items belonging to unpracticed categories. However, we think that in the present procedure the proper comparison involves one to five repetitions. Standard RIF with categorical material involves the presentation of common familiar concepts, whereas the L2 picture naming task in the present experiment (see also Levy et al., 2007) involves the presentation of new unfamiliar pictures (depicting common objects). Hence, the first naming trial would increase the familiarity with the picture and produce facilitation (see Johson and Anderson, 2004; and Levy et al., 2007, for further discussion and similar results in other inhibitory paradigms). Although not significant ($p > 0.05$), **Figure 1** shows that RT to items named for the first time is faster than the RT to new items, these differences in perceptual familiarity may obscure inhibitory effects when comparing 0–5 repetitions in incongruent trials ($p > 0.05$). However, the inhibitory effects clearly emerge when comparison involve already familiar items (1 vs. 5 presentations).

representations when they experience between-language competition. In this study we show that during an L2 naming task both L1 and L2 lexical representations are activated and compete, in particular this competition is evident when the grammatical gender information in the two languages is incongruent. The results of the picture naming task demonstrate that the participants took more time naming the pictures when the grammatical gender of the corresponding names was incongruent than when it was congruent. This congruency effect demonstrates that L1 was activated even when only L2 was needed for naming and that this activation included grammatical features.

More importantly, results of the second task involving retrieval of the article in L1 indicate that the grammatical competition during L2 naming was resolved by specifically inhibiting the competing grammatical gender in L1. Note that in the article naming task access to the gender information was needed, and therefore it is a task that specifically taps gender processing, in order to measure the access of gender representation properly (see Bajo et al., 2006, for the importance of task specificity to test inhibition). Although we found a significant gender congruency effect between objects practiced once and five times, the fact that this effect in L1 is larger as the number of repetitions in L2 increases clearly show that this gender congruency effect is the result of the previous naming in L2. In addition, the absence of such effect with pictures never presented for L2 naming also signals that the slower response times with repetition are due to the mechanism involved in reducing gender interference during picture naming. Nevertheless, direct evidence in favor of an inhibitory account is provided when we focus on the effect of repetition on incongruent pictures and the increment in L1 article retrieval for pictures named five times in L2. The fact that this effect was absent for congruent objects tell us that the impairment is caused for the competition arisen for the incongruent between-language gender for the nouns, and not for facilitatory effects in the congruent condition.

However, it could be argued that this data are open to alternative explanations. First, it could be argued that the congruency effect is not due to the co-activation of grammatical features that compete for selection, but to the effect of determiners similarity. This might be the case because of the particular form of the determiners used in Spanish and Italian. Thus, in Spanish they are "el" for masculine and "la" for feminine, whereas in Italian there are "il"/"lo" for masculine and "la" for feminine. So, the incongruency effect could be interpreted as due to the similarity in word form of the Spanish and Italian feminine determiners. To rule out this alternative explanation we performed additional analyses introducing gender as a variable. If the gender effect was due to form similarity we should find that in the L2 naming task the masculine condition should produce longer effects than the feminine condition. The results of the ANOVA on the L2 naming times with Gender, Congruency, and Repetition as independent variables showed a main effect of Congruency [F $(1, 31) = 5.3811$, $p < 0.05$], and Repetition [F $(1, 31) = 173.934$, $p < 0.05$]. However, Gender (feminine vs. masculine) was not significant and did not interact with any of the other variables (all ps > 0.05). This suggests that the congruency effect was not due to form similarity, but to gender incongruency. In addition, the ANOVA performed in the article naming task of the second phase showed that the critical Gender × Congruency × Repetition interaction was not significant [F $(2, 56) = 0.276$, $p > 0.05$]. Indicating that Congruency × Repetition (the inhibitory index) was similar for both feminine and masculine.

Similarly, it could be argued that the effect of repetition in incongruent trials might be due to associative interference. Within the memory field, some have argued that the forgetting induced by retrieval of information is due to the strengthening of the practiced items with the contextual cue, so that when that cue is later presented for recall, the strengthen representation is activated first and block the retrieval of the non-practiced items (Raaijmakers and Shiffrin, 1981). In this context, this associative account would suggest that practice in L2 naming would strengthen the relation between the presented pictures and the L2 name, so that later, when participants saw the pictures again the strengthen L2 representation would come to mind and block the retrieval of the L1 representation. In the memory literature, this interpretation has been ruled out by showing that retrieval induced forgetting is also produce when the task used to capture forgetting of the unpracticed items does not test the strengthened relation. This is done by presenting either novel cues (Anderson and Spellman, 1995; Bajo et al., 2006) or item specific tests (Román et al., 2009). Although our procedure is not exactly cue independent, in our experiment the particular tasks used during the first and second phase were selected so that associative interference was not present. Thus, in the first phase a bare noun naming task was used to avoid the presentation of the L2 determiner, whereas in the second phase we asked participants to only name the L1 determiner corresponding to the object in the picture. Hence, the picture and the L2 determiner were never presented together during the first phase to produce strengthening of the picture-determiner representation. Hence, the relation between the picture and the L2 determiner was never strengthened and there are no reasons to think that the determiner in L2 was blocking retrieval of the L1 determiner.

Hence, the results speak in favor of the importance of inhibitory control mechanisms in resolving between-language competition in bilinguals at the grammatical gender level. Levy et al. (2007) observed co-activation at the phonological level and were able to show that phonological competition was resolved by means of inhibition of phonological representations. In our experiment, we were able to find a similar pattern of inhibition at the grammatical gender level. Together, both studies highlight the importance of executive control mechanisms for controlling language production in bilinguals.

In conclusion, grammatical gender information is a lexical representation that is automatically activated and can cause competition between-languages with similar gender systems. This interference seems to be solved by inhibitory mechanisms that suppress momentarily the grammatical gender representation of specific lexical entries. Although more research is needed to isolate the specific inhibition of competitive traces, the fact that competition processes are required for inhibition to occur seems to be clear.

ACKNOWLEDGMENTS

This research was supported by doctoral research grant from the Andalusian Government (P08-HUM-03600) to Luis Morales; grants EDU2008-01111 and CSD2008-00048 from the Spanish Ministry of Education; and grants P07-HUM-02510 and P08-HUM-03600 from the Andalusian Government. We would like to thank Laure Ann Stowe, Wido la Heij, and Jared Linck for their helpful comments and suggestions.

REFERENCES

Abutalebi, J., and Green, D. W. (2008). Control mechanisms in bilingual language production: neural evidence from language switching studies. *Lang. Cogn. Process.* 23, 557–582.

Alameda, J. R., and Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

Anderson, M. C., Bjork, E. L., and Bjork, R. A. (2000). Retrieval-induced forgetting: evidence for a recall-specific mechanism. *Psychon. Bull. Rev.* 7, 522–530.

Anderson, M. C., Bjork, R. A., and Bjork, E. L. (1994). Remembering can cause forgetting: retrieval dynamics in long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 1063–1087.

Anderson, M. C., and Green, C. (2001). Suppressing unwanted memories by executive control. *Nature* 410, 131–134.

Anderson, M. C., and McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 608–629.

Anderson, M. C., and Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: memory retrieval as a model. *Psychol. Rev.* 102, 68–100.

Bajo, T., Gómez-Ariza, C. J., Fernández, A., and Marful, A. (2006). Retrieval-induced forgetting in perceptually driven memory tests. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1185–1194.

Berg, T., and Schade, U. (1992). The role of inhibition in a spreading activation model of language production. I. The psycholinguistic perspective. *J. Psychol. Res.* 21, 405–434.

Bertinetto, P. M., Burani, C., Laudanna, A., Marconi, C., Ratti, D., Rolando, C., and Thornton, A. M. (2005). *CoLFIS (Corpus e Lessico di Frequenza dell'Italiano Scritto)* [CoLFIS (Corpus and Frequency Lexicon of Writteen Italian]. Available at: http://www.istc.cnr.it/material/database [accessed July 6, 2006].

Bordag, D., and Pechmann, T. (2007). Factors influencing L2 gender processing. *Biling. (Camb. Engl.)* 10, 299–314.

Bordag, D., and Pechmann, T. (2008). Grammatical gender in translation. *Second Lang. Res.* 24, 139–166.

Caramazza, A., and Miozzo, M. (1997). The relation between syntactic and phonological knowledge in lexical access: evidence from the "tip-of-the-tongue" phenomenon. *Cognition* 64, 309–343.

Colomè, A. (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736.

Costa, A. (2005). "Lexical access in bilingual production," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York: Oxford University Press), 308–325.

Costa, A., and Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. *Biling. (Camb. Engl.)* 2, 231–244.

Costa, A., Caramazza, A., and Sebástian-Gallés, N. (2000). The cognate facilitation effect: implications for models of lexical access. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 1283–1296.

Costa, A., Kovacic, D., Franck, J., and Caramazza, A. (2003). On the autonomy of the grammatical gender systems of the two languages of a bilingual. *Biling. (Camb. Engl.)* 6, 181–200.

Costa, A., La Heij, W., and Navarrete, E. (2006). The dynamics of bilingual lexical access. *Biling. (Camb. Engl.)* 9, 137–151.

Costa, A., Miozzo, M., and Caramazza, A. (1999). Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *J. Mem. Lang.* 41, 365–397.

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., and Job, R. (2005). Grammatical gender is selected in bare noun production: evidence from the picture-word interference paradigm. *J. Mem. Lang.* 53, 42–59.

Cutting, J. C., and Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 318–344.

De Bot, K. (1999). "The psycholinguistics of language loss," in *Studies on Language Acquisition*, eds G. Extra and L. Verhoeven (Berlin: Walter de Gruyter), 345–361.

Dijkstra, T. (2005). "Bilingual word recognition and lexical access," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York: Oxford University Press), 179–201.

Finkbeiner, M., Almeida, J., Janssen, N., and Caramazza, A. (2006). Lexical selection in bilingual speech production does not involve language suppression. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1075–1089.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. (Camb. Engl.)* 1, 67–81.

Hermans, D., Bongaerts, T., de Bot, K., and Schreuder, R. (1998). Producing words in a foreign language: can speakers prevent interference from their first language. *Biling. (Camb. Engl.)* 1, 213–230.

Johson, S. K., and Anderson, M. C. (2004). The role of inhibitory control in forgetting semantic knowledge. *Psychol. Sci.* 15, 448–453.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence from asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174.

La Heij, W., Mark, P., Sander, J., and Willeboorsde, E. (1998). The gender-congruency effect in picture-word tasks. *Psychol. Res.* 61, 209–219.

Lemhöfer, K., Spalek, K., and Schriefers, H. (2008). Cross-language effects of grammatical gender in bilingual word recognition and production. *J. Mem. Lang.* 59, 312–330.

Levelt, W. J. M., Roelofs, A., and Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav. Brain Sci.* 22, 1–75.

Levy, B. J., McVeigh, N. D., Marful, A., and Anderson, M. C. (2007). Inhibiting your native language: the role of retrieval-induced forgetting during second-language acquisition. *Psychol. Sci.* 18, 29–34.

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515.

Lotto, L., Dell'Acqua, R., and Job, R. (2001). Le figure PD/DPSS. Misure di accordosulnome, tipicità, familiarità, età di acquisizione e tempi di denominazione per 266 figure [PD/DPSS pictures: name agreement, typicality, familiarity, age of acquisition norms and naming times of 266 pictures]. *GiornaleItaliano di Psicologia* 28, 231–245.

Macizo, P., and Bajo, M. T. (2006). Reading for understanding and reading for translation: are they equal? *Cognition* 99, 1–34.

Macizo, P., Bajo, T., and Martín, M. C. (2010). Inhibitory processes in bilingual language comprehension: evidence from Spanish-English interlexical homographs. *J. Mem. Lang.* 63, 232–244.

Martín, M. C., Macizo, P., and Bajo, T. (2010). Time course of inhibitory processes in bilingual language processing. *Br. J. Psychol.* 101, 679–693.

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40.

Morris, C. D., Bransford, J. D., and Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *J. Verbal Learn. Verbal Behav.* 16, 519–533.

Paolieri, D., Cubelli, R., Macizo, P., Bajo, M. T., Lotto, L., and Job, R. (2010a). Grammatical gender processing in Italian and Spanish bilinguals. *Q. J. Exp. Psychol.* 63, 1631–1645.

Paolieri, D., Lotto, L., Morales, L., Bajo, T., Cubelli, R., and Job, R. (2010b). Grammatical gender processing in romance languages: evidence from bare noun production in Italian and Spanish. *Eur. J. Cogn. Psychol.* 22, 335–347.

Paolieri, D., Lotto, L., Leoncini, D., Cubelli, R., and Job, R. (2011). Differential effects of grammatical gender and gender inflection in the picture-word interference paradigm. *Br. J. Psychol.* 102, 19–36.

Raaijmakers, J. G. W., and Shiffrin, R. M. (1981). Search of associative memory. *Psychol. Rev.* 88, 93–134.

Román, P. E., Soriano, M. F., Gómez-Ariza, C. J., and Bajo, M. T. (2009). Retrieval-induced forgetting and executive control. *Psychol. Sci.* 20, 1053–1058.

Salamoura, A., and Williams, J. N. (2007). The representation of grammatical gender in the bilingual lexicon: evidence from Greek and German. *Biling. (Camb. Engl.)* 10, 257–275.

Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime User's Guide (Version 1.1)*. Pittsburgh, PA: Psychology Software Tools.

Seliger, H. W., and Vago, R. M. (eds). (1991). *First Language Attrition.* Cambridge: Cambridge University Press.

Starreveld, P. A., and La Heij, W. (2004). Phonological facilitation of grammatical gender retrieval. *Lang. Cogn. Process.* 19, 677–711.

Tipper, S. P., Grison, S., and Kessler, K. (2003). Long-term inhibition of return of attention. *Psychol. Sci.* 14, 19–25.

Tulving, E., and Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychol. Rev.* 80, 352–373.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 April 2011; paper pending published: 28 May 2011; accepted: 09 October 2011; published online: 25 October 2011.

*Citation: Morales L, Paolieri D and Bajo T (2011) Grammatical gender inhibition in bilinguals. Front. Psychology **2**:284. doi: 10.3389/fpsyg.2011.00284*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDIX

Table A1 | Language history and self-evaluated proficiency scores of the Italian–Spanish bilinguals.

Age (years)	24.66 (4.83)
LANGUAGE HISTORY	
Use of L2 (years)	2.88 (4.41)
Living in Spain (years)	2.13 (3.03)
SELF-EVALUATED PROFICIENCY LEVEL TEST IN L2	
Production	7.22 (1.22)
Comprehension	8.00 (1.41)
Writing	6.22 (1.76)
Reading	7.75 (1.54)

The scores are on a 10-point scale, in which 10 represents native speakers level and one complete ignorance of the language. Mean are shown, with SD in parentheses.

Table A2 | Stimulus material.

Incongruent (Spanish–Italian)	Congruent (Spanish–Italian)
EXPERIMENTAL PICTURES	
Almohada-Cuscino (pillow)	Bufanda-Sciarpa (scarf)
Cama-Letto (bed)	Falda-Gonna (skirt)
Mesa-Tavolo (table)	Mariposa-Farfalla (butterfly)
Mochila-Zaino (backpack)	Maleta-Valigia (suitcase)
Tapadera-Coperchio (lid)	Ventana-Finestra (window)
Seta-Fungo (mushroom)	Manzana-Mela (apple)
Tirita-Cerotto (band-aid)	Calabaza-Zucca (pumpkin)
Gaviota-Gabbiano (seagull)	Iglesia-Chiesa (church)
Bota-Stivale (boot)	Golondrina-Rondine (swallow)
Mantequilla-Burro (butter)	Abeja-Ape (bee)
Nariz-Naso (nose)	Sartén-Padella (pan)
Flor-Fiore (flower)	Nuez-Noce (walnut)
Cepillo-Spazzola (brush)	Grifo-Rubinetto (faucet)
Columpio-Altalena (swing)	Loro-Pappagallo (parrot)
Mono-Scimmia (monkey)	Taladro-Trapano (power drill)
Trineo-Slitta (sled)	Apio-Sedano (celery)
Zapato-Scarpa (shoe)	Cazo-Mestolo (ladle)
Cigarro-Sigaretta (cigaret)	Corcho-Tappo (cork)
Globo-Mongolfiera (hot air)	Sombrero-Cappello (hat)
Mosquito-Zanzara (mosquito)	Queso-Formaggio (cheese)
Zorro-Volpe (fox)	Perro-Cane (dog)
Tornillo-Vite (screw)	Vaso-Bicchiere (glass)
Coche-Macchina (car)	Tomate-Pomodoro (tomato)
Enchufe-Spina (plug)	Reloj-Orologio (clock)
CONTROL PICTURES	
Tenedor-Forchetta (fork)	Buitre-Avvoltoio (vulture)
Sobre-Busta (envelope)	Avestruz-Struzzo (ostrich)
Despertador-Sveglia (alarm clock)	Tiburón-Squalo (shark)
Rallador-Grattugia (grater)	Paraguas-Ombrello (umbrella)
Hombro-Spalla (shoulder)	Gusano-Verme (warm)
Látigo-Frusta (whip)	Taburete-Sgabello (stool)
Araña-Ragno (spider)	Zanahoria-Carota (carrot)
Ardilla-Scoiattolo (squirrel)	Jarra-Brocca (pitcher)
Cartera-Portafoglio (wallet)	Olla-Pentola (pot)
Galleta-Biscotto (cookie)	Pata-Zampa (leg)
Escopeta-Fucile (shotgun)	Carretera-Strada (road)
Bata-Camice (white coat)	Cereza-Ciliegia (cherry)

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 June 2011
doi: 10.3389/fpsyg.2011.00123

The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition

Lily Tao[1], Anna Marzecová[2], Marcus Taft[1], Dariusz Asanowicz[2] and Zofia Wodniecka[2]*

[1] School of Psychology, University of New South Wales, Sydney, NSW, Australia
[2] Institute of Psychology, Jagiellonian University, Krakow, Poland

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
David William Green, University College London, UK
Juan Lupiáñez, University of Granada, Spain

***Correspondence:**
Zofia Wodniecka, Institute of Psychology, Jagiellonian University, Al. Mickiewicza 3, Krakow, Poland.
e-mail: zofia.wodniecka@uj.edu.pl

Previous studies have demonstrated a bilingual advantage in the efficiency of executive attention. A question remains, however, about the impact of the age of L2 acquisition and relative balance of the two languages on the enhancement of executive functions in bilinguals, and whether this is modulated by the similarity of the bilingual's two languages. The present study explores these issues by comparing the efficiency of attentional networks amongst three groups of young adults living in Australia: English monolinguals and early and late Chinese–English bilinguals. We also address the impact of bilingualism on hemispheric lateralization of cognitive functions, which is of interest since a recent study on early bilinguals revealed reduced hemispheric asymmetry in attentional functioning. In the present study, participants performed a modified version of the lateralized attention network test. Both early and late bilinguals were found to have more efficient executive network than monolinguals. The late bilinguals, who were also reported to be more balanced in the proficiency and usage of their two languages, showed the greatest advantage in conflict resolution, whereas early bilinguals seemed to show enhanced monitoring processes. These group differences were observed when controlling for non-verbal intelligence and socioeconomic status. Such results suggest that specific factors of language experience may differentially influence the mechanisms of cognitive control. Since the bilinguals had distinct language sets, it seems that the influence of bilingualism on executive functions is present regardless of the similarity between the two languages. As for hemispheric lateralization, although the results were not clear-cut, they suggest the reduced lateralization in early bilinguals.

Keywords: bilingualism, age of L2 acquisition, attentional networks, attention network test, lateralization

INTRODUCTION

A person who speaks two languages needs to attend to the language that is appropriate in the particular context and ignore the language that is irrelevant. This kind of experience may lead to development of more effective attentional mechanisms. Indeed, a substantive body of research has consistently shown benefits for bilinguals in some aspects of cognitive functioning, especially in executive control abilities (see Bialystok et al., 2009, for a review). The evidence for a bilingual advantage in tasks requiring resolution of conflict, or inhibition of non-relevant information, is consistent with the notion that bilinguals recruit the executive control system in order to manage the simultaneous activation of their two languages (Green, 1998; Abutalebi and Green, 2007; Kroll, 2008; van Heuven et al., 2008), as well as with the claim that the enhancement of such processes through continual practice may generalize to other domains of cognitive functioning (Bialystok et al., 2009; Ye and Zhou, 2009; Festman et al., 2010).

Although the impact of bilingualism on non-linguistic processes seems to be well-documented, many issues still remain open and there is a clear need to determine the boundaries of such influence. Hernandez et al. (2010) list two possible ways in which this goal may be achieved. The first is to identify the exact components of executive control that are modulated by bilingualism, and the second is to explore the extent to which bilingualism influences other aspects of attention. Both approaches seem to be of great importance since the relationship between bilingualism and executive functions appears to be more complex than initially claimed. Research practice shows that some of the effects indicating cognitive benefits in bilinguals are not always easy to replicate (cf. Bialystok et al., 2005b; Colzato et al., 2008), and the detectability of the bilingual advantage in conflict resolution may be limited to some specific experimental conditions (Colzato et al., 2008; Costa et al., 2009). Moreover, in order to fully understand the nature of the relationship between bilingualism and the reported cognitive gains, we also need to explore which aspects of bilingual experience are crucial for the effect to emerge (Kroll, 2009).

Factors that can potentially contribute to the emergence of the bilingual benefit include the speaker's language proficiency and relative balance between the two languages, the intensity of daily usage of each of the two languages, length of exposure, age of L2 acquisition, the degree of similarity between a bilingual's two languages, and specificities related to the context in which both languages are being used on a daily bases. The latter relates to whether the two languages are separated in time in daily use (one language at home, the other at work), or whether daily usage involves frequent mixing of languages. According to Costa et al. (2009), this sociolinguistic factor may impact on whether bilinguals show enhancement of the monitoring aspect of executive functions. Disentanglement of how each of the factors selectively contributes to the cognitive benefit is a challenging task. Let us consider briefly how two of the factors listed above may contribute to changes in attentional control in bilinguals.

SIMILARITY BETWEEN LANGUAGES

Although bilingual benefits in executive functions in children and the elderly have been replicated across different languages and cultural contexts, reports on similar advantages in young adults are scarce and so far limited to bilinguals with language sets that are relatively similar in terms of lexical and grammatical structure (mostly Catalan–Spanish; Costa et al., 2008, 2009; Hernandez et al., 2010; but see Bialystok et al., 2005a; see also **Table 1** for a review). It may therefore be that usage of two typologically similar languages requires a greater degree of attentional control, leading to more efficient executive and alerting networks. It remains an open question whether having two structurally distinct language sets will lead to similar advantages in young adults who are at the peak of their cognitive abilities. Costa et al. (2008) suggested that the need to monitor the two languages may be particularly strong in contexts in which bilinguals use their two languages interchangeably (such as Catalan–Spanish speakers in Barcelona) and less needed in the case of bilinguals who have a clear separation between the languages and daily activities.

AGE OF ACQUISITION

Does one need to be an early bilingual to enjoy the benefits of improved executive functions? Indeed, most research reporting cognitive gains in bilinguals examined bilinguals who learnt their two languages relatively early in life (see **Table 1**). The only two exceptions so far are studies by Bialystok et al. (2006) and Wodniecka et al. (2010). However, in both of these studies, the late bilinguals represented elderly participant groups. Does it mean that, at least for young adults, early age of L2 acquisition is necessary for the attentional benefit to emerge? If not, what are the critical conditions that must be fulfilled by late bilinguals to confer similar advantages? Age of acquisition might have an influence on bilinguals' efficiency in executive control not necessarily because of biological or maturational constraints on language learning, but because of a set of environmental factors that might be the consequence of early or late L2 learning. In the most obvious way, age of acquisition has an impact on (although it does not determine) the amount of input that one receives in each language. If the length of simultaneous exposure to two languages is critical for the cognitive advantage to emerge, then early bilinguals should naturally enjoy greater cognitive benefits than late bilinguals. On the other hand, there are grounds to suggest that late bilinguals may, in fact, train their executive control to a greater extent than early bilinguals and hence display a larger cognitive benefit related to the training. Abutalebi and Green (2007) demonstrated that bilingual language production engages the neural executive network to a greater extent than monolingual production, suggesting the importance of the network in selecting a language in the face of interference. Moreover, L2 processing is more effortful than L1 processing and involves more extensive activation in the left frontal region than processing of the same language by monolingual speakers (Wartenburger et al., 2003; Hernandez and Meschyan, 2006; Abutalebi, 2008; Kovelman et al., 2008). This seems to suggest that late bilinguals utilize the executive network to a greater extent than early proficient bilinguals, presumably because executive control not only helps them control interference from their other language, but also supports processing of the less automatic L2. A model developed by Hernandez et al. (2005) proposes that L2 learning involves a competitive interplay between a bilingual's two languages in which speakers must overcome interference from L1. The more solidified that L1 is, the more difficult L2 learning becomes. It seems plausible then that, although bilingualism may result in massive training of the executive network, late bilinguals are even more prone to this training due to greater interference of L1 during L2 learning.

Previous research has reported bilingual advantages in children and older adults, which are the two groups whose attentional capacities are either not fully developed or are in decline. The first study that demonstrated the effect of bilingualism on executive control in young adults in their twenties was carried out by Costa et al. (2008). The authors used the attention network test (ANT; originally developed by Fan et al., 2002) to compare the efficiency of three attentional networks in Catalan–Spanish bilinguals and Spanish monolinguals: alerting, orienting, and executive control. Attentional networks are a system of functionally and neuro-anatomically independent webs of neural areas, which are involved in three kinds of functions: achieving and maintaining an alert state, orienting to sensory or mental events, and monitoring and resolving competitions or conflicts (Raz and Buhle, 2006; Posner and Rothbart, 2007; Posner and Fan, 2008). Costa et al. (2008) found young adult bilinguals to be advantaged in conflict resolution. Moreover, bilinguals showed a larger alerting effect, but did not differ from monolinguals in their orienting of attention. In addition, bilinguals were overall faster than monolinguals in performing the task. In their subsequent ANT study with young adult bilinguals, Costa et al. (2009) argued that the overall reaction time (RT) advantage of bilinguals may indicate higher efficiency of the monitoring system, which evaluates the need to engage in conflict resolution processes. How exactly the monitoring and the conflict resolution processes interact with each other and to what extent they depend on one another is still an open issue (cf. Costa et al., 2009). Nevertheless, the available evidence suggests that bilingualism may impact various aspects of the cognitive control mechanism. A recent study by Marzecová, Asanowicz, Krivá, and Wodniecka (submitted for publication) replicated the results obtained by Costa et al. (2008) in relation to executive and alerting networks. However no overall advantage in RTs was observed. The results of Marzecová et al. (submitted for publication) suggest an advantage for bilinguals in conflict resolution *per se*, but not in the process of monitoring (cf. Costa et al., 2009). The participants in that study were early, relatively balanced speakers of two languages that are typologically similar to each other (mostly Czech–Slovak bilinguals). The participants tested in all three studies described above represented similar profiles: They were early bilinguals with life-long exposure to the two typologically similar languages. It is therefore impossible to determine which aspect of their experience was crucial for the attentional advantage that was observed. An important question stemming from previous research with the ANT task, then, is whether similar effects would be observed in a group of bilinguals whose two languages are more distinct from each other, and if so how the later age of acquisition would impact on the pattern of results.

Table 1 | Overview of past studies in which bilinguals and monolinguals were compared with regard to the efficiency of executive control. The table includes a summary of the data on language experience (age of acquisition, proficiency, percentage of usage, relative language balance) of participants if they were provided. The description of participants includes data on combinations of languages that were spoken by the bilingual participants. Such information is missing if groups of bilinguals were heterogeneous and spoke various sets of languages. The table summarizes the results on the efficiency of executive control and the advantage on overall RT across groups.

Study	Participants	Bilingual age of acquisition	Language proficiency	Bilingual language usage	Relative language balance	Task	Index of executive control effect	Difference between groups in global RTs	Difference in executive control index
Bialystok (2006)	(a) 19 video-game players BL (22.2 years) (b) 17 video-game players ML (21.6 years) (c) 30 non-video-game players BL (22.0 years) (d) 31 non-video-game players ML (22.0 years)	Early Before age 5	Self-rating (1–10): at least 6 (spoken L1)	L1 at home, L2 at school or work	Balanced	(i) Simon squares task (ii) Simon arrows task	(i) Incongruent vs. congruent trials (ii) Incongruent vs. congruent trials	(i) No (ii) Yes (in high-switch condition only)	(i) No (ii) No
Bialystok (2010)	Study 1 (a) 26 BL (6.0 years) (b) 25 ML (6.1 years) Study 2 (a) 25 BL (5.8 years) (b) 25 ML (5.8 years) Study 3 (a) 25 BL (6.1 years) (b) 25 ML (6.0 years)	Study 1, 2, 3 Early From birth or schooling	Study 1 PPVT-III score (English) (a) 104.2 (b) 107.8 Study 2 PPVT-III score (a) 104.3 (b) 105.0 Study 3 PPVT-III score (a) 100.9 (b) 107.7	Study 1 five-point scale (1 = mostly L1, 5 = mostly L2): 3.6 Study 2 five-point scale (1 = mostly L1, 5 = mostly L2): 3.2 Study 3 five-point scale (1 = mostly L1, 5 = mostly L2): 2.0	Study 1, 2, 3 Balanced	Study 1, 2, 3 (i) Trail-making task (ii) Global-local task	Study 1, 2, 3 (i) Trails B vs. Trails A (ii) Incongruent vs. congruent trials	Study 1, 2, 3 (i) Yes (ii) Yes	Study 1, 2, 3 (i) No (ii) No
Bialystok et al. (2005a)	(a) 10 French–English BL (b) 9 Cantonese–English BL (c) 10 English ML (29.0 years)	Early From early childhood	Bilinguals' L2 fluency equivalent to monolinguals	L1 at home, L2 at school	Balanced	Simon task	Incongruent vs. congruent trials	Yes (for Cantonese–English BL vs. ML only)	No
Bialystok et al. (2004)	Study 1 (a) 10 mid-age Tamil–English BL (43.0 years) (b) 10 mid-age English ML (43.0 years) (c) 10 older Tamil–English BL (72.3 years) (d) 10 older English ML (71.6 years)	Study 1 Later in childhood From age 6 Study 2 Later in childhood From age 6	Study 1 PPVT-R score: (English) (a) 91.8 (b) 91.0 (c) 91.9 (d) 85.8	Study 1 44.0% (L1) 56.0% (L2) Study 2 51.7% (L1) 48.3% (L2) Study 3 50.0% (L1) 50.0% (L2)	Study 1, 2, 3 Balanced	Study 1, 2, 3 Simon task	Study 1, 2, 3 Incongruent vs. congruent trials	Study 1 Yes Study 2 Yes (in three conditions only) Study 3 Yes (in blocks 1–7 only)	Study 1 Yes Study 2 Yes Study 3 Yes (in blocks 1–4, 8–9 only)

(Continued)

Table 1 | Continued

Study	Participants	Bilingual age of acquisition	Language proficiency	Bilingual language usage	Relative language balance	Task	Index of executive control effect	Difference between groups in global RTs	Difference in executive control index
	Study 2 (a) 32 mid-age Tamil–English (20) or Cantonese–English (12) BL (42.6 years) (b) 32 mid-age English ML (42.6 years) (c) 15 older English–Tamil (9) or English–French (6) BL (70.2 years) (a) 15 older English ML (70.4 years) Study 3 (a) 10 French–English BL (40.6 years) (b) 10 ML English (38.8 years)	Study 3 Early From childhood	Study 2 PPVT-III score: (a) 86.0 (b) 85.4 (c) 81.4 (d) 79.7 Study 3 PPVT-III score: (a) 91.0 (a) 89.1						
Bialystok et al. (2008)	(a) 24 young BL (19.7 years) (b) 24 young ML (20.7 years) (c) 24 older BL (68.3 years) (d) 24 older ML (67.2 years)	(a) Early Before age 6 (c) Late Before age 20	Self-rating (0–4): (a) 3.15 (L1) 3.83 (L2) (c) 3.65 (L1) 3.79 (L2)	Used both L1 and L2 daily	Balanced	(i) Simon arrows task (ii) Stroop color-naming task	(i) Incongruent vs. congruent trials (ii) Incongruent vs. congruent trials	(i) No (ii) No	(i) No (ii) Yes
Bialystok et al. (2006)	Study 1 (a) 24 young BL (20.8 years) (b) 24 young ML (20.7 years) (c) 24 older BL (71.3 years) (d) 24 older ML (70.4 years) Study 2 a) 24 young BL (23.9 years) b) 24 young ML (25.6 years) c) 24 older BL (64.5 years) d) 24 older ML (66.9 years)	Study 1, 2 a) Early About age 6 c) Late About age 12	Not reported	Study 1, 2 L1 at home, L2 at school or work	Study 1, 2 Balanced	Study 1, 2 Faces/ modified anti-saccade task	Study 1, 2 (i) Response suppression: red vs. green eye trials (ii) Inhibitory control: conflicting gaze vs. supporting gaze trials	Study 1 No Study 2 Yes (for older BL vs. older ML, in three conditions only)	Study 1 (i) No (ii) No Study 2 (i) Yes (ii) Yes (for older BL vs. older ML, in RTs only)
Bialystok and Martin (2004)	Study 1 (a) 31 Chinese–English BL (4.9 years) (b) 36 English ML (4.9 years)	Study 1, 2, 3 Early From birth	Study 1 PPVT-R score: (English) (a) 87.8 (b) 112.2 Study 2	Study 1 L1 at home, L2 at school and in the community	Study 1, 2, 3 Balanced	Study 1 Computerised dimensional change card sort	Study 1, 2, 3 Number of correct post-switch trials	RTs not reported	Study 1, 2, 3 Yes

	Study 2 (a) 15 French–English BL (4.6 years) (b) 15 English ML (5.1 years) Study 3 (a) 26 Chinese–English BL (4.4 years) (b) 27 English ML (4.2 years)	PPVT-R score: (a) 89.6 (b) 110.8 Study 3 PPVT-R score: (a) 84.3 (b) 109.7		Study 2 Both L1 and L2 at home, L2 at school Study 3 L1 at home, L2 at school and in the community		Study 2, 3 Manual dimensional change card sort			
Bialystok et al. (2005a)	Study 1 (a) 17 French–English BL (b) 17 English ML (5 years) Study 2 (a) 18 French–English BL (b) 22 English ML (5 years) Study 3 (a) 56 BL (b)40 ML (20–30 years) Study 4 (a) 10 mid-age BL in India (b) 10 mid-age ML in Canada (30–59 years) (c) 10 older BLin India (d) 10 older ML in Canada (60–80 years)	Study 1, 2, 3 Early From birth Study 4 Later in childhood	Study 1 PPVT-R score: (English) ML > BL EVIP score: L1 proficiency equivalent to L2 Study 2 PPVT-R score: ML > BL Study 3 Not reported Study 4 PPVT-R score ML = BL	Study 1, 2, 3 L1 at home, L2 in the community Study 4 Used both L1 and L2 daily	Study 1, 2, 3, 4 Balanced	Study 1, 2 Simon task for children Study 3 Simon task with control condition Study 4 Simon task from Study 1, 2	Study 1, 2, 3, 4 Incongruent vs. congruent trials	Study 1, 2 Yes Study 3 No Study 4 Yes	Study 1, 2, 3, 4 No
Bialystok and Viswanathan (2009)	(a) 30 BL in Canada (8.5 years) (b) 30 BL in India (8.6 years) (c) 30 ML in Canada (8.5 years)	Early From birth	PPVT-III score: (English) (a) 100.0 (b) 96.3 (c) 111.9	(a) 46% (L1) 54% (L2) (b) 53% (L1) 47% (L2)	Balanced	Faces/ modified anti-saccade task	(i) Response suppression: red vs. green eye trials (ii) Inhibitory control: gaze shift vs. straight eye trials	No	(i) No (ii) Yes
Carlson and Meltzoff (2008)	(a) 12 English–Spanish BL (6.0 years) (b) 21 English–Spanish (13) or Japanese–English (8) BL in language immersion (5.8 years) (c) 17 English ML (6.3 years)	(a) Early From birth (b) Later in childhood	EOWPVT-SBE score: (a) 100.0 (b) 136.0 (c) 135.0	(a) Both L1 and L2 at home, L2 with friends (b) L1 at home and half day at school, L2 for half day at school	(a) Balanced (b) Dominant in L1	(i) Children's attention network test (ii) Advanced dimensional change card sort	(i) Incongruent vs. congruent trials (ii) Number of correct conflict trials	RTs not reported	(i) No (ii) Yes

(Continued)

Table 1 | Continued

Study	Participants	Bilingual age of acquisition	Language proficiency	Bilingual language usage	Relative language balance	Task	Index of executive control effect	Difference between groups in global RTs	Difference in executive control index
Colzato et al. (2008); Experiment 1	(a) 16 Dutch–English BL (22 years) (b) 16 Spanish ML (22 years)	Early From birth	Self-rating (1–10): 8.9	Used both L1 and L2 daily	Balanced	Stop signal task	Inhibition of response	No	No
Costa et al. (2009)	Experiment 1 (a) 60 Catalan–Spanish BL (20.1 years) (b) 60 Spanish ML (20.0 years) Experiment 2 (a) 62 Catalan–Spanish BL (20.1 years) (b) 62 Spanish ML (20.7 years)	Experiment 1, 2 Early From early childhood	Not reported	Experiment 1 seven-point scale (1 = only L2, 7 = only L1): 5.3 Experiment 2 Seven-point scale (1 = only L2, 7 = only L1): 5.0	Experiment 1, 2 Balanced	Experiment 1 Attention network test (i) 8% congruent (ii) 92% congruent Experiment 2 Attention network test (i) 50% congruent (ii) 75% congruent	Experiment 1, 2 Incongruent vs. congruent trials	Experiment 1 (i) No (ii) No Experiment 2 (i) Yes (ii) Yes (in block 1 only)	Experiment 1 (i) No (ii) No Experiment 2 (i) No (ii) Yes (in block 1 only)
Costa et al. (2008)	(a) 100 Catalan–Spanish BL (22 years) (b) 100 Spanish ML (22 years)	Early From early childhood	Self-rating (1–4): 4.0 (L1) 3.9 (L2)	Seven-point scale (1 = only L2, 7 = only L1): 5.1	Balanced	Attention network test	Incongruent vs. congruent trials	Yes	Yes (in blocks 1–2 only)
Emmorey et al. (2008)	(a) 15 English ML (50.1 years) (b) 15 bimodal BL (46.2 years) (c) 15 unimodal BL (47.0 years)	(b) Early 0.9 years c) Later in childhood 6.1 years	Self-rating (1–5): (b) 4.5 (L1) 4.6 (L2) (c) 4.5 (L1) 3.1 (L2)	Used both L1 and L2 daily	(b) Balanced (c) Dominant in L1	Flanker task with baseline, neutral, congruent, and incongruent conditions	Incongruent vs. congruent trials	Yes (for unimodal bilinguals only)	No
Hernandez et al. (2010); Experiment 1	(a) 41 Catalan–Spanish BL (20.9 years) (b) 41 Spanish ML (21.4 years)	Early From early childhood	Not reported	Seven-point scale (1 = only L2, 7 = only L1): 5.1	Balanced	Numerical Stroop task	(i) Stroop interference: incongruent vs. neutral trials	Yes-tendency p = 0.061	(i) Yes (ii) Yes

Luk et al. (2010)	(a) 10 BL (20 years) (b) 10 ML (22 years)	Later in childhood From age 6	Self-rating (1–10): 7.1 (L1) 7.8 (L2) PPVT-III score: (English) (a) 94.8 (b) 105.8	Used both L1 and L2 regularly	Balanced	Flanker task with baseline, neutral, congruent, and incongruent conditions	(ii) Stroop facilitation: neutral vs. congruent trials Incongruent vs. congruent trials	No	No
Martin-Rhee and Bialystok (2008)	Study 1 (a) 17 French–English BL (5.0 years) (b) 17 English ML (4.7 years) Study 2 (a) 21 BL (4.6 years) (b) 20 ML (4.5 years)	Study 1, 2 Early From birth	Study 1 PPVT-R score: (a) 89.6 (L1) (b) 111.4 EVIP score: (a) 98.8 (L2) Study 2 PPVT-R score: (a) 86.4 (b) 96.4	Study 1, 2 Both L1 and L2 at home, L2 at school	Study 1, 2 Balanced	Study 1 Simon task (i) Immediate response (ii) Short delay (iii) Long delay Study 2 (i) Simon task (ii) Stoop picture naming task	Study 1, 2 Incongruent vs. congruent trials	Study 1 (i) Yes (ii) No (iii) No Study 2 (i) Yes (ii) No	Study 1 (i) No (ii) No (iii) No Study 2 (i) No (ii) No
Marzecová et al. (submitted for publication)	(a) 18 BL (23.5 years) (b) 17 ML (20.0 years)	Early Before age 4	Self-rating (1–7): 6.9 (L1) 6.3 (L2)	53% (L1) 32% (L2) 15% (L3)	Balanced	Lateralised attention network test	Incongruent vs. congruent trials	No (but overall advantage in ERR)	Yes (in RT only; tendency in ERR, p = .08)
Morton and Harper (2007)	(a) 17 French–English BL (6.9 years) (b) 17 English ML (6.9 years)	Early From birth Early Before age 6	PPVT-R score: (a) 100.3 (L1) (b) 110.1 EVIP score: (a) 97.8 (L2)	58.3% (L1) 41.7% (L2)	Balanced	Simon task	Incongruent vs. congruent trials	No	No
Prior and MacWhinney (2010)	(a) 44 BL (19.5 years) (b) 44 English ML (18.7 years)		Self-rating (1–10): (a) 7.8 (L1) 9.3 (L2) (b) 9.3 (L1) 3.1 (L2) PPVT-III score: (a) 102.30 (L2) (b) 109.95 (L1)	(a) 27% (L1) 73% (L2) (b) 97% (L1) 3% (L2)	Dominant in L2	Task switching paradigm (a) Color task (b) Shape task	(i) Switching cost: switch trials vs. non-switch trials (ii) Mixing cost: mixed task blocks vs. single task blocks	Not reported	(i) Yes (in RT only) (ii) No

(Continued)

Table 1 | Continued

Study	Participants	Bilingual age of acquisition	Language proficiency	Bilingual language usage	Relative language balance	Task	Index of executive control effect	Difference between groups in global RTs	Difference in executive control index
Soveri et al. (2010)	(a) 17 mid-age Finnish–Swedish BL (40.1 years) (b) 18 mid-age Finnish ML (38.5 years) (c) 16 older Finnish–Swedish BL (66.0 years) (d) 14 older Finnish ML (67.6 years)	Early Before age 7	Self-rating (0–6): (a) 5.8 (L1) 5.7 (L2) (b) 5.9 (L1) 3.2 (L2) (c) 5.5 (L1) 5.8 (L2) (d) 5.9 (L1) 3.7 (L2)	Used both L1 and L2 actively throughout life	Balanced	Forced-attention dichotic listening task	Identification of targets presented to either left (forced-left condition) or right (forced-right condition) ear	RTs not reported	Yes
Present study	(a) 36 early Chinese–English BL (18.9 years) (b) 30 late Chinese–English BL (20.8 years) (c) 34 English ML (20.4 years)	(a) Early 0.3 years (b) Late 16.2 years	Self-rating (1–7): (a) 3.6 (L1) 6.6 (L2) (b) 6.8 (L1) 4.9 (L2)	(a) 25% (L1) 75% (L2) (b) 59% (L1) 40% (L2) (c) 1% (L3)	(a) Strongly dominant in L2 (b) Moderately dominant in L1	Lateralised attentional network task	Incongruent vs. congruent trials	Yes (for early BL vs. ML only)	Yes (in RT only for early BL vs. ML; in ERR only for early vs. late BL; in both RT and ERR for late BL vs. ML)

BL, bilinguals; ML, monolinguals; PPVT-III, peabody picture vocabulary test – Third Edition; PPVT-R, peabody picture vocabulary test – Revised; EVIP, echelle vocabulaire en images peabody; EOWPVT-SBE, expressive one-word picture vocabulary test – Spanish bilingual edition.

The present study aimed at comparing early and late bilinguals in the efficiency of the three attentional networks, alerting, orienting, and executive control (cf. Posner and Rothbart, 2007; Posner and Fan, 2008), and thus to shed some light on the interaction between age of L2 acquisition and cognitive gains associated with bilingualism. We asked which aspects of attentional functions are modulated by early and late bilingualism and to what extent the L2 age of acquisition has a differential impact on attention. We aimed at extending earlier findings on attentional advantage observed in early young bilinguals by including a group of late bilinguals, who acquired L2 in their adolescent years. Additionally, the bilinguals in the current study were a linguistically homogeneous group whose two languages were very distinct from each other; namely, Chinese and English. This allowed for comparison between the results from the current study with earlier research with bilinguals who spoke two very similar languages.

An alphabetic language like English has a phonemically based script, which is a system of letters that each represents a unit of sound (phoneme). The Chinese language, in contrast, has a morphosyllabic script, which is a system of characters each representing a unit of semantic meaning (morpheme) and having little systematic correspondence to phonology. In order to vocalize and comprehend Chinese, one must memorize the phonology and meaning of each character (Chee et al., 1999; Luk and Bialystok, 2008). Further, neuroimaging evidence suggests that different cognitive and processing resources may be required for reading and understanding Chinese as opposed to English (e.g., Tan et al., 2003), which may result in a lesser conflict between the two languages. If the advantage of bilingualism is related to the two language systems being similar to each other, then the Chinese–English bilinguals, both early and late, would show no advantage over monolinguals in the functioning of attentional networks. On the other hand, if the typological distance between the two languages of a bilingual does not play a role, then the early Chinese–English bilinguals may show advantages in alerting and executive networks, corroborating previous findings with language sets that are more similar to each other (Costa et al., 2008; Marzecová et al., submitted for publication). Such a finding would reveal that bilingual experience influences attentional functioning irrespective of the degree of similarity between languages, and would be consistent with previous studies on Chinese–English bilingual children (Goetz, 2003; Bialystok and Martin, 2004). If advantages in attentional functioning were to be found in early Chinese–English bilinguals, late bilinguals may or may not show similar pattern. If the constant practice in monitoring and switching between languages since an early age is necessary to gain enhancement of attentional functioning, then it may not be observed in late bilinguals. If, however, late bilinguals train the executive network more intensively than early bilinguals (because they need to control interference from L1, which may be even greater than in early bilinguals, as well as to support the less automatized L2; Costa et al., 2009), then the late Chinese–English bilinguals may show even more enhanced efficiency of attention than early bilinguals, especially in conflict resolution processes.

Besides exploring the impact of bilingual experience on attentional networks, we sought to investigate pattern of lateralization in bilinguals' attentional functioning. The lateralization of attention in bilinguals was of interest since previous research indicated that bilingualism may modulate interhemispheric organization of the attentional networks, especially with regard to executive control (Marzecová et al., submitted for publication). To assess hemispheric asymmetries of three attentional networks (alerting, orienting, and executive control), we employed a lateralized attention network test (LANT; Greene et al., 2008). Reduced hemispheric asymmetry in bilinguals has previously been reported for language functions (Dehaene et al., 1997; Moreno et al., 2010), but also for non-verbal cognitive functions (Hausmann et al., 2004). Moreover, the study by Marzecová et al. (submitted for publication) showed right hemisphere dominance for conflict resolution in monolinguals, and a lack of such asymmetry in bilinguals. Both studies that reported the lack of hemispheric asymmetry tested early bilinguals (Hausmann et al., 2004; Marzecová et al., submitted for publication). It remains an open question, then, whether late onset bilinguals display a similar pattern of lateralization for non-linguistic functions as early bilinguals. The hemispheric asymmetry of linguistic processes appears to be influenced by the age of acquisition of the second language. Meta-analyses of language studies (Hull and Vaid, 2006, 2007) indicate that, regardless of proficiency, bilinguals who acquired their second language at an early age (typically before age six) consistently show more bilateral involvement in linguistic tasks for both L1 and L2. Late bilinguals, on the other hand, exhibit left hemispheric lateralization for both their languages, as is typically observed in monolinguals. Analogously, it is possible that early bilinguals would show reduced hemispheric asymmetry of attentional networks, while late bilinguals would show the same pattern of lateralization as monolinguals. If this is the case, it may be argued that early experience in learning a second language is critical in altering the functional cerebral organization of non-linguistic functioning.

MATERIALS AND METHODS

MATERIALS AND PROCEDURE

Background questionnaire

A language background questionnaire was used to obtain participant information in order to classify the bilinguals as either Early or Late. Demographics details were also collected to allow any major differences between groups, such as age, gender, and socioeconomic background, to be identified. In particular, lower socioeconomic status (SES) has been shown to be associated with deficits in aspects of attention, especially in tasks that require filtering information and managing response conflict (see e.g., Stevens et al., 2009). In the present study, parental occupation was used as an index of SES. Following the recommendations of McMillan (2010), specific occupations of the mother and father were coded using the Australian and New Zealand Standard Classification of Occupations (ANZSCO; Australia Bureau of Statistics/Statistics New Zealand, 2009), and then converted into a percentile score using the Australian Socioeconomic Index 2006 (AUSEI06; McMillan et al., 2009). The higher of the two parents' scores was retained as the SES score for each participant. Parental education level, determined as the average of the two parents' number of years of education, provided further information about socioeconomic background, as parental education is a good predictor of SES (see Marks et al., 2000).

In addition to sociodemographic details, bilingual participants provided information relating to language experience, so that effects of individual differences in factors such as proficiency and usage could be examined. Bilinguals rated their proficiency in both Chinese and English, separately for speaking, understanding speech, reading, and writing, using a seven-point scale (1 = Not at all; 7 = Native-like). Self-ratings were also provided for the amount of daily usage of each language (expressed in percentages), the frequency of mixing their two languages in the same sentence, on a five-point scale (1 = Rarely; 5 = Very frequently), and the frequency of deliberately refraining from uttering a Chinese word or phrase when speaking to an English speaker, on a five-point scale (1 = Rarely; 5 = Very frequently). The latter two self-ratings allowed differences in frequency of mixing and inhibiting to be examined. The questions pertaining to language experience were mostly adapted from questions in the L2 Language History Questionnaire (Li et al., 2006) and the Language Dominance Scale (Dunn and Fox Tree, 2009).

Lastly, the handedness of participants was established using a question adapted from the Edinburgh Handedness Inventory (Oldfield, 1971). Participants indicated whether they used their left hand for any of a list of eight activities. People who marked four or more activities were deemed to be left-handed.

Non-verbal intelligence test

In order to compare general non-verbal intelligence across the three groups, participants completed a shortened version of Raven's Advanced Progressive Matrices Set I (Raven et al., 1998). One point was given for each correct answer, with a maximum total of 12; the total score was used as an index of the person's general non-verbal intelligence.

Lateralized attention network test

The LANT is a computer-based task requiring manual responses to stimuli presented on screen. Stimuli were presented using the DMDX program (Forster and Forster, 2003). The main stimulus in the LANT comprised an array of five arrows, oriented and arranged in a vertical line. The middle arrow was the target and either pointed up or down. The target was flanked with other arrows that were either congruent, i.e., pointing in the same direction as the target; or incongruent, i.e., pointing in the opposite direction (see **Figure 1A**). The array of arrows subtended a height of 3.0° visual angle, and was presented at a distance of 2.2° visual angle from a central fixation cross to either left or right visual field. The stimuli were preceded by one of four types of cue: (i) a valid spatial cue, which was an asterisk presented in the same visual hemifield as the target; (ii) an invalid spatial cue, which was an asterisk presented in the opposite visual hemifield; (iii) a center cue, which was an asterisk presented at the location of the fixation cross; and (iv) no cue (see **Figure 1B**). For the spatial cue conditions, 80% were valid while the other 20% were invalid. The LANT procedure and stimuli parameters were based on the study by Greene et al. (2008).

Each trial of the LANT consisted of five events as follows: (i) a central fixation cross presented for a period of random variable duration (1300–2700 ms), ensuring that the onset of the target stimuli was predicted by the cue and not by the regular timing of the initial fixation period; (ii) a cue presented for 100 ms; (iii) a short fixation period for 400 ms; (iv) the target and flanker stimuli flashed randomly to either left or right of the fixation cross for 180 ms, in order to isolate the information to one hemisphere; and (v) a response period that ended once participants responded, or timed out after 1820 ms (see **Figure 1C**). The fixation cross remained on the screen throughout whole trial, until participants responded or the trial timed out. During the task, a chin rest with forehead bar was used to secure the position of the eyes of participants at a distance of 50 cm away from the center of the screen. Participants were instructed to keep their heads still and fixate on the central cross throughout the session. Their task was to respond as accurately and as quickly as possible to the direction of the target middle arrow, ignoring the four flanker arrows. First there were two 12-trial practice blocks, in which participants received feedback for the accuracy of response for each trial. Two experimental blocks then followed, each consisting of 144 trials. Within each block, trials were presented in a randomized order. The number of trials was divided equally across the two flanker types, as well as across the two visual hemifields. Responses were made on a mouse held sideways, so that the two buttons were oriented vertically. Response hand alternated between blocks in a counterbalanced order across participants.

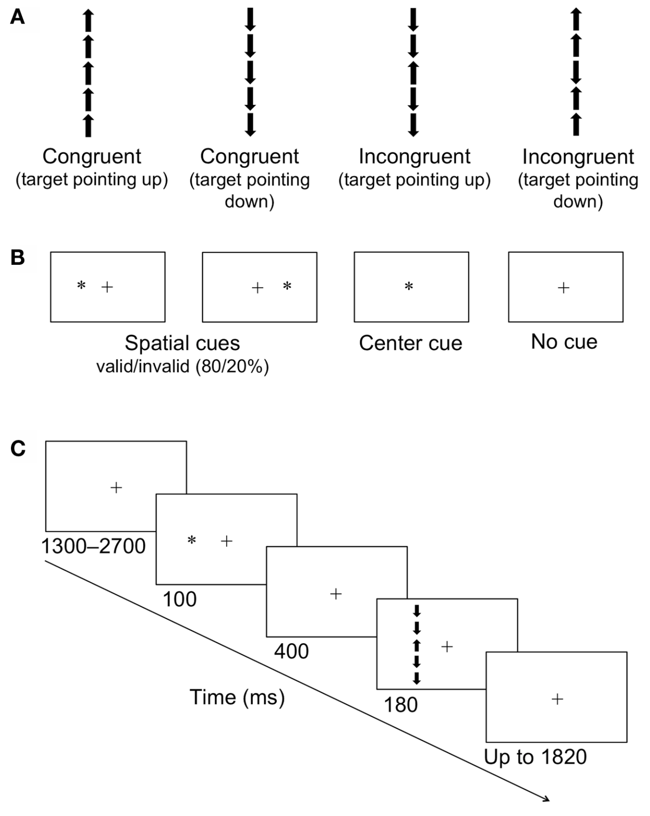


FIGURE 1 | Schematic representation of targets and flankers (A), types of cues (B), and events in the lateralized attention network test [LANT (C)].

The LANT provides indices for the efficiency of alerting, orienting, and executive networks (cf. Fan et al., 2002). Subtracting RT or accuracy in the center cue condition from no cue condition allows the efficiency of the alerting network to be measured. Typically, performance is much improved after occurrence of the center (warning) cue, which signals when target will appear next (Posner, 2008). Comparison between the results in a valid spatial cue condition (which informs participants where the target will occur) and the results in a center cue condition provide information about the efficiency of orienting to the target location. If a target is preceded by a valid spatial cue, responses are faster and more accurate, since attention is already focused on the target location (Posner, 1980). Additionally, orienting cost can be examined by comparing an invalid spatial cue condition with a center cue condition. The orienting cost reflects the efficiency of reorienting to the target presented outside the current focus of attention (Corbetta et al., 2008). Finally, a comparison between congruent and incongruent flanker conditions shows the cost of conflict resolution, which is an index of the executive network's efficiency. In order to respond quickly and accurately to a target in the incongruent condition one must inhibit the interference and resolve the conflict caused by flankers, which are incongruent with the target (Eriksen and Eriksen, 1974; Fan et al., 2003).

PARTICIPANTS

A total of 100 people participated in this study, each belonging to one of three groups: Early Bilinguals (n = 36), Late Bilinguals (n = 30), and Monolinguals (n = 34). Overall, the age of participants ranged from 18 to 48 years (M = 20.0, SD = 3.7); there were 56 females and 44 males. The participants were students undertaking a first year psychology course at the University of New South Wales, who received course credit in exchange for participation. The study was approved by the Human Research Ethics Advisory Panel of the School of Psychology at the UNSW and participants provided written informed consent prior to participation.

No left-handed participants were tested, as patterns of cerebral lateralization have been found to vary more in left-handers (Andreou and Karapetsas, 2001). The Chinese–English bilingual groups included both Mandarin and Cantonese speakers, as there is essentially no difference in structure between these two dialects. The age of arrival to Australia (or to another English-speaking country) was considered as the age of L2 exposure, thus the age of immersion in the L2 environment was considered as the age of onset of bilingualism and used in the classification of bilinguals as either Early or Late. **Table 2** presents the sociodemographic characteristics for the three groups, along with their Raven's non-verbal intelligence scores. The language characteristics of the two bilingual groups are also presented.

Monolinguals

All of the participants in the Monolingual group were born, and had spent most of their lives, in Australia or other English-speaking countries. All were of Caucasian descent. People whose parents spoke other languages were excluded, as they may have had some understanding and/or ability to communicate in a second language.

Early bilinguals

The Early Bilingual group consisted of those who had arrived in Australia at or before age six. The average age of arrival for the group was 0.3 years, due to the large majority (30) being born in Australia (i.e., age of arrival 0 years). For the six who were not born in Australia, the average age of arrival was 2.1 years. Given the average ages of first learning English and first being able to communicate in English (see **Table 2**), it can be assumed that most of the Early Bilinguals learned Chinese as their first language. Nearly all (32) of the participants in this group indicated a higher level of proficiency for English compared to Chinese, and a higher percentage of daily use of English over Chinese. Every Early Bilingual had received all of their formal education in English.

Table 2 | Characteristics of participant groups (SD in parentheses).

	Group		
Characteristic	**Early**	**Late**	**Mono**
Age	18.9 (1.3)	20.8 (2.5)	20.4 (5.5)
Gender (F:M)	19:17	19:11	18:16
Socioeconomic status (percentile score)	48.6 (24.2)	62.3 (21.3)	77.3 (17.0)
Parental education (years)	11.8 (3.7)	13.7 (3.4)	14.7 (3.2)
Non-verbal intelligence score (out of 12)	9.1 (2.3)	8.2 (3.0)	6.9 (3.0)
Age of first learning L2	2.9 (1.8)	7.8 (3.7)	–
Age of first able to communicate in L2	4.0 (1.7)	12.3 (4.7)	–
PROFICIENCY IN L1 (SEVEN-POINT SCALE)			
Speaking	4.9 (0.9)	6.7 (0.7)	–
Understanding	5.2 (1.0)	6.7 (0.6)	
Reading	2.4 (1.2)	6.6 (0.7)	
Writing	2.1 (1.1)	6.3 (0.9)	
PROFICIENCY IN L2 (SEVEN-POINT SCALE)			
Speaking	6.8 (0.4)	4.9 (1.0)	–
Understanding	6.9 (0.3)	5.0 (0.9)	
Reading	6.8 (0.4)	5.1 (0.9)	
Writing	6.7 (0.6)	4.7 (0.9)	
Percentage use of L1 and L2 (L1:L2)	25:75	59:40	–
Frequency of mixing L1 and L2 (five-point scale)	3.0 (1.4)	3.4 (1.0)	–
Frequency of inhibiting L1 (five-point scale)	1.6 (1.1)	2.1 (1.2)	–

Early = early bilinguals; late = late bilinguals; mono = monolinguals.

Late bilinguals

The Late Bilingual group consisted of those who had arrived in Australia at or after age 12. The average age of arrival in Australia for the group was 16.2 years (ranging from 12 to 19 years). As can be seen in **Table 2**, the Late Bilinguals first learned and were first able to communicate in English at substantially later ages than Early Bilinguals. Further, in contrast to the Early Bilinguals, the majority (24) of the Late Bilinguals indicated a higher level of proficiency for Chinese over English, and about equal or higher percentage of daily use for Chinese over English. Most had received more years of education in Chinese than in English (on average 9.7 and 4.1 years respectively).

Between-group comparisons

The three groups differed from each other in SES, parental education, and non-verbal intelligence. The differences in SES were significant between all three groups. Early Bilinguals had a lower average

SES than both Late Bilinguals and Monolinguals, $F(1, 97) = 32.37$, $p < 0.001$ and $F(1, 97) = 6.88$, $p = 0.010$ respectively, while Late Bilinguals had a lower score than Monolinguals, $F(1, 97) = 8.09$, $p = 0.005$. For parental education, the same trend across the three groups was observed as for SES, although only the comparison between Early Bilinguals and Monolinguals was statistically significant, $F(1, 97) = 12.02$, $p = 0.001$. For non-verbal intelligence score, the Early Bilinguals had the highest average score and the Monolinguals the lowest, but only the comparison between those two groups was statistically significant, $F(1, 97) = 11.15$, $p = 0.001$.

The comparisons between Early and Late Bilinguals on language characteristics revealed that Late Bilinguals had higher proficiency in L1 for each of the four language subskills (speaking, understanding, reading, and writing) than Early Bilinguals, smallest $t(64) = 9.07$, $p < 0.001$. Late Bilinguals also had greater percentage of use in L1 compared to Early Bilinguals, $F(1, 64) = 63.47$, $p < 0.001$. On the other hand, Early Bilinguals indicated higher proficiency in L2 for each of the subskills than Late Bilinguals, smallest $t(64) = 10.63$, $p < 0.001$, and greater percentage of use in L2, $F(1, 64) = 63.74$, $p < 0.001$. There were differences between the two groups, though not statistically significant, in the frequency of mixing, $F(1, 64) = 2.04$, $p = 0.159$, and inhibiting, $F(1, 64) = 2.46$, $p = 0.122$, where Late Bilinguals showed a greater average frequency in both (see **Table 2**).

RESULTS

OVERALL RESULTS AVERAGED ACROSS THREE LANGUAGE GROUPS

Trials with RTs faster than 200 ms or slower than 1200 ms (overall 1.8%) and trials with errors were excluded from the RT analysis. The mean RT was 648 ms (SD = 112.6). The mean error rate (ERR) yielded 12.6% (SD = 17.6). The mean RTs and ERRs broken by all conditions are presented in **Table 3**. The RT and ERR data were first analyzed by means of a 4 (cue condition: no cue, valid spatial, invalid spatial, center) × 2 (flanker type: congruent, incongruent) × 2 (visual field: left, right) ANOVA. The main effects of cue condition were significant both for RT, $F(3, 297) = 360.73$, $p < 0.0001$, and ERR, $F(3, 297) = 85.11$, $p < 0.0001$. The main effects of flanker type were also significant for RT, $F(1, 99) = 364.97$, $p < 0.0001$, and for ERR, $F(1, 99) = 169.92$, $p < 0.0001$. Importantly, the visual field asymmetry was found both for RT and ERR. Responses were 6 ms faster, $F(1, 99) = 4.77$, $p = 0.03$, and 2.4% more accurate, $F(1, 99) = 17.22$, $p < 0.001$, for targets presented in the left visual field (LVF) than in the right visual field (RVF). We also found significant cue × VF interaction for RT, $F(3, 297) = 4.24$, $p = 0.006$, showing the largest asymmetry in the invalid cue condition, and flanker type × VF interaction for ERR, $F(1, 99) = 22.31$, $p < 0.001$, which showed the LVF advantage (5%) in the incongruent condition and no asymmetry in the congruent condition. Description of other significant interactions obtained in the task goes beyond the research goals presented in this paper.

Attentional networks

The *alerting effect* was indexed by the difference between the center cue condition and no cue condition. Participants, averaged across three groups, responded 40 ms (SD = 28.7) faster on trials with a center cue than on trials with no cue, $t(99) = 14.01$, $p < 0.0001$, and made 2.7% (SD = 6.7) fewer errors, $t(99) = 4.03$, $p < 0.0001$. The *orienting benefit effect* was calculated by subtracting the RT or ERR of trials with a valid spatial cue from trials with a center cue. Participants took great benefit of a valid spatial cue, responding 67 ms (SD = 29.4) faster, $t(99) = 22.81$; $p < 0.0001$, and 4.6% (SD = 5.4) more accurately, $t(99) = 8.52$; $p < 0.0001$, than on trials with a center cue. The *orienting cost* was calculated by subtracting the RT and ERR of trials with a center cue from trials with an invalid

Table 3 | Mean reaction times of correct responses and error rates for all conditions.

Cue condition	Flanker type	VF	Reaction times						Error rates					
			Monolinguals		Early bilinguals		Late bilinguals		Monolinguals		Early bilinguals		Late bilinguals	
			RT (ms)	SD	RT (ms)	SD	RT (ms)	SD	ERR(%)	SD	ERR (%)	SD	ERR (%)	SD
No cue	Congruent	Left	670.5	83.0	603.8	68.3	657.7	94.2	2.6	3.8	2.1	4.0	670.5	83.0
		Right	674.3	84.9	596.0	73.2	655.9	96.6	3.3	5.2	4.0	6.7	674.3	84.9
	Incongruent	Left	757.5	105.1	681.5	84.6	729.8	86.4	25.0	17.1	19.4	17.8	757.5	105.1
		Right	762.8	117.3	680.3	79.3	712.6	82.4	29.0	21.5	22.4	16.9	762.8	117.3
Spatial valid	Congruent	Left	568.8	74.1	509.5	69.1	538.1	77.6	2.9	4.3	3.0	6.0	568.8	74.1
		Right	578.9	80.6	510.2	68.0	539.9	79.1	2.6	3.6	2.7	4.7	578.9	80.6
	Incongruent	Left	650.1	86.0	566.1	78.4	584.2	80.6	9.0	11.3	6.3	6.0	650.1	86.0
		Right	660.3	100.2	578.2	93.8	601.7	83.2	12.9	15.7	10.2	11.9	660.3	100.2
Spatial invalid	Congruent	Left	679.2	91.1	606.4	89.3	666.3	97.6	6.3	8.8	6.9	16.5	679.2	91.1
		Right	692.9	75.7	620.3	95.8	688.0	109.3	4.8	8.2	5.9	9.7	692.9	75.7
	Incongruent	Left	764.0	99.5	690.4	108.8	724.5	95.0	28.7	20.3	29.9	25.9	764.0	99.5
		Right	779.6	101.5	695.4	116.7	757.6	113.1	40.4	23.2	34.7	22.6	779.6	101.5
Center	Congruent	Left	619.4	77.6	561.4	73.5	599.5	82.3	2.9	5.2	3.0	4.8	619.4	77.6
		Right	627.7	79.3	552.9	69.3	583.3	81.5	2.4	4.9	3.5	6.1	627.7	79.3
	Incongruent	Left	736.9	90.3	647.1	80.0	684.5	79.0	19.5	16.8	15.3	15.0	736.9	90.3
		Right	742.8	103.1	647.7	98.4	693.0	90.1	25.4	21.4	17.7	19.5	742.8	103.1

spatial cue. The cost of 55 ms on speed (SD = 43.0) and 8.7% on accuracy (SD = 9.8) were both significant, $t(99) = 12.8$, $p < 0.0001$, and $t(99) = 8.8$, $p < 0.0001$ respectively. The *conflict effect*, indexed by the difference between congruent and incongruent flanker conditions, yielded 80.7 ms (SD = 42.2) for RT, $t(99) = 19.10$, $p < 0.0001$, and 15.9% (SD = 12.2) for ERR, $t(99) = 13.03$; $p < 0.0001$. All attention network indexes were similar for RT and ERR measurement; hence no speed–accuracy trade-off was observed.

BETWEEN-GROUP COMPARISONS

Because the groups differed on SES, parental education, and non-verbal intelligence, it is important to ensure that any differences between the groups in attentional functioning were not due to pre-existing differences other than language background. Therefore, in all between-group analyses presented below, the parental education and Raven's non-verbal intelligence were included as covariates. SES was not controlled for as a third covariate due to its highly significant correlation with parental education ($r = 0.51$, $p < 0.001$). Parental education was chosen over SES as the more objective measure of potential environmental influence since (a) SES had a moderate significant correlation with Raven's score ($r = -0.28$, $p = 0.005$), while parental education did not ($r = -0.05$, $p = 0.61$); and (b) some of the responses for parental occupation were too vague to properly classify and several of the parents were retired, making education level a more objective measure.

Overall RT and ERR

The results of ANCOVA showed that groups differed significantly on overall RT, $F(2, 95) = 4.59$, $p = 0.012$, when controlling for parental education and intelligence. Early bilinguals responded the fastest (609 ms), monolinguals exhibited the longest time of reactions (685 ms), while RT of late bilinguals fell in between (651 ms). Subsequent tests showed that the difference between monolinguals and early bilinguals was significant, $F(1, 66) = 11.15$, $p = 0.001$, whereas the differences between monolinguals and late bilinguals, and between early and late bilinguals were not significant: $F(1, 62) = 1.76$, $p = 0.19$, and $F(1, 62) = 2.35$, $p = 0.13$, respectively. There was no significant group effect, $F < 1$, on the overall ERR measure.

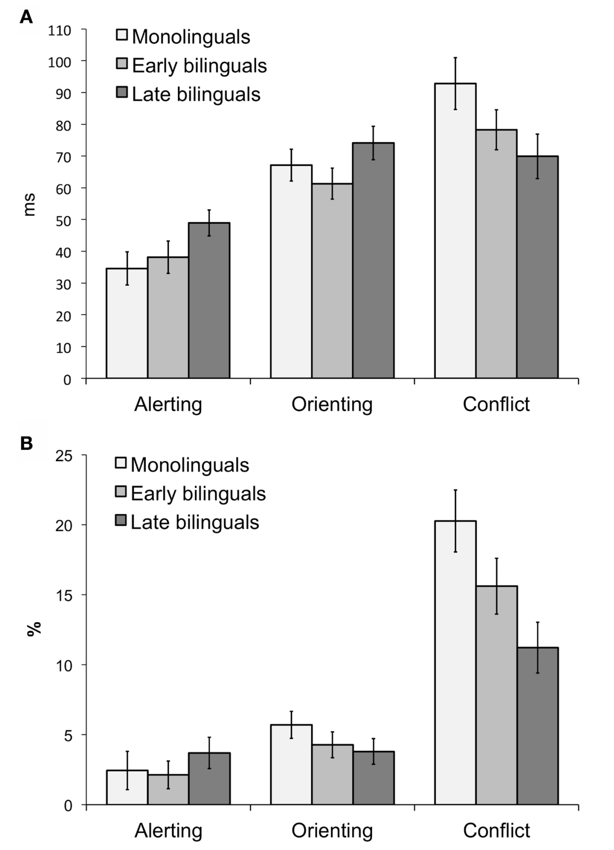


FIGURE 2 | Attentional networks in terms of RT (A) and ERR (B).

Attentional networks

Alerting. Monolinguals showed the smallest alerting affect (34.6 ms), the intermediate result was obtained for early bilingual group (38 ms), and late bilinguals obtained the largest effect (49 ms); see **Figure 2** for the attentional network indexes in the three groups. This trend is in line with previous studies, in which bilingual advantage in the alerting network was observed (Costa et al., 2008; Marzecová et al., submitted for publication). However, the effect did not reach significance, $F(1, 95) = 2.19$, $p = 0.12$. In the ERR analysis, the effect of group was not significant ($F < 1$).

Orienting. The three groups did not differ significantly in the orienting benefit effect, either in terms of RT, $F(2, 95) = 1.34$, $p = 0.27$, or ERR, $F < 1$. For the orienting cost, the late bilinguals showed the greatest cost (69 ms), while the early bilinguals and the monolinguals showed notably lesser costs (51 and 47 ms, respectively). However, the trend did not reach significance, $F(2, 95) = 2.2$, $p = 0.11$. To further investigate the effect of orienting cost in RT, three between-group comparisons were carried out: monolinguals vs. early bilinguals, monolinguals vs. late bilinguals, and early vs. late bilinguals. The difference between late bilinguals and monolinguals was significant, $F(1, 60) = 4.55$, $p = 0.037$, but the other two comparisons were not. Also, the ERR analysis for the three groups did not reveal any significant differences, $F(2, 95) = 1.09$, $p = 0.34$.

Conflict. Crucially, the three groups differed in the efficiency of the executive network. The late bilinguals were found to be most efficient in resolution of conflict, with the cost of 69.8 ms in terms of RT and 11.2% in terms of ERR. The conflict cost in the early bilingual group was 78.2 ms for RT and 15.6% for ERR. The largest effect was observed in the monolingual group: participants in this group were 92.8 ms slower and 20.2% less accurate in the conflict than in the non-conflict trials. The main effect of group was significant for both RT, $F(2, 95) = 3.06$, $p = 0.051$, and ERR, $F(2, 95) = 3.76$, $p = 0.027$. To explore these effects, and to test specific hypotheses on differences between the three groups, separate analyses were carried out for three comparisons: monolinguals vs. early bilinguals, monolinguals vs. late bilinguals, and early vs. late bilinguals.

Monolinguals vs. early bilinguals. Compared to monolinguals, the early bilinguals showed significantly reduced conflict cost in RT (difference of 14.6 ms) $F(1, 66) = 4.74$, $p = 0.033$, but not in the accuracy of conflict resolution, $F(1, 66) = 1.21$, $p = 0.27$.

Monolinguals vs. late bilinguals. Late bilinguals were more efficient than monolinguals in the resolution of conflict, for both RT (23 ms), $F(1, 60) = 4.91$, $p = 0.031$, and ERR (9%), $F(1, 60) = 8.56$, $p = 0.005$.

Early vs. late bilinguals. The lower conflict cost in terms of RT was observed for late bilinguals in comparison to the early bilinguals (69.8 vs. 78.2 ms), although the effect was not significant, $F < 1$. The magnitude of conflict in the ERR measure was significantly lower for late bilinguals (10 vs. 16.3%), $F(1, 62) = 4.46$, $p = 0.039$. Because the two language groups differed in self-rated proficiency for their two languages, we compared early and late bilinguals again, adding as covariates the L1 and L2 proficiency. When the L1 and L2 proficiency was controlled for, the between-group effects proved non-significant both for RT and ERR, Fs < 1. However, because the two groups differed especially in reading and writing skills in Chinese (the early bilinguals reported very poor writing and reading skills in their L1 Chinese), for further analyses we calculated the index of L1–L2 balance in speaking and listening only. The index was a result of subtraction of the mean L2 proficiency in speaking and listening from the mean L1 proficiency. When this index was included as a covariate, the RT difference in conflict was still non-significant, $F < 1$, while there was a strong trend indicating higher ERR in early than in late bilinguals, $F(1, 61) = 3.42$, $p = 0.069$. A similar pattern of results was observed when an index of balance of daily use was used instead of the balance of proficiency. The "balance of use" was a subtraction of percentage of L2 use from percentage of L1 use. When this covariate was included, the differences for RTs remained non-significant, $F < 1$, while for ERR the difference was again marginally significant, $F(1, 61) = 3.52$, $p = 0.066$.

Hemispheric asymmetry in monolinguals and bilinguals

In order to investigate the functional hemispheric asymmetry in mono- and bilingual participants, we conducted ANCOVA with three within subject factors: cue condition (no cue, valid spatial, invalid spatial, center), flanker type (congruent, incongruent), and VF (left, right), and the group of participants as between subject factor. There were no significant interactions between VF and group, all Fs < 1.5, suggesting no between-group differences in the hemispheric asymmetry. However, based on previous research, we expected the reduced hemispheric asymmetry to be particularly apparent in the group of early bilinguals when compared to monolinguals. Therefore, to further explore the issue of hemispheric asymmetry, we carried out separate tests for the three following comparisons: monolinguals vs. early bilinguals, monolinguals vs. late bilinguals, and early vs. late bilinguals. While monolinguals responded 13 ms faster to the LVF than to the RVF targets, the early bilinguals did not exhibit such asymmetry (LVF – RVF = 1 ms); the pattern of lateralization in the three participant groups is presented in **Figure 3**. However, the interaction was only marginally significant, $F(1, 66) = 3.80$, $p = 0.055$. The other comparisons between groups and visual fields were not significant.

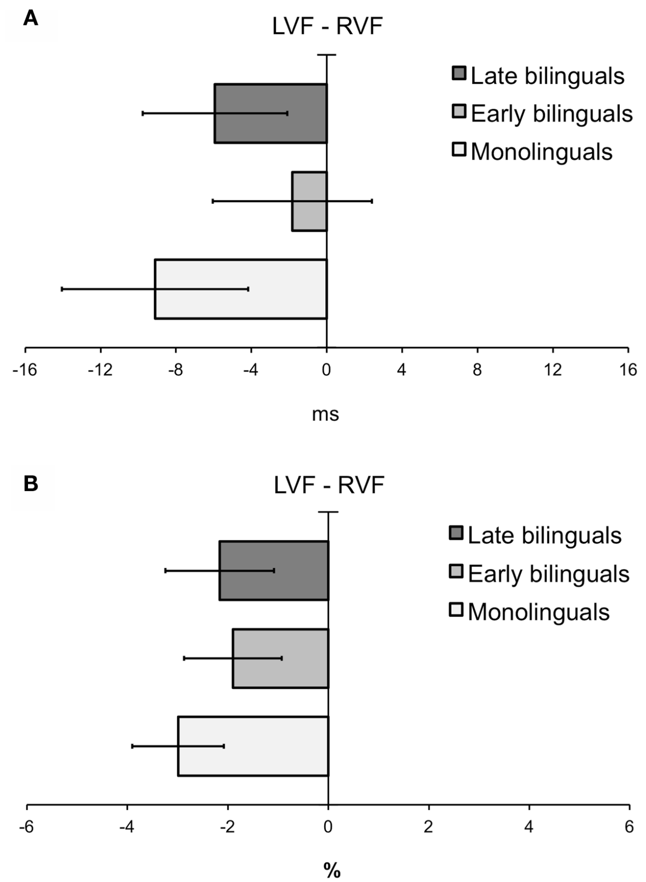


FIGURE 3 | Left visual field advantage (LVF *minus* RVF) in three groups, calculated from RT (A) and ERR (B) data.

DISCUSSION

The goal of the present study was to investigate the effects of managing two structurally distinct languages on the efficiency of attentional networks in early and late Chinese–English bilinguals. Additionally, we aimed to investigate the influence of early and late bilingualism on hemispheric asymmetries of attentional networks.

The overall pattern of results for the attentional networks across the three groups replicated findings from previous studies (Fan et al., 2002; Greene et al., 2008; see MacLeod et al., 2010, for a meta-analysis). Participants responded more quickly and made fewer errors when (a) there was a warning cue presented before the target stimuli (alerting network), (b) the spatial cue correctly indicated the location of the target (orienting network), and (c) flanking arrows pointed in the same direction as the target arrow (conflict/executive network). All these effects were robust and highly significant. The current variant of LANT was slightly more demanding than the previous ANT (Fan et al., 2002) and LANT (Greene et al., 2008), as revealed by the slower overall RT and higher ERR. It was presumed that greater demands on attention would circumvent the usually observed ceiling effect in accuracy, as well as improve the reliability of the measured effects (cf. Evert et al., 2003; Verleger et al., 2009; Asanowicz et al., submitted for publication). To this end, the eccentricity of the targets of the original LANT (Greene et al., 2008) was doubled. Presenting stimuli more peripherally was expected to decrease visual acuity and contrast sensitivity; hence, more attention would be needed for proper target discrimination (cf. Carrasco, 2011).

Consistent with our predictions, the English monolinguals were less efficient in resolution of conflict than each of the two Chinese–English bilingual groups. Importantly, the effects were not attributable to differences in socioeconomic background or non-verbal intelligence, as these factors were statistically controlled for. The results are in line with previous studies that used the ANT to compare young adult bilinguals and monolinguals (Costa et al., 2008, 2009; Marzecová et al., submitted for publication) and extend their scope by including bilinguals who speak two languages that are distinct from each other (i.e., Chinese and English).

THE DIFFERENTIAL IMPACT OF EARLY AND LATE BILINGUALISM ON EXECUTIVE CONTROL

In the present study, both bilingual groups outperformed monolinguals. However, the difference between early bilinguals and monolinguals seemed to be qualitatively different from the difference between late bilinguals and monolinguals. On the one hand, the results for early bilinguals showed a reduced conflict cost in RT (but not in ERR) as well as an advantage in overall RT. On the other hand, there was an advantage for late bilinguals in conflict resolution both in terms of RT and ERR, without significant differences in overall RT or ERR.

In the vast majority of studies reporting a bilingual advantage in conflict resolution, the benefit has been present not just selectively for trials that require resolution of conflict, but also in the overall RT measure (see **Table 1** for an overview). Such results have led researchers to propose that bilingualism may not only influence the efficiency of conflict resolution, but also another aspect of cognitive control, referred to as the "monitoring system," which evaluates the need to engage the conflict resolution mechanism (Bialystok et al., 2009; Costa et al., 2009). According to Costa et al. (2009), if the task at hand engages the monitoring system to a large extent, the advantage for bilinguals on overall RTs emerges. Costa et al. (2009) proposed two alternative ways in which the interplay between conflict resolution and monitoring processes might be explained. According to the first hypothesis, bilingualism may independently influence both monitoring and conflict resolution processes. According to the second hypothesis, the monitoring system may account for the observed bilingual advantage on both overall RT and conflict cost. The fact that the bilingual benefit in conflict resolution in most of the previous studies co-occurs with the overall RT benefit seems to support the latter claim (cf. Costa et al., 2009). However, recent findings by Marzecová et al. (submitted for publication) do not bear out this alternative and instead support the first hypothesis, according to which the two types of benefits might be dissociable. These authors reported the advantage for bilinguals over monolinguals in conflict resolution with no group differences on overall RT; moreover, their results were obtained in a condition in which high monitoring should have been involved (i.e., with a 50/50 proportion of congruent and incongruent trials; cf. Costa et al., 2009). In the present study, only early bilinguals outperformed monolinguals in overall RTs; the late bilinguals showed an advantage over monolinguals only in the conflict resolution *per se*. These results seem to indicate that specific bilingual experience may differentially influence the conflict resolution and/or monitoring systems. Let us consider some aspects of bilingual experience that may lead to enhancement of these particular cognitive processes.

The bilinguals tested in the two previous studies by Costa et al. (2008, 2009) were highly proficient and balanced early bilinguals. Since the bilingual advantage in overall RTs observed in those studies was not always accompanied by the reduced cost of conflict resolution, authors concluded that bilingualism primarily influences the monitoring system rather than the conflict resolution processes. In the present study, the bilingual advantage on overall RTs was observed only in combination with a reduced conflict cost for early bilinguals, whose L1 proficiency was rather limited. Although bilinguals from the studies by Costa et al. (2008, 2009) and the early bilinguals from the current study differed from each other in many aspects of language experience, they shared one common characteristic – the early age of acquisition. It seems plausible then to speculate that early, simultaneous consolidation of two language systems may bring about enhanced monitoring processes. However, it seems that not all early bilinguals show an advantage in monitoring (see Marzecová et al., submitted for publication), and that even bilinguals who acquired their two languages later in childhood (around age six) may exhibit such advantages (Emmorey et al., 2008). On the other hand, the late bilinguals in the present study, who were at the same time more balanced in their proficiency and use of their two languages (see **Table 2**), displayed the reduced conflict cost without any effects on overall performance. The results observed in the group of late bilinguals were similar to those reported in the study by Marzecová et al. (submitted for publication) on a group of early but moderately unbalanced bilinguals. Although bilinguals from the experiment conducted by Marzecová et al. (submitted for publication) and the late bilinguals from the present study differed in their age of L2 acquisition, they were similar with regard to balance and proficiency of the two languages. It seems that the common factor in their language experience – the moderate balance – might be responsible for the dissociation in the pattern of results: The lack of evidence for specific enhancement of monitoring processes (i.e., lack of advantage in overall RT) along with the clear advantage of a reduced conflict cost. Hence, although at this point it seems rather difficult to disentangle the factors that may lead to specific enhancement of cognitive processes in bilinguals, the present study shows the necessity of such an endeavor.

It seems important to note that alternative interpretations of overall RT advantage other than the monitoring account put forward by Costa et al. (2009) are plausible. The overall RT advantage may be equally interpreted as a measure of tonic alertness or vigilance (Roca et al., 2011). Therefore, the advantage of early bilinguals on overall performance may result from their greater vigilance. By this account, early bilinguals would be more focused on the task at hand and therefore more efficient in executing correct responses (cf. Marzecová et al., submitted for publication).

In additional analyses of the executive network efficiency, we compared the two bilingual groups. The late bilinguals showed reduced conflict cost in ERR when compared with early bilinguals. This result is consistent with our initial hypothesis that late bilinguals would show a greater advantage in conflict resolution than early bilinguals, since they may utilize the executive network to a greater extent in order to control the interference from L1 and to support processing of their less automatized L2. The bulk of evidence for a bilingual advantage in executive functions was based on research with bilinguals who used both languages regularly since early childhood, and were relatively proficient in both; that is, they were early balanced bilinguals (e.g., Carlson and Meltzoff, 2008; Costa et al., 2008, 2009; Hernandez et al., 2010; see also Bialystok, 2009, for a review). Because in most of these studies, the experiential factors were correlated with each other, it was impossible to disentangle the relative importance of each of them. The present study indicates clearly, that early L2 acquisition is not essential for the enhancement of conflict resolution processes, although it may play a part in the emergence of efficient monitoring processes.

Although in the present study, neither the early nor the late bilingual group could be regarded as perfectly balanced (considering the self-ratings of proficiency and percentage of use in L1 and L2 reported in **Table 2**), the early bilinguals were significantly less balanced than the late group. When the balance of proficiency (in terms of speaking and listening comprehension) or balance of use was controlled for, the differences between early and late bilinguals became markedly reduced. Thus, the present finding showing greater enhancement in conflict resolution for the late bilingual group seems to be in line with previous studies showing greater efficiency of executive control in balanced bilinguals (Carlson and Meltzoff, 2008; Luk and Bialystok, 2008). Furthermore, the advantage over monolinguals observed for the early bilingual group adds to the existing literature in providing evidence that enhancement of executive control is plausible for bilinguals who are far from being balanced. Taken together, the results from the present study suggest that the degree of balance between the bilinguals' two languages may have a greater impact on conflict resolution than the age of onset of bilingualism, but that the age of L2 acquisition may play an important role in mediating the monitoring advantage.

ALERTING NETWORK

Regarding the efficiency of the alerting, there was a trend for late bilinguals to exhibit greater benefit from the alerting cue than did the other two groups. Such a trend accords with previous studies in which a larger alerting effect was found for bilinguals compared to monolinguals (Costa et al., 2008; Marzecová et al., submitted for publication). There is no apparent explanation for the lack of significant group differences in alerting, apart from the concern of a methodological nature. It has been reported that the reliability of the alerting index as measured by the ANT is considerably lower than indexes of orienting and executive networks (MacLeod et al., 2010). This especially holds true for the LANT designed by Greene et al. (2008), in which the reliability of the alerting index is even lower than in the ANT. Hence, with regard to the alerting network, the experimental design might not be have been sensitive enough to capture the potentially small between-group differences, especially using a participant sample that is smaller relative to the Costa et al. (2008) study ($n = 200$).

At present not enough experimental evidence is available to provide an account of the mechanisms underlying the effects of bilingualism on the alerting network (cf. Costa et al., 2008). It is speculated that, unlike Chinese–English bilinguals in the present study, bilinguals with two structurally similar language sets (e.g., Catalan–Spanish) may need to achieve and maintain a higher state of alertness in monitoring and switching between their languages, thus gaining significant enhancement in the alerting network. Furthermore, since the central cue was always predictive of the time of target presentation, the alerting index may be seen as a combination of two processes: alertness and response preparation based on temporal expectancy. Temporal preparation has been shown to enhance not only perceptual processing, but also motor processing, thus leading to faster RTs as well as higher accuracy (Correa et al., 2005). Moreover, such an anticipatory process has been shown to enhance controlled stimulus–response selection (Correa et al., 2009). The trend for late bilinguals to show higher alerting may therefore suggest an enhanced efficiency of response anticipation mechanisms (Marzecová et al., submitted for publication), which are known to be supported by the executive control network (cf. Fan et al., 2007; Correa et al., 2009).

ORIENTING NETWORK

In relation to the orienting network, the absence of between-group differences in orienting benefit is consistent with previous studies (Costa et al., 2008, 2009). However, in the current study, the late bilinguals showed significantly greater orienting cost compared to monolinguals, i.e., they were slower to reorient attention to a target occurring in an invalidly cued location. It has been shown that in tasks with highly predictive spatial cues (as was the case in the current study), the orienting cost is associated with deactivation of the temporo-parietal junction (TPJ; Doricchi et al., 2010) – the structure that regulates reorienting of attention to uncued locations (Corbetta et al., 2008). The inhibition of TPJ seems to lead to greater filtering of stimuli occurring in the uncued location (Doricchi et al., 2010; Lasaponara et al., 2011). Therefore, the observed effect may indicate that late bilinguals have a greater capacity to inhibit stimuli that occur in an invalid location, which helps them use the predictive cue more efficiently by filtering out the uncued stimuli in the anticipatory period.

HEMISPHERIC ASYMMETRY

The LVF advantage was observed in both overall RT and ERR measures across all three groups, generally suggesting right hemisphere superiority in attentional processing (cf. Heilman, 1995; Mesulam, 1999), as assessed by behavioral measures of attentional networks. Additionally, the RT in the invalid spatial cue condition indicated right hemisphere specialization in reorienting of attention to targets occurring outside the current focus of attention. This is consistent with the neuroanatomical model of orienting networks proposed by Corbetta and Shulman (2002). The accuracy measure in the incongruent flanker condition seems to point to dominance of the right executive network in conflict resolution, which accords with several behavioral and imaging studies (Hazeltine et al., 2003; Aron et al., 2004; Asanowicz et al., submitted for publication; but see Fan et al., 2003). For the alerting effect, as in earlier LANT studies (Greene et al., 2008; Poynter et al., 2010), we did not observe any VF affects.

In line with our hypothesis, the comparisons of VF effects for overall RT in monolingual and early bilingual groups revealed a strong trend toward a reduced LVF advantage in early bilinguals. Additionally, the comparison of VF asymmetry for overall RT between monolinguals and late bilinguals showed no significant difference. These results are consistent with our predictions and with previous findings suggesting that bilinguals, particularly those who have acquired L2 at an early age, display reduced right hemisphere dominance for non-linguistic cognitive processing. It has been put forward that bilinguals who learned their second language before age six show bilateral, rather than left hemisphere dominant, cerebral organization of language as a result of early use of multiple languages, as the brain is undergoing extensive neuron wiring and synaptic changes from age three to six (Peng and Wang, 2011). It is plausible, then, that such cortical changes at an early age may have similar effects in the cerebral organization in non-linguistic domains of cognitive processing. In the study by Marzecová et al. (submitted for publication), in which attentional functioning was examined using the LANT task, bilinguals displayed reduced hemispheric asymmetry in the executive network as compared to monolinguals. In addition, Hausmann et al. (2004) observed that bilinguals displayed reduced right hemisphere involvement relative to monolinguals in face discrimination, a process that is typically more dominant in the right hemisphere. Furthermore, these results are in accordance with studies on language processing which show reduced hemispheric asymmetry in bilinguals (see Hull and Vaid, 2007, for a meta-analysis). However, no other differences in lateralization between monolinguals and bilinguals were observed in the current study. In particular, we did not observe the reduced asymmetry of executive network in early bilinguals, which was reported by Marzecová et al. (submitted for publication). There were also no group differences in lateralization for the ERR measure; all three groups of participants had a similar, small but reliable, LVF advantage in performance accuracy. It is important to note that these results should be interpreted with caution, since there have been arguments made that behavioral laterality measures do not provide a reliable measurement of hemispheric asymmetry (cf. Paradis, 2009).

In addition, several methodological factors might have led to the pattern of results that are much less straightforward than those obtained by Marzecová et al. (submitted for publication). The LANT task used in the current study was based to a large extent on the procedure proposed by Greene et al. (2008), in which generally no asymmetries were observed. Thus, the fact that some VF effects were obtained, and were even quite consistent between groups, is noteworthy. Considering that (1) effects of attentional asymmetries are generally small and may be affected by many factors (cf. Jewell and McCourt, 2000); (2) between-group differences in attentional asymmetries must therefore be even smaller and, thus, we need even more statistical power; (3) behavioral measures of hemispheric asymmetries are indirect and inherently noisy (Zaidel, 1995), we can conclude that using an almost four times larger sample than Greene et al. (2008) increased statistical power, which in turn allowed us to observe the asymmetries. However, the power might still be insufficient.

To summarize, the results seem to be consistent with the hypothesis according to which early bilingualism reduces hemispheric asymmetry of attentional networks. However, the results are far from conclusive and more research is needed to explore this issue in greater depth. Of particular note, the question of the possible modulating effects of age of L2 acquisition on the interhemispheric organization of cognitive functions remains open.

CONCLUSION

The present study demonstrates that continual practice in monitoring and switching between two language systems can lead to the enhanced executive control due to involvement of inhibitory control processes that are required to select and produce the intended language (Green, 1998). This seems to hold true regardless of the age at which bilinguals have acquired their second language, and regardless of the similarity between the two languages. Furthermore, the benefit from the continual practice in keeping two languages apart appears to be present even for bilinguals who are strongly dominant in one of their languages, although late and more balanced bilinguals appear to show a greater enhancement in conflict resolution. The results also suggest that the age of L2 acquisition may mediate the impact of bilingualism on monitoring processes; in the current study the bilingual advantage in overall RTs was only observed in the group of early bilinguals. Such a result seems to indicate that early (and continuous) contact with two languages may be critical for the monitoring advantage to emerge. Therefore, the results clearly suggest a pattern of dissociation in the influences of bilingual experience on conflict monitoring and conflict resolution processes. Further research should aim at a scrupulous disentanglement of specific factors related to language experience, which might differentially influence cognitive control processes in bilinguals.

ACKNOWLEDGMENTS

This research was supported by a subsidy from Foundation for Polish Science and by a Group of Eight European Fellowship to Zofia Wodniecka. It was also supported by a grant from the Australian Research Council awarded to Marcus Taft. The authors wish to thank Lidija Krebs-Lazendic for technical assistance with the study and the Reviewers for their invaluable comments and suggestions.

REFERENCES

Abutalebi, J. (2008). Neural processing of second language representation and control. *Acta Psychol. (Amst)* 128, 466–478.

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

Andreou, G., and Karapetsas, A. (2001). Hemispheric asymmetries of visual ERPs in left-handed bilinguals. *Cogn. Brain Res.* 12, 333–335.

Aron, A. R., Robbins, T. W., and Poldrack, R. A. (2004). Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci. (Regul. Ed.)* 8, 170–177.

Australia Bureau of Statistics/Statistics New Zealand. (2009). *Australian and New Zealand Standard Classification of Occupations*, 1st Edn. Belconnen, ACT: Australia Bureau of Statistics/Statistics New Zealand.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79.

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Biling. (Camb. Engl.)* 12, 3–11.

Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: beyond inhibition. *Dev. Psychol.* 46, 93–105.

Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873.

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

Bialystok, E., Martin, M. M., and Viswanathan, M. (2005b). Bilingualism across the lifespan: the rise and fall of inhibitory control. *Int. J. Bilingualism* 9, 103–119.

Bialystok, E., Craik, F. I. M., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Public Interest* 10, 89–129.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1352.

Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339.

Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500.

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298.

Carrasco, M. (2011). Visual attention. *Vision Res.* doi: 10.1016/j.visres.2011.04.012

Chee, M. W. L., Tan, E. W. L., and Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *J. Neurosci.* 19, 3050–3056.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

Corbetta, M., Patel, G., and Shulman, G. L. (2008). The reorienting system of the human brain: from environment to theory of mind. *Neuron* 58, 306–324.

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.

Correa, A., Lupiáñez, J., and Tudela, P. (2005). Attentional preparation based on temporal expectancy modulates processing at the perceptual level. *Psychon. Bull. Rev.* 12, 328–334

Correa, A., Rao, A., and Nobre, A. C. (2009). Anticipating conflict facilitates controlled stimulus-response selection. *J. Cogn. Neurosci.* 21, 1461–1472.

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

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

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., and Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8, 3809–3815.

Doricchi, F., Macci, E., Silvetti, M., and Macaluso, E. (2010). Neural correlates of the spatial and expectancy components of endogenous and stimulus-driven orienting of attention in the posner task. *Cereb. Cortex* 20, 1574–1585.

Dunn, A. L., and Fox Tree, J. E. (2009). A quick, gradient Bilingual Dominance Scale. *Biling. (Camb. Engl.)* 12, 273–289.

Emmorey, K., Luk, G., Pyers, J. E., and Bialystock, E. (2008). The source of enhanced cognitive control in bilinguals: evidence from bimodal bilinguals. *Psychol. Sci.* 19, 1201–1206.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

Evert, D. L., McGlinchey-Berroth, R., Verfaellie, M., and Milberg, W. P. (2003). Hemispheric asymmetries for selective attention apparent only with increased task demands in healthy participants. *Brain Cogn.* 53, 34–41.

Fan, J., Flombaum, J., McCandliss, B., Thomas, K., and Posner, M. (2003). Cognitive and brain consequences of conflict. *Neuroimage* 18, 42–57.

Fan, J., Kolster, R., Ghajar, J., Suh, M., Knight, R. T., Sarkar, R., and McCandliss, B. D. (2007). Response anticipation and response conflict: an event-related potential and functional magnetic resonance imaging study. *J. Neurosci.* 27, 2272–2282.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347.

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

Forster, K. I., and Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* 35, 116–124.

Goetz, P. J. (2003). The effects of bilingualism on theory of mind development. *Biling. (Camb. Engl.)* 6, 1–15.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. (Camb. Engl.)* 1, 67–81.

Greene, D. J., Barnea, A., Herzberg, K., Rassis, A., Neta, M., Raz, A., and Zaidel, E. (2008). Measuring attention in the hemispheres: the lateralized attention network test (LANT). *Brain Cogn.* 66, 21–31.

Hausmann, M., Durmusoglu, G., Yazgan, Y., and Gunturkun, O. (2004). Evidence for reduced hemispheric asymmetries in non-verbal functions in bilinguals. *J. Neurolinguistics* 17, 285–299.

Hazeltine, E., Bunge, S. A., Scanlon, M. D., and Gabrieli, J. D. (2003). Material-dependent and material-independent selection processes in the frontal and parietal lobes: an event-related fMRI investigation of response competition. *Neuropsychologia* 41, 1208–1217.

Heilman, K. M. (1995). "Attentional asymmetries," in *Brain Asymmetry*, eds R. J. Davidson and K. Hugdahl (Cambridge, MA: The MIT Press), 217–234.

Hernandez, A. E., Li, P., and MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends Cogn. Sci. (Regul. Ed.)* 9, 220–225.

Hernandez, A. E., and Meschyan, G. (2006). Executive function is necessary to enhance lexical processing in a less proficient L2: evidence from fMRI during picture naming. *Biling. (Camb. Engl.)* 9, 177–188.

Hernandez, M., Costa, A., Fuentes, L. J., Vivas, A. B., and Sebastian-Galles, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Biling.(Camb. Engl.)* 13, 315–325.

Hull, R., and Vaid, J. (2006). Laterality and language experience. *Laterality* 11, 436–464.

Hull, R., and Vaid, J. (2007). Bilingual language lateralization: a meta-analytic tale of two hemispheres. *Neuropsychologia* 45, 1987–2008.

Jewell, G., and McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia* 38, 93–110.

Kovelman, I., Baker, S. A., and Petitto, L. A. (2008). Bilingual and monolingual brains compared: an fMRI investigation of syntactic processing and a possible "neural signature" of bilingualism. *J. Cogn. Neurosci.* 20, 153–169.

Kroll, J. F. (2008). Juggling two languages in one mind. *Psychol. Sci. Agenda* 22, 3–7.

Kroll, J. F. (2009). The consequences of bilingualism for the mind and the brain. *Psychol. Sci. Public Interest* 10, i–ii.

Lasaponara, S., Chica, A. B., Lecce, F., Lupiáñez, J., and Doricchi, F. (2011). ERP evidence for selective drop in attentional costs in uncertain environments: Challenging a purely premotor account of covert orienting of attention. *Neuropsychologia*. doi: 10.1016/j.neuropsychologia.2011.05.012

Li, P., Sepanski, S., and Zhao, X. (2006). Language history questionnaire: a web-based interface for bilingual research. *Behav. Res. Methods* 38, 202–210.

Luk, G., Anderson, J. A. E., Craik, F. I. M., Grady, C., and Bialystok, E. (2010). Distinct neural correlates for two types of inhibition in bilinguals: response inhibition versus interference suppression. *Brain Cogn.* 74, 347–357.

Luk, G., and Bialystok, E. (2008). Common and distinct cognitive bases for reading in English-Cantonese bilinguals. *Appl. Psycholinguist.* 29, 269–289.

MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., and Shore, D. I. (2010). Appraising the ANT: psychometric and theoretical considerations of the attention network test. *Neuropsychology* 24, 637–651.

Marks, G. N., McMillan, J., Jones, F. L., and Ainley, J. (2000). *The Measurement of Socioeconomic Status for the Reporting of Nationally Comparable Outcomes of Schooling*. Australian Council for Educational Research Draft Report. Available at: http://www.curriculum.edu.au/verve/_resources/socioeconomicstatus_file.pdf

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. (Camb. Engl.)*11, 81–93.

McMillan, J. (2010). Occupation-based conceptualisations of socioeconomic status. *Paper Presented at the Socioeconomic Status and Australian Higher Education Students Symposium*, University of Technology, Sydney.

McMillan, J., Beavis, A., and Jones, F. L. (2009). The AUSEI06: a new socioeconomic index for Australia. *J. Sociol.* 45, 123–149.

Mesulam, M. M. (1999). Spatial attention and neglect: parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 1325–1346.

Moreno, S., Bialystok, E., Wodniecka, Z., and Alain, C. (2010). Conflict resolution in sentence processing by bilinguals. *J. Neurolinguistics* 23, 564–579.

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Paradis, M. (2009). *Procedural and Declarative Determinants of Second Languages*. Amsterdam: John Benjamins.

Peng, G., and Wang, W. S. (2011). Hemisphere lateralization is influenced by bilingual status and composition of words. *Neuropsychologia* 49, 1981–1986.

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25.

Posner, M. I. (2008). Measuring alertness. *Ann. N. Y. Acad. Sci.* 1129, 193–199.

Posner, M. I., and Fan, J. (2008). "Attention as an organ system," in *Topics in Integrative Neuroscience: From Cells to Cognition*, ed. J. R. Pomerantz (New York: Cambridge University Press), 31–61.

Posner, M. I., and Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annu. Rev. Psychol.* 58, 1–23.

Poynter, W., Ingram, P., and Minor, S. (2010). Visual field asymmetries in attention vary with self-reported attention deficits. *Brain Cogn.* 72, 355–361.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. (Camb. Engl.)* 13, 253–262.

Raven, J., Raven, J. C., and Court, J. H. (1998). *Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.

Raz, A., and Buhle, J. (2006). Typologies of attentional networks. *Nat. Rev. Neurosci.* 7, 367–379.

Roca, J., Castro, C., López-Ramóna, M. F., and Lupiáñez, J. (2011). Measuring vigilance while assessing the functioning of the three attentional networks: the anti-vigilance task. *J. Neurosci. Methods* 198, 312–324.

Soveri, A., Laine, M., Hamalainen, H., and Hugdahl, K. (2010). Bilingual advantage in attentional control: evidence from the forced-attention dichotic listening paradigm. *Biling. (Camb. Engl.)* 14, 371–378.

Stevens, C., Lauinger, B., and Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: an event-related brain potential study. *Dev. Sci.* 12, 634–646.

Tan, L. H., Spinks, J. A., Feng, C.-M., Siok, W. T., Perfetti, C. A., Xiong, J., and Gao, J.-H. (2003). Neural systems of second language reading are shaped by native language. *Hum. Brain Mapp.* 18, 158–166.

van Heuven, W. J. B., Schriefers, H., Dijkstra, T., and Hagoort, P. (2008). Language conflict in the bilingual brain. *Cereb. Cortex* 18, 2706–2716.

Verleger, R., Sprenger, A., Gebauer, S., Fritzmannova, M., Friedrich, M., Kraft, S., and Jaśkowski, P. (2009). On why left events are the right ones: neural mechanisms underlying the left-hemifield advantage in rapid serial visual presentation. *J. Cogn. Neurosci.* 21, 474–488.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., and Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–170.

Wodniecka, Z., Craik, F. I. M., Luo, L., and Bialystok, E. (2010). Does bilingualism help memory? Competing effects of verbal ability and executive control. *Int. J. Biling. Educ. Biling.* 13, 575–595.

Ye, Z., and Zhou, X. (2009). Executive control in language processing. *Neurosci. Biobehav. Rev.* 33, 1168–1177.

Zaidel, E. (1995). "Interhemispheric transfer in the split brain: long-term status following complete cerebral commissurotomy," in *Brain Asymmetry*, eds R. J. Davidson and K. Hugdahl (Cambridge, MA: The MIT Press), 491–532.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 March 2011; accepted: 28 May 2011; published online: 10 June 2011.

*Citation: Tao L, Marzecová A, Taft M, Asanowicz D and Wodniecka Z (2011) The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. Front. Psychology **2**:123. doi: 10.3389/fpsyg.2011.00123*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in PSYCHOLOGY
ORIGINAL RESEARCH ARTICLE
published: 28 October 2011
doi: 10.3389/fpsyg.2011.00309

The influence of expertise in simultaneous interpreting on non-verbal executive processes

Carolina Yudes*, Pedro Macizo and Teresa Bajo

Department of Experimental Psychology, University of Granada, Granada, Spain

Edited by:
Judith F. Kroll, Penn State University, USA

Reviewed by:
Gigi Luk, Rotman Research Institute at Baycrest, Canada
Annette de Groot, University of Amsterdam, Netherlands

****Correspondence:***
Carolina Yudes, Departamento de Psicología Experimental, Campus de Cartuja s/n, Universidad de Granada, 18071, Granada, Spain.
e-mail: cyudes@ugr.es

This study aimed to explore non-verbal executive processes in simultaneous interpreters. Simultaneous interpreters, bilinguals without any training in simultaneous interpreting, and control monolinguals performed the Wisconsin card sorting task (WCST; Experiment 1) and the Simon task (Experiment 2). Performance on WCST was thought to index cognitive flexibility while Simon task performance was considered an index of inhibitory processes. Simultaneous interpreters outperformed bilinguals and monolinguals on the WCST by showing reduced number of attempts to infer the rule, few errors, and few previous-category perseverations. However, simultaneous interpreters presented Simon effects similar to those found in bilinguals and monolinguals. Together, these results suggest that experience in interpreting is associated with changes in control processes required to perform interpreting tasks.

Keywords: simultaneous interpreting, bilingualism, executive processes, cognitive flexibility, inhibitory processes

INTRODUCTION

In the last two decades, an important issue within bilingual studies has been whether proficiency in two or more languages results in cognitive advantages. Previous studies have shown that becoming expert in a motor or cognitive domain sometimes leads to generalization of the acquired advantage to other domains (Die et al., 2009; Gruber et al., 2010). We understand expertise as the set of special skills and knowledge derived from extensive experience within a knowledge domain (Hoffman, 1998). Expertise may lead to a reorganization of the cortical functions as the result of this extensive experience (Maguirre et al., 2000; Mechelli et al., 2004; Gruber et al., 2010). For example, skilled video-game players have been found to develop better attentional processing (Green and Bavelier, 2003) and better skills to perform mental rotations and to work with iconic representations than non-players in playing the game Tetris (Sims and Mayer, 2002). Similarly, frequent internet communicators have been found to be more skilled at attending visual stimuli, and at planning and processing simultaneous information than infrequent internet communicators (Johnson, 2008). In addition, extensive training on dividing attention improves performance on complex concurrent tasks (Spelke et al., 1976).

According to this view about expertise, bilingual speakers who have to negotiate the use of their two languages in their daily lives can be considered experts at managing competition and resolving conflicts (Bialystok et al., 2005; Kroll and Link, 2007; Bialystok, 2008). The constant use of language selection processes to maintain activation of one language and avoiding competition from the other language may increase the ability to ignore irrelevant information and develop efficient attentional control across all domains of perceptual and cognitive processing. The idea is that the executive control mechanism in charge of resolving competition in language related tasks is similar to the control mechanism acting in the domains of perception, attention, or action (Bialystok, 2001; Kan and Thompson-Schill, 2004; Abutalebi and Green, 2007). Thus, numerous studies have examined whether control processes in language selection generalize to non-linguistic tasks involving conflict resolution (Costa et al., 2006, 2009; Bialystok, 2007; Bialystok et al., 2008). Results of these studies have provided evidence for this superior executive functioning in bilinguals when they perform tasks such as the Simon task (Bialystok, 2006; Bialystok et al., 2008), the flanker task (Costa et al., 2008), the task-switching paradigm (Prior and MacWhinney, 2010), or the anti-saccade task (Bialystok et al., 2006a).

An extreme situation for between-language control is simultaneous interpreting (SI). In this task, a spoken message in a source language (SL) must be reformulated and then produced into the target language (TL). The challenge for control comes from the fact that these processes occur in simultaneity. The interpreter receives part of a message in the SL, while she/he is mentally translating and verbally producing previous parts of the message in the TL (Gerver, 1971). Thus, the two language systems have to be simultaneously active for comprehension and production (de Groot and Christoffels, 2006). Executive control is considered essential for this task since the SL has to be selected for comprehension, while the TL has to be selected for production; therefore, strong coordination between the languages is needed to move from one language to the other (Gile, 1991, 1997; Lambert et al., 1995; Danks et al., 1997; Christoffels and de Groot, 2004). In fact, learning to interpret involves attentional training to achieve mental flexibility and coordination so that no information loss and interference between languages occurs. As experience increases, resource allocation is carried out more automatically and efficiently (Gile, 1995, 2009; Liu, 2008). In this regard, simultaneous interpreters acquire important skills for controlling their attentional resources, so they can be considered as "experts in executive control." Hence, our argument here is that SI is an extreme situation for language control and, as a consequence, extensive experience in interpreting may result in superior executive functioning.

Interestingly, language control in interpreting may differ from language control in other bilingual contexts. It has been observed that both interpreters and ordinary bilinguals experience interference from the language that is not in use (e.g., Rodriguez-Fornells et al., 2005; Kaushanskaya and Marian, 2007). However, the challenge for the bilingual is to select the appropriate language and to avoid this interference from the non-TL (Grosjean, 2001), whereas the challenge for the interpreter is to keep the two languages active and continuously switch from one language to another. Bilingual models of language control propose that language selection is regulated by inhibitory processes in both language comprehension (BIA model, Dijkstra and van Heuven, 1998) and language production (Green, 1998). These models propose that selection of the appropriate language is achieved by inhibition of the competing non-appropriate language. However, there is some evidence that inhibition may not be the mechanism by which the interpreters achieve language control. In a recent study, Ibáñez et al. (2010) asked bilinguals without any training or experience in interpreting, not in any other form of formal translation (hence forth these ordinary bilinguals will be referred simply as "bilinguals") and professional translators matched in language proficiency to read sentences word-by-word at their own pace. In all the trials participants were asked to read and understand the sentences, and to repeat them in the language of presentation. The input language (Spanish: L1 and English: L2) varied from trial to trial in an unpredictable manner. In addition, cognate words (words that share similar form and meaning in two languages) were included in some of the sentences. These two manipulations were critical: the cognate effect (that is, the difference in processing time and/or errors between cognates and non-cognates) is often thought to indicate that bilingual's two languages are simultaneously activated (Kroll and Stewart, 1994; Dijkstra et al., 1999; Macizo and Bajo, 2006); whereas switching between the languages of input provides a way to examine whether the non-appropriate language is indeed inhibited. Specifically, if the latter holds, an asymmetrical switching cost may be expected to occur, the switching cost being larger when the input language changes from the less dominant L2 to the more dominant L1 than when it changes from L1 to L2 (Meuter and Allport, 1999). The results of this experiment indicated that lexical processing depended on the participants' experience in professional translation. Experienced translators were faster at processing cognate words relative to control words, indicating that the two languages were active during the course of reading. In addition, the translators did not seem to inhibit the irrelevant language because there was no asymmetrical switching cost. In contrast, the bilinguals presented larger switching cost when switching to L1 than when switching to L2 (asymmetrical pattern of switching cost) indicating that they inhibited the non-TL when they understood sentences in their alternative language. Moreover, the bilinguals processed cognate and control words equally rapid indicating that only the language in which the sentences were presented was active in each trial. This would suggest that bilinguals and translators negotiate their two languages in different ways and it is possible that these differences extend to differences between bilinguals and interpreters (translators with professional experience in interpreting tasks) in executive functions. More specifically, it might be possible that interpreters had formed strong connections between lexical equivalents as a result of their practice in interpreting which would favor the automatic activation of their two languages. In addition, it is also possible that difference between ordinary bilinguals and interpreters extends to the enhancement of different executive functions.

Hence, we aimed to explore this last hypothesis by comparing professional interpreters with bilinguals and monolingual participants in two tasks tapping different aspects of executive control: the Wisconsin card sorting test (WCST) and the Simon task. These tasks were selected following the theoretical framework provided by Miyake and colleagues (Miyake et al., 2000; Friedman and Miyake, 2004; Friedman et al., 2006). In different studies they have investigated the psychometric relationships between the tasks that are commonly used to assess executive control and they have identified three separable control functions: "shifting" between tasks and mental sets (also called "flexibility"), "inhibition" of unwanted responses, and "updating" and monitoring of working memory (WM) representations. In the current study, we focused on two of these control functions, shifting and inhibition. From our previous analyses, we hypothesized that the interpreters should show superior performance in tasks requiring "shifting" (e.g., WCST), whereas bilinguals may be superior in tasks requiring "inhibiting" unwanted responses (e.g., Simon task). The WCST is a stimulus categorization task in which the participants have to infer a sorting rule that allows them to arrange a set of cards. This rule is modified during the task and the participants have to infer new rules continuously. The participants receive information on whether their responses are correct or not, but they are not informed about the underlying rule. Thus, this test reflects the participant's ability to switch their mental set and, therefore, her/his mental flexibility to infer the rule. On the other hand, the Simon task is used to capture inhibitory control of prepotent responses in the presence of conflicting information. In the task, the participants have to pay attention to one stimulus dimension (i.e.,color) while ignoring another irrelevant dimension (i.e., spatial position). However, the typical result is that participants cannot ignore the information about the stimulus location and they show longer reaction times when there is conflict between the spatial information provided on the screen (left or right) and the response key (left or right). This result is known as Simon effect (Lu and Proctor, 1995, for a review, Simon, 1990) which indicates that participants are not able to resist the misleading information and that they have difficulties inhibiting the response. A more detailed description of WCST and Simon tasks used in this study will be provided in the next sections.

In the present study professional interpreters were compared with bilinguals and monolinguals in two cognitive functions. In Experiment 1, we used the WCST to evaluate the cognitive function of shifting, while in Experiment 2, we used the Simon task to evaluate the ability of inhibiting irrelevant information.

EXPERIMENT 1 WISCONSIN CARD SORTING TEST

MATERIALS AND METHODS

Participants

Forty-eight participants served as volunteers in this study. Participants were paid for their participation. The first group was composed of 16 Spanish monolingual speakers from the University of Granada (11 female). The second group was composed of 16 fluent bilingual speakers (10 female) with Spanish as their native language and English as their L2. Finally, the third group was composed of 16 professional interpreters (8 female) with a mean of 10.83 years of experience in interpreting (participants' characteristics can be seen in **Table 1**).

Participants completed the Raven progressive matrices intelligence test to control for general intelligence. An ANOVA conducted on the total scores indicated that the groups did not differ, $F < 1$. Hence, possible between-group differences cannot be due to unspecific global skills (see **Table 1**).

In addition, since studies on executive functioning have observed differences associated to WM span (Padilla et al., 2005), the participants were assessed in their WM amplitude. Thus, participants performed a Spanish version of the Reading Span Test (Daneman and Carpenter, 1980) to assess their WM capacity. In this test sets of sentences are shown and participants are instructed to read each sentence aloud and to recall the last word of each sentence at the end of the set. The number of sentences in the set increases gradually from two to six. The size of the largest set of sentences in which all last words are recalled correctly represents the participant's memory span. Subjects with 3.5 or higher scores are usually considered to have a high memory span (Miyake et al., 1994). An ANOVA conducted on the mean WM span showed significant differences among the groups, $F(2, 45) = 14.19$, MSE = 0.39, $p < 0.05$. These differences were due to the higher memory span for the interpreters relative to the monolinguals, $F(1, 45) = 24.52$, $p < 0.05$, and bilingual speakers, $F(1, 45) = 17.94$, $p < 0.05$ (means can be seen in **Table 1**). There were no differences in WM span between monolinguals and bilinguals ($p > 0.05$). The larger memory span of the interpreters replicates previous results (Bajo et al., 2000; Padilla et al., 2005).

The interpreters were also older than the bilinguals and monolinguals, $F(2, 45) = 17.33$, MSE = 53.08, $p < 0.05$ (see **Table 1**). Therefore, in the analyses that we report below we first include the entire group of interpreters, then we performed the same analyses with a smaller interpreter group ($N = 8$) equated in age (and other demographic variables) to the bilingual and monolingual groups. Since the pattern of results was identical, we are reporting only those in which the complete group of 16 interpreters was included.

We also asked the bilinguals and interpreters to fill out a language history questionnaire (see Macizo and Bajo, 2006; Macizo et al., 2010) to assess their language proficiency and the history of their two languages. The mean scores for each group in reading, writing, speaking, and speech comprehension are reported in **Table 1**. The analyses carried out on these data revealed that there were no differences between bilinguals and interpreters in their general L2 proficiency, $F(1, 30) = 1.15$, MSE = 0.48, $p > 0.05$, or in the frequency of use of their L2, $F(1, 30) = 0.30$, MSE = 1.93, $p > 0.05$. In order to guarantee maximal comparability within our bilingual groups (interpreters and bilinguals), we selected only non-balanced bilingual speakers (the interpreters were all unbalanced-late bilinguals). Consequently, both groups had similar proficiency, history, and use of their second language, although they differed in their interpreting experience.

In addition to the participants' mean age, mean WM span, Raven scores, and proficiency measures, **Table 1** shows the time that the participants spent living in L2 speaking countries and their profession/occupation. The main difference between ordinary bilinguals and interpreters was their educational training in translation and interpreting, only the latter having been formally trained in interpreting. In addition, the interpreters had practiced professional interpreting for a long time ($M = 10.83$ years) while the bilinguals lacked such professional experience. Professional interpreting was the interpreters' main occupation, although they also had occupations similar to the main occupations of ordinary bilinguals.

Table 1 | Characteristics of participants in the study.

	Monolinguals	Bilinguals	Interpreters
Age	21.65 (2.91)	25.68 (3.17)	36.31 (11.85)
WM span	3 (0.67)	3.26 (0.65)	4.29 (0.54)
Raven	27 (30.11)	31.93 (29.98)	35.62 (24.34)
SECOND LANGUAGE (L2) PROFICIENCY QUESTIONNAIRE			
Fluency (total)		8.35 (0.47)	8.62 (0.86)
Reading		8.78 (0.75)	8.87 (0.95)
Writing		7.93 (1.06)	8.49 (0.96)
Speaking		8.15 (0.88)	8.43 (1.03)
Speech comprehension		8.56 (0.81)	8.68 (1.01)
Frequency of use (days per week)		4.64 (1.29)	4.91 (1.48)
Write		4.75 (1.84)	4.68 (1.66)
Read		5.49 (1.41)	5.37 (1.54)
Speak		3.68 (2.35)	4.68 (1.88)
Time living in L2 speaking countries (months)		14.48 (6.49)	12.11 (4.07)
PROFESSION			
Main	Under graduated students	Touristic guide/ English teacher/ Ph. D. students on English philology	Interpreting
Secondary			Touristic guide/ English teacher/ international business

The self-report in the language history questionnaire ranged from 1 to 10 where 1 was not fluent and 10 very fluent. Means and SD (in brackets) are reported.

Materials and procedure

We used the Spanish version of the WCST (Cruz-Lopez, 2001). The test is composed of 128 response-cards and 4 stimulus-cards depicting geometric figures. The figures differ along three dimensions: shape (cross, circle, triangle, or star), color (red, blue, yellow, or green), and number of items (one, two, three, or four). These dimensions are combined to compose the response-cards which included one or more figures with the same shape and color, for example, cards with one green cross, cards with three yellow circles, cards with four blue triangles, etc. On the other hand, the stimulus-cards depicted one red triangle; two green stars; three yellow crosses; and four blue circles.

The experiment was conducted in a quiet and well lit room. The stimulus-cards were given to the participants and they were asked to sort each response-card placing it on one of the stimulus-cards according to a sorting rule. Participants were not informed of the particular sorting rule, but every time a response-card was sorted, they received positive or negative feedback depending on whether the response matched the sorting rule. This feedback was provided to allow the participant to guess the rule and to make correct responses. However, after 10 consecutive correct responses the sorting rule changed. The participants were not informed of this change of rule but they received negative feedback if they continued sorting the cards with the previous rule. Thus, through negative and positive feedback participants should again guess the correct rule. The first sorting rule was based on the color dimension, the second on the shape, and the third on the number. The rules were repeated twice before completing the test. The task finished either when the participant inferred the six rules (color, shape, number, twice each) successfully or when the participant reached the maximum of 128 trials.

RESULTS

First, we report analyses on global performance (number of completed categories, number of attempts, and number of errors). Then, we report detailed analyses on different types of error to capture differences in mental flexibility (Barceló and Knight, 2002).

Global performance

Number of completed categories. The number of categories ranged from 0 to 6 (0 meant that the participant was not able to complete 10 consecutive correct responses to any of the categories and six meant that the participant successfully achieved all the series). The results of the ANOVA on the number of completed categories indicated that there were no differences among the groups, $F(2, 45) = 2.62$, $\text{MSE} = 2.21$, $p > 0.05$, $\eta_p^2 = 0.11$ (see **Table 2**).

Number of attempts. The analysis on the number of attempts to find the correct sorting rule (max. 128) revealed a main effect of group, $F(2, 45) = 7.92$, $\text{MSE} = 426.21$, $p < 0.05$. The interpreters needed fewer attempts to guess the rule (90.68 out of the 128 possible attempts) than the rest of the groups. The differences were significant when compared to the monolinguals, $F(1, 45) = 14.01$, $p < 0.05$, and to the bilinguals, $F(1, 45) = 9.29$, $p < 0.05$, $\eta_p^2 = 0.26$, whereas there were no differences between monolinguals and bilinguals, $F < 1$ (see **Table 2**).

Table 2 | Mean number of completed categories, attempts, and errors (and SD) for each group of participants. CI: 95% confidence interval.

	Global performance		
	Completed categories	**Number of attempts**	**Number of errors**
Monolinguals	4.37 (1.66)	118 (14.85)	41.59 (16.64)
CI (95%)	3.48–5.26	110.08–125.91	32.73–50.47
Bilinguals	4.56 (1.45)	112.93 (21.05)	40.62 (21.30)
CI (95%)	3.78–5.34	101.72–124.15	29.27–51.97
Interpreters	5.50 (1.31)	90.68 (24.79)	22.37 (20.12)
CI (95%)	4.79–6.19	77.47–103.89	11.65–33.09

Maximum number of completed categories = 6; maximum number of attempts = 128.

Number of errors. The analysis on the number of errors showed a reliable main effect of group, $F(2, 45) = 4.95$, $\text{MSE} = 378.59$, $p < 0.05$, $\eta_p^2 = 0.18$, so that the interpreters had the lower percentage of errors as compared to the monolinguals, $F(1, 45) = 7.81$, $p < 0.05$, and the bilinguals, $F(1, 45) = 7.03$, $p < 0.05$. There were not significant differences between monolinguals and bilingual participants, $F < 1$ (see **Table 2**).

Types of error. The WCST manual (Heaton et al., 1993) distinguishes between perseverative and non-perseverative errors. The perseverative errors are failures to change the mental rule after receiving negative feedback so that the person continues sorting the cards according to the previous-category dimension despite feedback indicating that the response was wrong. The non-perseverative errors are the normal errors needed to learn the new rule. This type of error reflects an attitude to change the response after receiving disconfirming feedback (Barceló, 1999). An ANOVA was conducted to examine the distribution of these types of error in each group. The results of this analysis yielded a significant interaction between-group and type of error, $F(2, 45) = 11.92$, $\text{MSE} = 20.29$, $p < 0.05$, $\eta_p^2 = 0.18$. This interaction indicated that there were no group differences when analyzing non-perseverative errors, $F(2, 45) = 1.38$, $p > 0.05$. However, the effect of group was significant when analyzing the percentage of perseverative errors, $F(2, 45) = 8.39$, $p < 0.05$. In this case, interpreters showed fewer errors than the monolinguals, $F(1, 45) = 15.29$, $p < 0.05$, and the bilinguals, $F(1, 45) = 9.07$, $p < 0.05$. There were no differences between monolinguals and bilinguals, $F < 1$ (see **Table 3**).

To further understand the effect of expertise in interpreting, we performed additional analyses on the perseverative errors. Thus, we categorized these errors into perseverations to the immediately preceding category and perseverations to a different-category (Hartman et al., 2001). Previous-category perseverations reflect lack of flexibility to change the mental set to a new rule, while different-category perseverations reflect the understanding that the previous rule is no longer correct but there is an unsuccessful attempt to infer a new rule. The ANOVA performed on the number of previous-category perseverations revealed a significant effect of group, $F(2, 45) = 6.16$, $\text{MSE} = 51.56$, $p < 0.05$, $\eta_p^2 = 0.21$, with

Table 3 | Mean number (and SD) of different types of errors (perseverative and non-perseverative), and mean number (and SD) of types of perseverations (previous-category and different-category perseverations) in each group of participants.

Group	Types of error		Types of perseverative errors	
	Perseverative	Non-perseverative	Previous-category	Different-category
ERROR SCORES				
Monolinguals	19.46 (9.29)	15.42 (8.25)	12.98 (7.19)	3.03 (3.24)
Bilinguals	16.50 (10.13)	18.50 (9.24)	11.12 (8.94)	3.18 (4.95)
Interpreters	6.56 (8.49)	13.50 (8.21)	4.50 (4.77)	0.56 (1.15)

the interpreters showing fewer previous-category perseverations than the monolinguals, $F(1, 45) = 11.16$, $p < 0.05$, and the bilinguals, $F(1, 45) = 6.79$, $p < 0.05$. There were no differences between monolinguals and bilingual participants ($p > 0.05$). In contrast, the ANOVA on the number of different-category perseverations showed marginally significant differences between the groups, $F(2, 45) = 2.82$, MSE = 12.22, $p = 0.06$, $\eta_p^2 = 0.11$ (means can be seen in **Table 3**). To avoid the problem of unequal variance across cells (because of the reduced number of errors) we performed also analyses with the arcsine transformation of these values. The results of these analyses were the same as those reported here. In addition, to control for the possible non-parametric distribution of errors, non-parametric analyses were performed (Friedman ANOVAs; see Friedman, 1940). The results of these analyses were the same as those reported here.

Because the group of interpreters had larger WM capacity than the groups of monolinguals and bilinguals, we decided to explore whether the observed differences were due to differences in this capacity. Thus, we ran a new series of analyses exploring the role of WM span. Since monolingual and bilingual participants had a similar performance on the WCST and comparable WM capacity (smaller than the interpreters), we pooled them out and divided them up according to their WM. In this way, we composed a group of 16 non-interpreters (eight monolinguals and eight bilinguals) with low WM capacity (with scores below 3.5 in the Reading span test; $M = 2.63$, SD = 0.39), and a group of 16 non-interpreters (eight monolinguals and eight bilinguals) with high WM capacity (with scores greater than 3.5 in the Reading span test; $M = 3.62$, SD = 0.59), and compared them with the group of professional interpreters with high WM span ($M = 4.29$, SD = 0.54). An ANOVA revealed that WM span was similar in the high span participants and interpreters ($p > 0.05$) and both groups scored higher than the low span participants (all ps < 0.05).

In these new analyses, the ANOVA on the number of completed categories did not reveal differences among the span groups, $F(2, 45) = 2.62$, MSE = 2.21, $p > 0.05$, $\eta_p^2 = 0.11$ (low span: $M = 4.56$, SD = 1.45; high span: $M = 4.37$, SD = 1.66; interpreters: $M = 5.5$, SD = 1.31). The analysis on the number of attempts to guess the sorting rule showed a main effect of group, $F(2, 45) = 7.65$, MSE = 430.09, $p < 0.05$, $\eta_p^2 = 0.25$. The interpreters needed fewer attempts to guess the rule ($M = 90.68$, SD = 24.79) than both low span ($M = 116.45$, SD = 18.06), $F(1, 45) = 12.34$, $p < 0.05$, and high span participants ($M = 114.48$, SD = 18.68), $F(1, 45) = 10.53$, $p < 0.05$. No differences were found between high and low span participants ($p > 0.05$). In addition, the interpreters showed fewer perseverative errors ($M = 6.56$, SD = 8.49) than the low span ($M = 18.08$, SD = 9.65), $F(1, 45) = 11.98$, $p < 0.05$, and the high span participants ($M = 17.88$, SD = 10.03), $F(1, 45) = 11.56$, $p < 0.05$. No differences were found between the low and high span groups ($p > 0.05$). Similarly, the number of perseverations to the previous-category was significantly lower for the interpreters ($M = 4.5$, SD = 4.77) than for the low span ($M = 12.57$, SD = 8.26), $F(1, 45) = 10.02$, $p < 0.05$, and the high span participants ($M = 11.53$, SD = 8.05), $F(1, 45) = 7.62$, $p < 0.05$. No differences were found between low and high span groups, $p > 0.05$. Finally, the number of perseverations to a different-category was smaller for the interpreters ($M = 0.56$, SD = 1.15) than for the low span ($M = 2.31$, SD = 3.34), $F(1, 45) = 2.07$, $p < 0.05$, and the high span participants ($M = 3.89$, SD = 4.78), $F(1, 45) = 7.58$, $p < 0.05$. The low and high span groups did not differ, $p > 0.05$.

DISCUSSION

In summary, although the interpreters did not differ from monolinguals or bilinguals in the global number of completed categories, they were able to complete the task in a more efficient way. This efficiency was observed in the reduced number of attempts to infer the sequence of rules and in the smaller number of errors. Importantly, analyses on the type of errors indicated that the main differences between the interpreters and the other groups were observed in the reduced number of perseverative errors in the group of interpreters. Furthermore, when we examined the types of perseverations, we found reliable group differences in the perseveration from previous-category with the interpreters having the lowest number of this type of error. This pattern of results suggests that the interpreters were able to update the task-relevant information efficiently and rapidly change their hypothesis when needed. The interpreters looked for alternative solutions to negative feedback and they reorganized the elements of the problem faster than monolingual or bilingual speakers. Interpreters had better performance in the WCST even when they were compared with high span bilingual/monolingual participants, suggesting that their advantage on "shifting" or mental flexibility was due to their interpreting experience and not to their larger WM capacity.

EXPERIMENT 2 SIMON TASK

In Experiment 2, we examined whether the interpreters would also show better performance than the bilinguals and monolinguals in tasks requiring inhibition of conflicting responses (e.g., the Simon task). As we mentioned, whereas interpreting requires excellent switching skills and mental flexibility, there are data suggesting that interpreters might not inhibit the alternative language while interpreting (Ibáñez et al., 2010). If this was the case, very likely the interpreters would not show superior performance in the Simon task.

MATERIALS AND METHODS

Participants

The same participants that carried out Experiment 1 also participated in Experiment 2.

Materials and procedure

In the Simon task, participants had to respond to color stimuli presented on an irrelevant spatial location. In the task, each trial started with a fixation point (+) that remained on the center of the screen for 350 ms. Then, a colored square (red or blue) appeared on either the left or the right side of the fixation point and participants were instructed to press the response key corresponding to the color of the square as fast as possible. The response keys were also located left or right on the keyboard. Thus, if the colored square was red, the participant had to press the response key marked with a red sticker located on the right side of the keyboard ("intro" key); if the square was blue, the participant had to press the response key marked with a blue sticker located on the left side of the keyboard ("tab" key). If there was no response, the colored square remained on the screen for 2000 ms.

Depending on the location of the colored square on the screen, the trials could be congruent or incongruent. Thus, in congruent trials the location of the stimulus coincided with the position of the response key (e.g., red square on the right), whereas in the incongruent trials the position of the stimulus and the response key did not match, that is, the stimulus was presented on the opposite side of the correct response key (e.g., red square on the left side of the screen). In addition, there were control trials in which stimuli were centered on the screen. There were a total of 150 trials. Participants were given 24 practice trials with feedback before the experimental trials to familiarize them with the response keys. The remaining 126 trials were divided in three blocks of 42 experimental trials each. In each block there were 14 incongruent trials, 14 congruent trials, and 14 control trials, which were randomly presented. Participants received instructions to respond to the color of the squares, and they were told that the locations of the stimuli were irrelevant to perform the task.

The sequence of events and data collection was controlled by E-prime experimental software, 1.1 version (Schneider et al., 2002).

RESULTS

We performed separate ANOVAs (Group × Type of Cue) on the number of correct responses and on response times (RT).

Response times

Response times faster than 200 ms or slower than 1200 ms were excluded from the analysis (0.23%). RTs associated to incorrect responses were also filtered out (2.5%).

The analysis on RTs indicated that the effect of group was not significant, $F < 1$. A main effect of type of cue was observed, $F(2, 90) = 59.79$, MSE = 549, $p < 0.05$, $\eta_p^2 = 0.57$. This effect showed the typical Simon effect, that is, significantly slower RTs to incongruent trials ($M = 474.01$ ms, SD = 62.01) relative to the congruent trials ($M = 437.02$, SD = 66.75; see **Table 4**). The interaction between type of cue and group was not significant, $F < 1$, indicating that the Simon effect was equivalent for all the groups (Simon effect (RT): monolinguals = 34.49 ms; bilinguals = 38.74 ms; interpreters = 37.69 ms; Simon effect (correct responses): monolinguals = −5.25; bilinguals = −5.65; interpreters = −2.97)[1].

Table 4 | Mean reaction times (RT; in milliseconds) and percentage of correct responses (% CR and SD) for each group of participants for the Simon task. CI: 95% confidence interval.

	Type of trials		
	Congruent	**Incongruent**	**CI (95%) congruent/incongruent**
MONOLINGUALS			
RT	441.37 (66.58)	475.86 (53.23)	405.7–477.1/447.8–504.2
% CR	98.37 (1.82)	93.11 (7.82)	97.4–99.3/88.9–97.2
BILINGUALS			
RT	424.34 (57.26)	463.11 (60.43)	393.8–454.8/430.8–495.3
% CR	98.66 (1.21)	93.01 (7.94)	98.1–99.3/88.7–97.2
INTERPRETERS			
RT	445.35 (77.07)	483.06 (73.01)	404.2–486.4/444.1–522.1
% CR	99.25 (1.13)	96.28 (3.88)	98.6–99.8/94.2–98.3

As in the previous experiment, we explored the role of WM on performance in the Simon test. Thus, we carried out new analyses grouping monolinguals and bilinguals and then dividing them up in two groups based on their WM span (see Experiment 1). The analysis on RT when WM span was considered revealed that equivalent Simon effects were present in low span participants, $F(1, 15) = 12.37$, MSE = 785, $p < 0.05$, high span participants, $F(1, 15) = 27.71$, MSE = 426, $p < 0.05$ and interpreters, $F(1, 15) = 25.96$, MSE = 438, $p < 0.05$.

Accuracy analyses

The analysis performed on the number of correct responses revealed that there were no significant differences among the groups, $F(2, 45) = 1.45$, MSE = 28.71, $p > 0.05$, $\eta_p^2 = 0.05$. However, the effect of type of cue was significant, $F(2, 45) = 25.97$, MSE = 19.79, $p < 0.05$, $\eta_p^2 = 0.33$. Thus, all participants made fewer correct responses to incongruent trials ($M = 94.13$, SD = 6.84) relative to congruent trials ($M = 98.76$, SD = 1.44). The interaction between-group and type of trial was not significant, $F < 1$, revealing similar Simon effect for the three groups (see **Table 4**).

The analyses in which WM span was considered showed that the decrease in correct responses for incongruent trials relative to the congruent trials was significant in all the groups independently of their WM capacity (low span participants,

[1] Simon effect is calculated by subtracting the reaction times (or percentage of correct responses) to congruent trials from those to incongruent trials.

$F(1, 15) = 8.81$, MSE = 32.11, $p < 0.05$, Simon effect = −5.95; high span participants, $F(1, 15) = 9.63$, MSE = 20.39, $p < 0.05$, Simon effect = −4.96; interpreters, $F(1, 15) = 10.72$, MSE = 6.59, $p < 0.05$, Simon effect = −2.97).

DISCUSSION

In summary, the interpreters showed Simon effects that were similar in magnitude to those observed in the monolingual and bilingual groups. This finding suggests that experience in interpreting does not necessarily improve the functioning of all executive processes since the type of inhibitory control required by the Simon task seemed to be independent of the exact nature of the prior bilingual language experience. This is important because it suggests that the cognitive advantage of the interpreters is not general, but restricted to the exact cognitive operations needed to perform the interpreting task.

Surprisingly, the bilinguals did not show a reduced Simon effect relative to the monolinguals and interpreters. The lack of superiority of the bilinguals when compared to the monolingual group is inconsistent with other studies that show a bilingual advantage in tasks that involve inhibitory control (Bialystok et al., 2004, 2005, 2006a,b; Bialystok, 2006; Costa et al., 2008; Hernández et al., 2010). For example, Linck et al. (2008) compared the magnitude of the Simon effect in monolingual and bilingual participants and they found the bilinguals to have a significantly smaller Simon effect relative to monolingual participants (see also Bialystok et al., 2004; Bialystok, 2006). In the Section "General Discussion" we will go back to this finding and discuss several possible reasons for not finding a bilingual advantage in our study.

GENERAL DISCUSSION

The results of Experiment 1 showed that experience in interpreting enhances cognitive flexibility as measured by the WCST. Thus, the interpreters showed fewer attempts to infer the rule, a smaller number of errors and, crucially, fewer previous-category perseverations. This pattern of results suggests that the interpreters were able to update the task-relevant information efficiently and to change their responses accordingly.

In contrast, as shown in Experiment 2, experience in interpreting did not affect performance in the Simon task. Interpreters were not able to avoid interference from the irrelevant location dimension and they showed Simon effects similar in magnitude than those observed in the bilingual and monolingual groups, suggesting that interpreting does not enhance the ability to reduce interference from conflicting responses. Different analyses showed that our results were not due to the participants' WM capacities since interpreters had a better performance than untrained high span participants in the WSCT and both groups showed similar Simon effects.

Our results are consistent with a recent study by Köpke and Nespoulous (2006) comparing professional interpreters, interpreting students, and control subjects in the Stroop task. The results showed that the interpreters were not better than the students and controls in avoiding interference in the Stroop situation. Köpke and Nespoulous (2006) hypothesized that the normal performance of the interpreters in the Stroop task could be due to the reduced validity of this task to measure the attentional skills in SI. They suggest that because the Stroop task is visual in nature, whereas SI engages meaningful auditory material, there is no transfer from interpreting to Stroop. However, our results suggest that it is not the visual nature of the task that is causing the normal performance of the interpreters since they were superior in the WCST that also involved visual materials. In our opinion, this pattern of results would rather be caused by the processes underlying interpreting than with the modality involved in the tasks. As we mentioned, the results of Ibáñez et al. (2010) suggest that interpreting does not require language inhibition, since the two languages have to be active for comprehension and production during the task. However, mental flexibility to switch from one language to another is crucial during interpreting. Consistent with these ideas, the interpreters in our study show enhanced flexibility, but normal inhibitory control. The overall pattern of performance in the interpreters suggests that, in accordance with our expectations, the interpreters' advantage is selective and restricted to the processes directly involved in interpreting.

In contrast, the bilinguals did not behave as we expected. We predicted that experience in inhibitory control for language selection would result in superior performance in the Simon task; however this prediction was not confirmed. As mentioned, this finding is not consistent with other studies showing bilingual advantages in conflict resolution in tasks such as Simon tasks, flanker tasks (Bialystok et al., 2004; Costa et al., 2008), or numerical Stroop tasks (Hernández et al., 2010). There are several reasons for the discrepancy between our study and previous studies showing a reduced Simon effect for the bilinguals (Bialystok et al., 2004, 2005, 2008; Bialystok, 2006; Costa et al., 2008; Martin-Rhee and Bialystok, 2008).

First, the cognitive consequences of bilingualism are usually more salient for balanced than for unbalanced-late bilinguals (Bialystok, 1988; Kroll and Stewart, 1994). The bilinguals in our study were of the second type (unbalanced-late bilinguals) so that they were equated in language proficiency, history, and use to our group of interpreters. In order to isolate the effect of interpreting experience, our bilinguals and interpreters had to be as similar as possible in their L2 language history, for this reason, we selected participants that acquired their second language late, namely, in adolescence or adulthood. Furthermore, the use of the two languages was also unbalanced, so they used mostly one of their languages during their daily life. Therefore, the fact that our bilinguals did not behave differently from the monolinguals in our study suggests that the cognitive advantages related to bilingualism might only be evident in balanced bilinguals. Second, and despite our previous observations, most of the data reporting a reduced Simon effect come from children and elderly bilinguals, that is, populations with restricted executive functions, while sometimes this bilingual advantage is hard to observe in the case of young adults who are at the peak of their attentional capacities (Bialystok et al., 2005, 2008; Morton and Harper, 2007; Colzato et al., 2008; see Hilchey and Klein, 2011, for a recent review). For example, Bialystok et al. (2005) compared the performance of monolinguals and bilinguals in an age range between 30 and 80 years in the Simon task. Bilingual advantages were found only from 60 years old onward (see also Ryan et al., 2004; Bialystok et al., 2006a). Therefore, these data support the idea that the age-related rate of decline is significantly less severe for bilinguals, but also that the bilingual advantage in inhibitory control may be more evident in populations that usually show deficits in executive control. Similarly, Salvatierra and Roselli (2010) compared young and old, balanced and non-balanced bilinguals, and monolinguals in a Simon task. There were simple (two colors) and complex (four colors) Simon conditions. Results indicated that the older bilinguals had better performance than older monolinguals under a simple Simon condition. However, there was no bilingual advantage in the younger sample. Moreover, the advantage was found in bilinguals who despite having acquired their second language later in life used their two languages equally often everyday. Therefore, the authors concluded that the bilingual advantage in inhibitory control might depend on the level of linguistic activation rather than on the level of proficiency or age of acquisition.

CONCLUSION

In conclusion, our results show that experience in interpreting has positive consequences for executive processing. Interestingly, this advantage was only evident in executive functions directly involved in the interpreting tasks. Thus, interpreters showed more mental flexibility than the bilinguals and were faster in changing hypotheses online. This ability is probably associated with the interpreters' skills to alternate between languages continuously and to monitor and correct their own output while reformulating and producing speech in the TL. In contrast, and in agreement with previous data showing that inhibition may not be involved in interpreting (Ibáñez et al., 2010), interpreters were not better than monolinguals or bilinguals at ignoring conflicting information. Future research is required to determine which other aspects of executive processing are modulated by interpreting.

ACKNOWLEDGMENTS

This research was supported by the doctoral research Grant SEJ2005-00842 to Carolina Yudes, by the research program "Ramón y Cajal" to Pedro Macizo; by grants PSI2009-11094/PSI to Pedro Macizo, and EDU2008-01111 to Teresa Bajo, and CSD2008-00048 Consolider Ingenio 2010 from the Ministry of Science and Innovation of the Spanish Government; and by Proyecto de Excelencia de la JA-2007 to Teresa Bajo and Pedro Macizo and JA-2008_HUM 3600 to Teresa Bajo.

REFERENCES

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

Bajo, M. T., Padilla, F., and Padilla, P. (2000). "Comprehension processes in simultaneous interpreting," in *Translation in Context*, eds A. Chesterman, N. Gallardo-San Salvador, and Y. Gambier (Amsterdam: John Benjamins), 127–142.

Barceló, F. (1999). Electrophysiological evidence of two different types of error in the Wisconsin Card Sorting Test. *Neuroreport* 10, 1299–1303.

Barceló, F., and Knight, R. (2002). Both random and perseverative errors underlie WCST prefrontal patients. *Neuropsychologia* 40, 349–356.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Dev. Psychol.* 24, 560–567.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. New York: Cambridge University Press.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79.

Bialystok, E. (2007). Cognitive effects of bilingualism: how linguistic experience leads to cognitive change. *Int. J. Biling. Educ. Biling.* 10, 210–223.

Bialystok, E. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn.* 34, 859–873.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., Craik, F. I. M., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilingual. *J. Exp. Psychol. Learn.* 34, 859–873.

Bialystok, E., Craik, F. I. M., and Ryan, J. (2006a). Executive control in a modified anti-saccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn.* 32, 1341–1354.

Bialystok, E., Craik, F. I. M., and Ruocco, A. (2006b). Dual-modality monitoring in a classification task: the effects of bilingualism and ageing. *Q. J. Exp. Psychol.* 59, 1968–1983.

Bialystok, E., Martin, M., and Viswanathan, M. (2005). Bilingualism across the lifespan: the rise and fall of inhibitory control. *Int. J. Biling.* 9, 103–119.

Christoffels, I. K., and de Groot, A. M. B. (2004). Components of simultaneous interpreting: a comparison with shadowing and paraphrasing. *Biling. Lang. Cogn.* 7, 1–14.

Colzato, L., Bajo, M. T., van den Wildenberg, W., Paolieri, D.,Nieuwenhuis, S., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanism. *J. Exp. Psychol. Learn.* 34, 302–312.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Galles, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don´t. *Cognition* 113, 135–149.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from ANT task. *Cognition* 106, 59–86.

Costa, A., La Heij, W., and Navarrete, E. (2006). The dynamics of bilingual lexical access. *Biling. Lang. Cogn.* 9, 137–151.

Cruz-Lopez, M. V. (2001). *Test de clasificación de tarjetas de Wisconsin. Adaptación española.* Madrid: Manual Tea Ediciones S.A.

Daneman, M., and Carpenter, J. (1980). Individual differences in working memory and reading. *J. Verb. Learn. Verb. Behav.* 19, 450–466.

Danks, J. H., Shreve, G. M., Fountain, S. B., and McBeath, M. K. (1997). *Cognitive Processes in Translation and Interpreting.* Thousand Oaks: Sage.

de Groot, A. M. B., and Christoffels, I. (2006). Language control in bilinguals: monolingual tasks and simultaneous interpreting. *Biling. Lang. Cogn.* 9, 189–201.

Die, M. W. G., Green, C. S., and Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia* 47, 1780–1789.

Dijkstra, A. F. J., Grainger, J., and van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: the neglected role of phonology. *J. Mem. Lang.* 41, 496–518.

Dijkstra, A. F. J., and van Heuven, W. J. B. (1998). "The BIA-model and bilingual word recognition," in *Localist Connectionist Approaches to Human Cognition*, eds J. Grainger and A. M. B. de Groot (Mahwah, NJ: Erlbaum), 189–225.

Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *Ann. Math. Stat.* 11, 86–98.

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control processes: a latent variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., and Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychol. Sci.* 17, 172–179.

Gerver, D. (1971). *Aspects of Simultaneous Interpretation and Human Information Processing.* Unpublished Ph.D. dissertation. University of Oxford, UK.

Gile, D. (1991). Methodological aspects of interpretation (and translation) research. *Target* 3, 153–174.

Gile, D. (1995, 2009). *Basic Concepts and Models for Interpreter and Translator Training.* Amsterdam: John Benjamins.

Gile, D. (1997). "Conference interpreting as a cognitive management problem," in *Cognitive Processes in Translation and Interpretation*, eds. J. H. Danks, G. M. Shreve, S. B. Fountain, and M. K. McBeath (Thousand Oaks: Sage), 196–214.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Green, S., and Bavelier, D. (2003). Attention video game modifies visual selective attention. *Nature* 423, 534–537.

Grosjean, F. (2001). "The bilingual's language modes," in *One mind, Two Languages: Bilingual Language Processing*, ed J. Nicol (Oxford: Blackwell), 1–22.

Gruber, H., Jansen, P., Marienhagen, J., and Altenmueller, E. (2010). Adaptations during the acquisition of expertise. *Talent Dev. Excell.* 2, 3–15.

Hartman, M., Bolton, E., and Fehnel, S. (2001). Accounting for age differences on the Wisconsin card sorting test: decreased working memory, not inflexibility. *Psychol. Aging* 16, 385–399.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., and Curtis, G. (1993). *Wisconsin CARD SORTING TEST. Manual Revised and Expanded*. Odessa: Psychological Assessment Resources Inc.

Hernández, M., Costa, A., Fuentes, L., Vivas, A., and Sebastián-Galles, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Biling. Lang. Cogn.* 13, 315–325.

Hilchey, M., and Klein, R. (2011). Are there bilingual advantages on nonlinguistic intereference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658.

Hoffman, R. R. (1998). "How can expertise be defined? Implications of research from cognitive psychology," in *Exploring Expertise*, eds R. Williams, W. Faulkner, and J. Fleck (New York: Macmillan), 81–100.

Ibáñez, A., Macizo, P., and Bajo, M. T. (2010). Language access and language selection in professional translators. *Acta Psychol.* 135, 257–266.

Johnson, G. M. (2008). Cognitive processing differences between frequent and infrequent Internet users. *Comput. Hum. Behav.* 24, 2094–2106.

Kan, I., and Thompson-Schill, S. (2004). Selection from perceptual and conceptual representations. *Cogn. Affect. Behav. Neurosci.* 4, 466–482.

Kaushanskaya, M., and Marian, V. (2007). Non-target language recognition and interference in bilinguals: evidence from eye tracking and picture naming. *Lang. Learn.* 57, 119–163.

Köpke, B., and Nespoulous, J. (2006). Working memory performance in expert and novice interpreters. *Interpreting* 8, 1–23.

Kroll, J. F., and Link, J. (2007). "Representation and skill in second language learners and proficient bilinguals," in *Cognitive Aspects of Bilingualism*, eds I. Kecskes and L. Albertazzi (New York: Springer), 237–269.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174.

Lambert, S., Daró, V., and Fabbro, F. (1995). Focalized attention on input vs. output during simultaneous interpretation: possibly a waste of effort. *Meta Trans. J.* 40, 39–46.

Linck, J. A., Hoshino, N., and Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment. Lex.* 3, 349–374.

Liu, M. (2008). "How do experts interpret? Implications from research in interpreting studies and cognitive science," in *Efforts and Models in Interpreting and Translation Research. A Tribute to Daniel Gile*, eds G. Hansen, A. Chestermanand, and H. Gerzymisch-Arbogast (Amsterdam: John Benjamins), 159–178.

Lu, C., and Proctor, R. (1995). The influence of irrelevant location information on performance: a review of the Simon and spatial strop effects. *Psychon. Bull. Rev.* 2, 174–207.

Macizo, P., and Bajo, M. T. (2006). Reading for repetition and reading for translation: do they involve the same processes? *Cognition* 99, 1–34.

Macizo, P., Bajo, M. T., and Martín, M. C. (2010). Inhibitory processes in bilingual language comprehension: evidence from Spanish-English interlexical homographs. *J. Mem. Lang.* 63, 232–244.

Maguirre, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S., and Frith, C. D. (2000). Navigation-related structural changes in the hippocampus of taxi drivers. *Proc. Natl. Acad. Sci. U.S.A.* 97, 4398–4403.

Martin-Rhee, M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., and Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature* 431, 757.

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching and naming: a metrical costs of language selection. *J. Mem. Lang.* 40, 25–40.

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

Miyake, A., Just, M. A., and Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. *J. Mem. Lang.* 33, 175–202.

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726.

Padilla, F., Bajo, M. T., and Macizo, P. (2005). Articulatory suppression in language interpretation: working memory capacity, dual tasking and word knowledge. *Biling. Lang. Cogn.* 8, 207–213.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. Lang. Cogn.* 13, 253–262.

Rodriguez-Fornells, A., Van der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., and Munte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J. Cogn. Neurosci.* 17, 422–433.

Ryan, J. D., Bialystok, E., Craik, F. I. M., and Logan, J. (2004). Bilingual advantages in attentional control. *Poster Presented at the Meeting of the Cognitive Neuroscience Society*, San Francisco, CA.

Salvatierra, J., and Roselli, M. (2010). The effect of bilingualism and age on inhibitory control. *Int. J. Biling* 15, 26–37.

Schneider, W., Eschman, A., and Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh, PA: Psychology Software Tools.

Simon, J. R. (1990). "The effects of an irrelevant directional cue on human information processing," in *Stimulus-Response Compatibility: An Integrated Perspective*, eds W. Proctor and T. G. Reeve (Amsterdam: North-Holland), 31–88.

Sims, V. K., and Mayer, R. E. (2002). Domain specificity of spatial expertise. The case of video game players. *Appl. Cogn. Psychol.* 16, 97–115.

Spelke, E., Hirst, W., and Neisser, U. (1976). Skills of divided attention. *Cognition* 4, 215–230.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 March 2011; accepted: 13 October 2011; published online: 28 October 2011.

*Citation: Yudes C, Macizo P and Bajo T (2011) The influence of expertise in simultaneous interpreting on non-verbal executive processes. Front. Psychology **2**:309. doi: 10.3389/fpsyg.2011.00309*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

ORIGINAL RESEARCH ARTICLE
published: 10 November 2011
doi: 10.3389/fpsyg.2011.00273

Bilingualism and creativity: benefits in convergent thinking come with losses in divergent thinking

Bernhard Hommel[1]*, Lorenza S. Colzato[1], Rico Fischer[2] and Ingrid K. Christoffels[1]

[1] Institute for Psychological Research and Leiden Institute for Brain and Cognition, Leiden University, Leiden, Netherlands
[2] Cognitive Psychology Unit, Technische Universität Dresden, Dresden, Germany

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
Julia Festman, Otto-von-Guericke University Magdeburg, Germany
Anatoliy Vladimirovich Kharkhurin, American University of Sharjah, United Arab Emirates

***Correspondence:**
Bernhard Hommel, Department of Psychology, Cognitive Psychology Unit, Leiden University, Wassenaarseweg 52, 2333 AK Leiden, Netherlands.
e-mail: hommel@fsw.leidenuniv.nl

Bilingualism is commonly assumed to improve creativity but the mechanisms underlying creative acts, and the way these mechanisms are affected by bilingualism, are not very well understood. We hypothesize that learning to master multiple languages drives individuals toward a relatively focused cognitive-control state that exerts strong top-down impact on information processing and creates strong local competition for selection between cognitive codes. Considering the control requirements posed by creativity tasks tapping into convergent and divergent thinking, this predicts that high-proficient bilinguals should outperform low-proficient bilinguals in convergent thinking, while low-proficient bilinguals might be better in divergent thinking. Comparing low- and high-proficient bilinguals on convergent-thinking and divergent-thinking tasks indeed showed a high-proficient bilingual advantage for convergent thinking but a low-proficient bilingual advantage for fluency in divergent thinking. These findings suggest that bilingualism should not be related to "creativity" as a unitary concept but, rather, to the specific processes and mechanisms that underlie creativity.

Keywords: bilingualism, creativity, divergent thinking, convergent thinking

INTRODUCTION

Increasing evidence suggests that speaking more than one language does not only improve one's verbal skills but also more general, non-linguistic cognitive abilities. For instance, bilingual individuals have been demonstrated to outperform monolinguals in problem solving (Bain, 1975), perceptual focusing (Duncan and De Avila, 1979), and the Simon task (e.g., Bialystok et al., 2004; for a general review, see Bialystok and Craik, 2010). According to a growing consensus, the bilingual benefit is related to executive control functions, which are assumed to improve by learning multiple languages. To account for the bilingual benefit, some of the earlier approaches have considered that dealing with a new language might require the suppression of the dominant language, which might imply improvements in inhibitory control (Green, 1998; Bialystok, 2001). Other approaches have argued that preventing conflict between languages does not necessarily require direct inhibition but the combination of attentional top-down biasing together with local competition (i.e., direct interactions between alternative cognitive codes) may do (Poulisse and Bongaerts, 1994; Dijkstra and van Heuven, 1998; La Heij, 2005; Bialystok et al., 2006). This latter approach has received support from the observation that bilinguals are no more efficient in inhibiting unwanted responses than monolinguals but are less efficient than monolinguals in distributing attentional resources over multiple visual target events (Colzato et al., 2008). This suggests that learning multiple languages does not improve inhibitory skills but, rather, leads to a stronger, more selective focusing of cognitive control (Colzato et al., 2008).

The aim of the present study was to further characterize this focusing of control by investigating the impact of bilingualism on different types of creativity. Authors have argued at length about how the concept of creativity should be defined, whether creativity research should focus on the creative individual, the creative act, or the cognitive processes leading to it, and there is accordingly no consensus as to how creativity should be measured (for an overview, see Runco, 2007). Massive research from the last 40 years or so provides strong evidence that bilingualism somehow supports creativity, and a recent report of the European commission has listed more than 200 articles demonstrating this connection (European Commission, 2009). Unfortunately, however, the methodological diversity and sample characteristics of these studies are enormous, which renders it more than questionable whether they were actually assessing the same construct and processes. Moreover, there is still no mechanistic model explaining how creative processes operate and how bilingualism might affect these operations, which in view of the lack of conceptual clarity may not be surprising.

To address this issue, we tried to avoid addressing creativity as a whole. Instead, we compared high-proficient bilinguals with low-proficient bilinguals in two tasks that are likely to represent relatively process-pure measures of components of creativity: divergent thinking and convergent thinking. We do not claim that these are the only processes involved in creative acts (even though Guilford, 1967, considers them the by far most important) nor that individuals showing good performance with regard to these components need to be considered creative in general. Rather, we considered these two components and the two related assessments tasks as – in contrast to other components and tasks, and to creativity as a whole – the cognitive-control operations they are likely to rely on are relatively well understood (Hommel et al., manuscript submitted). Moreover, the fact that they are uncorrelated (Akbari Chermahini and Hommel, 2010) suggests that they are measuring different components of creativity indeed.

Divergent thinking can be defined as the process that allows people to generate as many responses as possible based on relatively weak constraints. As an example, in Guilford's (1967) alternate uses task (AUT) people are presented with a simple object, such as a pen, and asked to generate as many uses for that object they can think of. The results are commonly scored regarding the number of responses (fluency), the number of different categories being used (flexibility), the degree to which the responses differ from the standard or group mean (originality), and the amount of detail (elaboration). In contrast, convergent thinking can be defined as a more strongly constrained process that searches for one possible outcome. As an example, in Mednick's (1962) remote associates task (RAT) people are presented with three concepts, such as "hair," "stretch" and "time," and they are to identify the one concept that fits with all three in terms of association, meaning, or abstraction, such as "long" in the example.

Even though one can argue that both types of processes and tasks share a number of aspects, they are likely to require, or at least benefit from two different configurations of cognitive control. In its most elementary form, any type of biologically plausible decision-making can be considered a competition between alternative codes or representations (Bogacz, 2007) – such as between representations of the English word "frog" and the semantically equivalent Dutch word "kikker" in a picture-naming English–Dutch bilingual, between representations of the alternative uses of a "pen" in the AUT, or between representations of close associates of the three defining words in a trial of the RAT. This kind of "local competition," as we will call it, is likely to generate random results unless it is steered by the current task goal. Duncan et al. (1996) have suggested that the impact of task goals on behavioral control might consist in providing top-down support for those representations that are most consistent with the current goal, so that the competition between cognitive representations can be considered a top-down "biased competition." As suggested by Colzato et al. (2008), people might differ with respect to the degree to which they experience local competition and/or the degree to which they bias this competition by top-down processes. Consider what possible differences regarding cognitive-control states (strong versus weak top-down bias and/or local competition) imply for different types of creativity tasks.

Divergent thinking (as assessed by the AUT) is likely to benefit from a cognitive-control state that provides a minimum of top-down bias and local competition, so that the individual can easily and quickly "jump" from one thought to the other in an only weakly guided fashion (Hommel et al., manuscript submitted). In contrast, convergent thinking (as assessed by the RAT) is likely to benefit from strong top-down bias (that is representing the greater number of constraints that possible solutions need to meet) and strong local competition (as only one solution can be right). If so, engaging in divergent thinking should facilitate subsequent performance in tasks that require weak, "distributed" control while engaging in convergent thinking should be beneficial for subsequent performance in tasks requiring a more focused, exclusive control style, that is, strong top-down control and local competition. Indeed, previous divergent thinking was found to improve performance in tasks that require the distribution of attention to two successive visual targets (Hommel et al., manuscript submitted) and increased inter-task interactions between two overlapping tasks (Fischer and Hommel, submitted), while previous convergent-thinking improved performance in selective-attention and response-competition tasks (Hommel et al., manuscript submitted).

Relating the findings from research on bilinguals to the observations from creativity studies suggests that the cognitive-control style that seems to be acquired by learning multiple languages fits well with the style implied by convergent thinking. If so, a straightforward hypothesis presents itself: high-proficient bilinguals should outperform low-proficient bilinguals in convergent thinking while the opposite should be the case with divergent thinking. At first sight, this hypothesis seems to be disproved by the available evidence. Numerous studies have claimed that bilingualism has a specific, positive effect on divergent thinking (for an overview, see Ricciardelli, 1992). However, previous studies have a number of characteristics that make it difficult to relate them to our hypothesis, such as reporting only aggregated total scores of creativity (across often heterogeneous scales) and the use of tasks that are unlikely to provide sufficiently process-pure measurements of convergent versus divergent thinking. For instance, the recent studies of Kharkhurin (2009, 2010) employed the abbreviated torrance test for adults (ATTA, Goff and Torrance, 2002) to assess divergent thinking. The test consisted of three activities: identifying the troubles that one may encounter when walking on air, completing incomplete pictures, and drawing as many pictures as possible based on a given group of triangles. Even though all these activities certainly require some form of creativity, they all seem to be much more balanced with respect to the generation aspect (that would benefit from weak top-down control and local competition) and the number of constraints (which call for strong top-down control and local competition) than the AUT (which has fewer constraints) on the one hand and the RAT (which allows for one result only) on the other. Hence, the ATTA arguably mixes divergent and convergent operations more than necessary, which makes it difficult to make predictions from a control-state point of view. Moreover, Ricciardelli (1992) already identified a number of studies that did not show the expected divergent-thinking benefits in bilinguals or even an advantage for monolinguals. More recently, two other studies reported better performance in monolinguals than bilinguals in a verbal fluency task (requiring to generate as many exemplars of a given category as possible) that bears some similarities to the divergent-thinking AUT (Rosselli et al., 2000; Gollan et al., 2002). Hence, a closer look reveals that the evidence for better divergent thinking in bilinguals is rather mixed (we consider some possible reasons in the Discussion), which is why we considered it still reasonable to compare high-proficient with low-proficient bilinguals in divergent thinking (assessed by means of the AUT) and convergent thinking (assessed by means of the RAT) separately.

MATERIALS AND METHODS

PARTICIPANTS

Forty-two young healthy adults served as participants for partial fulfillment of course credit or a financial reward and constituted the two language groups: low-proficient and high-proficient

Table 1 | Demographic data measures, vocabulary proficiency scores, originality, fluency, flexibility, and elaboration scores from the alternate uses task (AUT), the number of correct items from the remote associates task (RAT), means, and SD for the low-proficient and high-proficient bilinguals are shown.

Sample	Low-proficient	High-proficient
N (F:M)	12:9	12:9
Age (in years)	19.6 (3.2)	20.7 (1.1)
IQ	121.8 (4.1)	123.7 (6.0)
English vocabulary score*	3754 (748)	4164 (520)
AUT		
Elaboration	4.2 (3.1)	3.3 (2.3)
Fluency*	40.7 (11.6)	32.5 (11.1)
Flexibility	27.2 (7.5)	26.4 (8.3)
Originality	17.4 (16.1)	12.8 (6.7)
RAT*	13.3 (3.1)	15.8 (4.2)

**p* < 0.05 (significant group difference).*

bilinguals. All reported having normal or corrected-to-normal vision, and were not familiar with the purpose of the experiment. Half of the participants were low-proficient bilingual German native-speaker students of the Technische Universität Dresden (Germany), and the other half were Dutch–English high-proficient bilinguals living in the Netherlands. All participants were tested by the same instructional protocols, although the actual testing was carried out in two different countries. The high-proficient bilingual participants attended the high school in English and some of them had lived part of their life in an English-speaking country. They used both Dutch and English on a daily basis throughout their lives. As research with bilingual adults (Kroll and Stewart, 1994) and children (Bialystok, 1988) has revealed that the cognitive and linguistic consequences of bilingualism are more salient for those bilinguals who are relatively balanced in their proficiency, we only considered balanced bilinguals for the present study. The low-proficient bilingual German participants were not functionally fluent in any other language despite the inevitable language courses in school. All participants in both groups attended university and shared similar middle-class socioeconomic backgrounds, and they were matched for age, sex, and IQ (measured by Raven's standard progressive matrices, SPM), see **Table 1**. Participants gave their written informed consent prior to their inclusion in the study in accordance to the 1964 Declaration of Helsinki.

PROCEDURE AND DESIGN

The experiment consisted of a 45-min session in which participants completed the AUT to assess divergent thinking, the RAT to assess convergent thinking, a vocabulary test to access participants' proficiency in English, and a short-version of a reasoning-based intelligence test (Raven's SPM; Raven et al., 1988). After completion of the tasks, the participants were debriefed and paid.

AUT (divergent thinking)

In this task (based on Guilford, 1967, and translated into Dutch and German), participants were asked to list as many possible uses for three common household items (brick, shoe, and newspaper) as they can within 10 min. Scoring comprised of four components:

Originality: Each response is compared to the total amount of responses from all of the subjects. Responses that were given by only 5% of the group count as unusual (1 point) and responses given by only 1% of them count as unique (2 points).

Fluency: The total of all responses.

Flexibility: The number of different categories used.

Elaboration: The amount of detail (e.g., "a doorstop" counts 0, whereas "a door stop to prevent a door slamming shut in a strong wind" counts 2 (1 point for explanation of door slamming and another for further detail about the wind).

RAT (convergent thinking)

In this task (based on Mednick, 1962, and translated into Dutch and German), participants were presented with three unrelated words (such as time, hair, and stretch) and asked to find a common associate (long). Our version comprised of 30 items, which were to be worked through within 10 min.

English vocabulary test

Participants' proficiency in English was assessed with a vocabulary test consisting of a non-speeded lexical decision task (Christoffels et al., 2006, manuscript submitted). The test consisted of words selected from five different word frequency bins and non-words. Participants were required to indicate whether or not they knew the meaning of the English letter strings. The score of the test ranges from 0 to 5,000 and is corrected for misattribution of non-words.

SPM (intelligence test)

The SPM assesses the individual's ability to create perceptual relations and to reason by analogy independent of language and formal schooling; it is a standard, widely used test to measure Spearman's *g* factor and of fluid intelligence in particular.

STATISTICAL ANALYSIS

Independent t-tests were performed to test differences between the two groups. From the two tasks, five measures were extracted for each participant: originality, fluency, flexibility, and elaboration scores from the AUT, the number of correct items from the RAT. The measures were scored by two independent readers (Cronbach's alpha = 0.90). A significance level of $p < 0.05$ was adopted for all tests.

RESULTS

No significant group differences were obtained for age, $t(40) = 1.47$, $p = 0.15$, and intelligence, $t(40) = 1.17$, $p = 0.25$. As expected, high-proficient bilinguals were significantly more proficient in the English vocabulary test than low-proficient bilinguals, $t(40) = 2.06$, $p < 0.05$, see **Table 1**.

Performance in the RAT and AUT was good and comparable to performance in other studies (e.g., Akbari Chermahini and Hommel, 2010). As expected, high-proficient bilinguals showed better performance in the RAT task than low-proficient bilinguals, $t(40) = 2.22$, $p < 0.05$. Also as expected, all four scores of the AUT showed an advantage for low-proficient over high-proficient bilinguals. While this advantage did not reach significance for flexibility, $t(40) < 1$, originality, $t(40) = 1.19$, $p = 0.24$, and elaboration, $t(40) = 1.10$, $p = 0.28$, it was reliable for fluency, $t(40) = 2.31$, $p < 0.05$, see **Table 1**.

DISCUSSION

The guiding hypothesis of this study assumes that speaking multiple languages leads to the adoption of a relatively focused cognitive-control state, at least as a default, which is characterized by a strong top-down biasing of information processing and strong local competition for selection between cognitive codes. Considering that this control-state fits with the control requirements of convergent thinking but not of divergent thinking, we predicted that high-proficient bilinguals should outperform low-proficient bilinguals in convergent thinking, whereas low-proficient bilinguals should perform better in divergent thinking. And this is exactly what the data show: high-proficient bilinguals excel in convergent thinking while low-proficient bilinguals excel in divergent thinking, at least with regard to the fluency score.

On the one hand, this outcome is consistent with previous observations that monolinguals outperform bilinguals in verbal fluency tasks (Rosselli et al., 2000; Gollan et al., 2002). On the other hand, however, it does not seem to fit with the general expectation that bilingualism is associated with greater cognitive flexibility and numerous studies that seem to support this expectation (cf., Ricciardelli, 1992; European Commission, 2009). One possible interpretation, which has been suggested by Kharkhurin (2010), is that verbal and non-verbal creativity tasks assess different skills and might be differently affected by bilingualism. Kharkhurin's specific suggestion is that bilinguals might excel in the non-verbal domain but be outperformed by monolinguals if it comes to verbal domains. However, not only is this suggestion difficult to combine with the fact that bilinguals must master language-related skills that monolinguals do not need to master, which again implies that bilinguals do have some unique expertise that relates to the verbal domain. It is also refuted by our findings. Both of our thinking tasks were clearly verbal and drawing on verbal skills, so that the better performance of bilinguals in the convergent-thinking task does not fit with the claim that bilinguals' performance is necessarily inferior to that of monolinguals. But, on the other hand, it is true that we demonstrated the predicted double dissociation for the verbal domain only, and it may very well be that non-verbal tasks show a different pattern.

Another aspect that might explain at least some inconsistencies in the field in general and with regard to the present findings in particular relates to the age of the investigated participants. The majority of studies on the relationship between bilingualism and creativity have focused on children (Kharkhurin, 2010). In line with earlier speculations (e.g., Eysenck, 1993; Ashby et al., 1999), there is increasing evidence for a reliable connection between dopamine and creativity performance, which for instance has been shown to vary with the individual density of dopamine receptors (de Manzano et al., 2010), genetic variability associated with striatal dopamine production (Reuter et al., 2006), and the Parkinson-related loss of dopaminergic neurons in the striatal pathway (Batir et al., 2009). Using the same tasks used in the present study, it could be demonstrated in healthy subjects that convergent and divergent thinking are mediated by both the individual tonic dopamine level (Akbari Chermahini and Hommel, 2010) and phasic changes of this level (Akbari Chermahini and Hommel, submitted). Interestingly, the brain systems that are targeted by the two major dopaminergic pathways in humans (the frontal and the striatal pathway; see Cools, 2008) are particularly strongly affected by developmental factors and are suspected to keep developing into early adulthood (Gogtay et al., 2004). This fits with observations of considerable variability of individual creativity measures over the lifespan (e.g., Simonton, 1997; Wohl, 2003). Hence, it remains to be seen whether observations from children really generalize to adults, and whether individual differences in creativity tasks are stable over time.

Finally, we would like to emphasize that we do not consider our measures of performance in divergent and convergent-thinking tasks to represent creativity as a whole – be it as a state or a personal trait. Divergent and convergent thinking are likely to be very important, if not crucial (Guilford, 1967), for many creative acts, but such acts can be suspected to comprise of a whole sequence of processes and components. Many authors since Wallas (1926) have assumed that creative acts run through at least four stages including *preparation*, which involves investigating the problem; *incubation*, which involves (often unconscious) thinking about the problem; *illumination*, where ideas come together to form a possible solution; and *verification*, which involves evaluating the chosen option. It makes sense to characterize the first two stages as emphasizing divergent processes and the final two stages as emphasizing convergent processes (Hommel, in press), and it may very well be that individual performance therein can be predicted to some degree based on the measures used in the present study. Nevertheless, the complexity of the respective stages strongly suggests that a number of other cognitive operations are involved. Accordingly, it would be far-fetched to consider good performance in the AUT and the RAT sufficient to categorize individuals as "creative." Indeed, we sincerely believe that unpacking complex concepts like "creativity" into their component processes represents a crucial step in constructing mechanistic models that are sufficiently transparent to undergo rigorous empirical testing.

And the same goes for the cognitive benefits that come along with bilingualism: modeling them in detail will also require the unpacking into component processes. The present findings suggest that one of these component processes is responsible for the regulation of cognitive-control states. If we assume that one dimension on which these states vary relates to the degree of top-down biasing of information processing and of local competition between alternative cognitive codes (Colzato et al., 2008; Hommel et al., manuscript submitted), having learned to handle multiple tasks seems to drive individuals toward the pole of the dimension that represents more biasing and more competition. Accordingly, bilinguals are likely to show particularly good performance on tasks that are relying or benefiting from this control state but to show relatively poor performance on tasks that require or benefit from weaker top-down control and less local competition. Examples for the former are tasks inducing response conflict (Bialystok et al., 2004) and convergent thinking, while examples for the latter are tasks that call for the distribution of attention (Colzato et al., 2008) and divergent thinking.

ACKNOWLEDGMENTS

The research of Lorenza S. Colzato is supported by NWO (Netherlands Organization for Scientific Research).

REFERENCES

Akbari Chermahini, S., and Hommel, B. (2010). The (b)link between creativity and dopamine: spontaneous eye blink rates predict and dissociate divergent and convergent thinking. *Cognition* 115, 458–465.

Ashby, F. G., Isen, A. M., and Turken, A. U. (1999). A neuro-psychological theory of positive affect and its influence on cognition. *Psychol. Rev.* 106, 529–550.

Bain, B. (1975). Toward an integration of Piaget and Vygotsky: bilingual considerations. *Linguistics* 16, 5–20.

Batir, A., Lhommee, E., Ardouin, C., Fraix, V., Seigneuret, E., Chabardes, S., Benabid, A.-L., Pollak, P., and Krack, P. (2009). Creativity induced by dopamine agonists in Parkinson's disease. *Mov. Disord.* 24, 234.

Bialystok, E. (1988). Levels of bilingualism and levels of linguistic awareness. *Dev. Psychol.* 24, 560–567.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge: Cambridge University Press.

Bialystok, E., and Craik, F. I. M. (2010). Cognitive and linguistic processing in the bilingual mind. *Curr. Dir. Psychol. Sci.* 19, 19–23.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified anti-saccade task: effects on aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354.

Bogacz, R. (2007). Optimal decision-making theories: linking neurobiology with behavior. *Trends Cogn. Sci. (Regul. Ed.)* 11, 118–125.

Christoffels, I. K., de Groot, A. M. B., and Kroll, J. F. (2006). Memory and language skills in simultaneous interpreting: expertise and language proficiency. *J. Mem. Lang.* 54, 324–345.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S. T., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

Cools, R. (2008). Role of dopamine in the motivational and cognitive control of behaviour. *Neuroscientist* 14, 381–395.

de Manzano, O., Cervenka, S., Karabanov, A., Farde, L., and Ullén, F. (2010). Thinking outside a less intact box: thalamic dopamine D2 receptor densities are negatively related to psychometric creativity in healthy individuals. *PLoS ONE* 5, e10670. doi:10.1371/journal.pone.0010670

Dijkstra, T., and van Heuven, W. J. B. (1998). "The BIA model and bilingual word recognition," in *Localist Connectionist Approaches to Human Cognition*, eds J. Grainger and A. Jacobs (Hillsdale, NJ: Erlbaum), 189–225.

Duncan, J., Emslie, H., Williams, P., Johnson, R., and Freer, C. (1996). Intelligence and the frontal lobe: the organization of goal-directed behavior. *Cogn. Psychol.* 30, 257–303.

Duncan, S. E., and De Avila, E. (1979). Bilingualism and cognition: some recent findings. *NABE J.* 4, 15–50.

European Commission. (2009). *Study on the Contribution of Multilingualism to Creativity: Final Report*. Available at: http://eacea.ec.europa.eu/llp/studies/study_on_the_contribution_of_multilingualism_tt_creativity_en.php

Eysenck, H. J. (1993). Creativity and personality: suggestions for a theory. *Psychol. Inq.* 4, 147–178.

Goff, K., and Torrance, E. P. (2002). *Abbreviated Torrance Test for Adults*. Bensenville, IL: Scholastic Testing Service.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., Nugent, T. F., Herman, D. H., Clasen, L. S., Toga, A. W., Rapoport, J. L., and Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101, 8174–8179.

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

Green, D. W. (1998). Mental control of the bilingual lexico-sematic system. *Biling. (Camb. Engl.)* 1, 67–81.

Guilford, J. P. (1967). *The Nature of Human Intelligence*. New York: McGraw-Hill.

Hommel, B. (in press). "Convergent and divergent operations in cognitive search," in *Cognitive search: Evolution, Algorithms, and the Brain. Strüngmann Forum Reports,* Vol. 9, eds P. M. Todd, T. T. Hills, and T. W. Robbins (Cambridge, MA: MIT Press).

Kharkhurin, A. V. (2009). The role of bilingualism in creative performance on divergent thinking and invented alien creatures tests. *J. Creat. Behav.* 43, 59–71.

Kharkhurin, A. V. (2010). Bilingual verbal and nonverbal creative behavior. *Int. J. Biling.* 14, 211–226.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174.

La Heij, W. (2005). "Lexical selection in monolinguals and bilinguals," in *Handbook of Bilingualism*, eds J. F. Kroll and A. M. B. de Groot (New York: Oxford University Press), 289–307.

Mednick, S. (1962). The associative basis of creative problem solving process. *Psychol. Rev.* 69, 200–232.

Poulisse, N., and Bongaerts, T. (1994). First language use in second language production. *Appl. Linguist.* 15, 36–57.

Raven, J. C., Court, J. H., and Raven, J. (1988). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. London: Lewis.

Reuter, M., Roth, S., Holve, K., and Hennig, J. (2006). Identification of first candidate genes for creativity: a pilot study. *Brain Res.* 1069, 190–197.

Ricciardelli, L. A. (1992). Creativity and bilingualism. *J. Creat. Behav.* 26, 242–254.

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., and Ostrosky-Solís, F. (2000). Verbal fluency and repetition skills in healthy older Spanish–English bilinguals. *Appl. Neuropsychol.* 7, 17–24.

Runco, M. A. (2007). *Creativity. Theories and Themes: Research, Development, and Practice*. San Diego, CA: Academic Press.

Simonton, D. K. (1997). Creative productivity: a predictive and explanatory model of career trajectories and landmarks. *Psychol. Rev.* 104, 66–89.

Wallas, G. (1926). *The Art of Thought*. New York: Harcourt Brace.

Wohl, E. (2003). *Creativity and Affective Traits Across the Lifespan: Developmental Influences Among Adolescents and Older Adults*. Ann Arbor, MI: UMI/ProQuest.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 May 2011; accepted: 29 September 2011; published online: 10 November 2011.

Citation: Hommel B, Colzato LS, Fischer R and Christoffels IK (2011) Bilingualism and creativity: benefits in convergent thinking come with losses in divergent thinking. Front. Psychology ***2****:273. doi: 10.3389/fpsyg.2011.00273*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

ORIGINAL RESEARCH ARTICLE
published: 18 April 2012
doi: 10.3389/fpsyg.2012.00115

Cognitive control in Russian–German bilinguals

Julia Festman[1]* and Thomas F. Münte[2]

[1] Potsdam Research Institute for Multilingualism, University of Potsdam, Potsdam, Germany
[2] Department of Neurology, University of Lübeck, Lübeck, Germany

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
Susanne Maria Reiterer, University of Tuebingen, Germany
Zofia Wodniecka, Jagiellonian University, Poland

***Correspondence:**
Julia Festman, Potsdam Research Institute for Multilingualism, University of Potsdam, Haus 2, Campus Golm, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany.
e-mail: festman@uni-potsdam.de

Bilingual speakers are faced with the problem to keep their languages apart, but do so with interindividually varying success. Cognitive control abilities might be an important factor to explain such interindividual differences. Here we compare two late, balanced and highly proficient bilingual groups (mean age 24 years, L1 Russian, L2 German) which were established according to their language control abilities on a bilingual picture-naming task. One group had difficulties to remain in the instructed target language and switched unintentionally to the non-target language ("switchers"), whereas the other group rarely switched unintentionally ("non-switchers"). This group-specific behavior could not be explained by language background, socio-cultural, or demographic variables. Rather, the non-switchers also demonstrated a faster and better performance on four cognitive control tests (Tower of Hanoi, Ruff Figural Fluency Test, Divided Attention, Go/Nogo). Here, we focus on two additional executive function tasks, the Wisconsin Card Sorting Test (WCST) and the Flanker task requiring conflict monitoring and conflict resolution. Non-switchers outperformed switchers with regard to speed and accuracy, and were better at finding and applying the correct rules in the WCST. Similarly, in the Flanker task non-switchers performed faster and better on conflict trials and had a higher correction rate following an error. Event-related potential recordings furthermore revealed a smaller error-related negativity in the non-switchers, taken as evidence for a more efficient self-monitoring system. We conclude that bilingual language performance, in particular switching behavior, is related to performance on cognitive control tasks. Better cognitive control, including conflict monitoring, results in decreased unintentional switching.

Keywords: Flanker task, ERN, Wisconsin Card Sorting Test, conflict monitoring, inhibition, late bilinguals, cognitive control, executive function

INTRODUCTION

The recent interest in the relation between bilingual language processing and non-linguistic control abilities has been fueled by research showing that bilinguals outperform monolinguals on tasks involving executive functions. These are often distinguished into three subdomains: inhibition, shifting of mental sets (also referred to as task switching or cognitive flexibility), and updating information in working memory (Miyake et al., 2000). Inhibition is necessary to resist distraction in order to stay focused on the currently relevant task or information. Inhibition thus enables us to perform goal-driven, task-relevant, and appropriate behavior (as regards social context, communication constraints, etc.). Set shifting refers to the ability to detach oneself from one task in order to turn to something else. It reflects the ability to adjust to changing demands, priorities, or information. In working memory we hold currently necessary important information and can constantly add more, newer information which might complete, replace, or change previous information.

The explanatory link between cognitive control and bilingualism is that early bilinguals have to constantly control interfering information from the two active and competing language systems which might train and enhance their cognitive control abilities (Martin-Rhee and Bialystok, 2008). Thus, it appears that bilingualism taxes inhibitory control by requiring speakers to suppress one language when using another, cognitive flexibility by requiring speakers to switch between languages, and working memory by having to keep track of swift changes in multilingual communication, e.g., which language is most appropriate with whom.

Research in executive function abilities in relation to bilingualism is a rather new field with an interdisciplinary and multimethodological approach. Whereas mostly dichotomous groups such as balanced and unbalanced bilinguals (e.g., Vega and Fernandez, 2011) have been used in this research, individual differences have been largely neglected. Reiterer (2009) has listed a number of "factors that matter" in language acquisition and ultimate attainment of L2 proficiency, including biological, psychological, linguistic, and socio-cultural variables. Based on brain imaging studies, she concluded that brain organization is highly dependent on a number of these factors. For example, individual differences in response inhibition are correlated with differences in functional magnetic resonance imaging (fMRI) activation patterns related to inhibitory control (Garavan et al., 2006). Ye and Zhou (2008) grouped monolingual students according to their performance in a color–word Stroop task into readers with higher and lower control abilities, which were found to modulate the resolution of conflict between sentential representations. Consequently, we suggest that individual differences in executive functions may also play a major role in bilingual language performance.

Children who acquire two languages early in life develop the ability to solve problems that contain conflicting or misleading cues better than their monolingual peers (Martin-Rhee and Bialystok, 2008). Carlson and Meltzoff (2008) examined three groups of kindergarten children. The bilingual group outperformed the monolingual and the English-language-learner group on a variety of conflict tasks. In a recent study in young Spanish undergraduate students (Costa et al., 2009) bilinguals responded generally faster across trial types during high conflict-monitoring conditions in the attention network task (ANT), but did not show an advantage on low conflict-monitoring conditions. In sum, bilinguals are better in "conflict" situations which require the ability to resolve interference among competing representations and thus parallel the situation in which two languages compete and create a conflict for selection in bilingual speech production.

The mechanisms of executive control responsible to resolve conflict in language tasks have been suggested to be similar to those engaged in other cognitive domains (for a review see Abutalebi and Green, 2007; Ye and Zhou, 2009). Bilinguals who frequently use both languages train conflict resolution constantly. Due to the parallel activation of both languages in bilinguals, conflict resolution seems to be inherent to "monolingual language mode" production (for the language mode model, see Grosjean, 1982), i.e., when only one of these languages is required for verbal output. As a consequence, bilinguals are more efficient in dealing with conflicting and distracting information also in other domains, e.g., incongruent flanking information or bivalent displays (see Bunge et al., 2002).

While an advantage on conflict trials can be accounted for by extensive training in resolving the conflict produced by incompatible competing representations or responses (Abutalebi and Green, 2007), the advantage on overall reaction time (RTs) found in many (but not all) bilingual studies can not, as many trials do not require conflict resolution. As an explanation, Costa et al. (2009) suggested that the bilinguals' monitoring process also kicks in on congruent trials and is responsible for the RT advantage on these trials. According to models of bilingual language production, first, one of the two available language schemas needs to be selected (Green, 1998). Costa et al. (2009) suggested that the bilingual monitoring system controls for the continuous use of the initially selected language in further processing stages. This monitoring activity for production in the target language is only necessary in bilinguals but not in monolinguals and could account for the bilingual RT advantage. Lexical competition according to Costa et al. (2009) can be thought of as the bilinguals' training stage for conflict resolution on conflict trials.

As a first step toward a more fine-grained look at the bilingual population we previously investigated individual differences in language control abilities (i.e., switching between languages on command) in a group of Russian–German bilinguals with high and balanced proficiency in both languages (Festman, 2011). Bilinguals were grouped according to their errors of cross-language interference (CLI) on a bilingual picture-naming task. Those who switched – although not required – were called "switchers," those who did not switch were called "non-switchers." A number of additional executive function tasks (Tower of Hanoi, Ruff Figural Fluency Task, Divided Attention, Go/Nogo) taxed inhibition of irrelevant information, problem solving, planning efficiency, generative fluency, and self-monitoring (Festman et al., 2010). A strong relationship between language control abilities and executive functions could be established in that "switchers" performed worse on all executive function tests compared to "non-switchers." This suggests that these groups may differ in their susceptibility to CLI because of individual differences in executive control functions. In Festman et al. (2010), we observed that non-switchers were significantly better able to produce correct responses in the verbal part of the WAIS, but not in the general performance part of the WAIS. Intelligence is thus not a likely candidate to explain between-group differences in all other cognitive tasks. We speculated that non-switchers were more efficient in suppressing irrelevant and conflicting information and thereby reduce response conflict earlier than in the course of switchers' processing. This advantage might have helped to facilitate response selection and response execution. Some indications of a likely difference in these processes have been found in our earlier tasks already, as non-switchers showed better monitoring abilities in the Tower of Hanoi task (less errors) and in the Ruff Figural Fluency task (fewer perseverations).

To investigate conflict monitoring and conflict resolution more directly, the Wisconsin Card Sorting Test (WCST) and the Eriksen flanker task were employed in the current study. The WCST requires participants to select and apply rules in order to sort trivalent displays (color, shape, number are the three features that constitute each stimulus). More importantly, on every few trials the participant has to shift rules according to cues while on the remaining trials he has to stick to the same rule as used for the previous trial (similar to the required use of one language until language change is signaled). The WCST requires both, staying on a rule and shifting from one rule to another, and thus allows us to determine whether the results of the WCST parallel the groups' language control abilities. Accordingly, if (a) switchers have difficulties to remain in the target language, because the non-target language continues to cause strong conflict with target language production, this should be paralleled by errors on trials when no rule change is necessary. Non-switchers were thus expected to be better in continuous rule application. If (b) switchers have difficulties to switch to the other language, because they have problems to inhibit the current target language in order to switch to the new target language, this should be paralleled by perseveration errors in the WCST (switch/shift to the new rule while inhibiting the former rule). As non-switchers previously demonstrated superior inhibitory control abilities (Festman et al., 2010), they were expected to show superior set shifting abilities resulting in faster and better performance.

The flanker task provided several measures of executive control and conflict monitoring. First, performance on conflict ("incongruent") and no-conflict ("congruent") trials was compared between the groups. Second, the number of errors as well as the number of corrected errors was assessed. If non-switchers indeed have better conflict resolution abilities, we predicted less interference and hence faster performance on incongruent trials. Moreover, we expected higher correction rates for non-switchers than for switchers. The flanker task, while originally introduced for behavioral studies, has been used extensively in conjunction with event-related brain potentials (ERP) and fMRI over the past 10 years. In the ERP, an "error-related negativity" (ERN; Gehring et al., 1993) emerges in the response-locked averages which is generated when participants make errors. The ERN peaks around 50–80 ms post-error and has a frontocentral maximum. It has been interpreted as an on-line index of performance monitoring and has been related to response conflict (Swick and Turken, 2002).

MATERIALS AND METHODS

PARTICIPANTS

This study had been approved by the ethics committee of the Otto-von-Guericke University, Magdeburg, the affiliation of the authors at the time of the experiment, and was performed in accordance with the 1964 Declaration of Helsinki. Informed consent was obtained from every participant, and participants were paid for their participation. **Table 1** summarizes the characteristics of the switcher and non-switcher group which had been established based on their language control as described in the introduction.

WISCONSIN CARD SORTING TEST

The WCST requires to apply rules continuously, to perform organized search, as well as to shift cognitive sets (rule changes) and to use feedback from the environment (Spreen and Strauss, 1998). We used a simplified computer version of the WCST developed by Barceló et al. (2002) which was presented using Presentation software[1].

The trial structure is depicted in **Figure 1**. Participants were instructed to respond as quickly and as accurately as possible. After training for 6 rule changes (1 rule change after 5–7 trials), participants were asked to complete 3 blocks of 12 rule changes each with the possibility to rest between blocks.

Trials were analyzed according to their position in the run. We distinguished trials immediately following a "change-rule-sound." According to Barceló et al. (2002), these are three-dimensional shift trials (3D) during which participants have to handle three rules in working memory (i.e., inhibit the previous rule and consider the other two rules for responding). On the first trial after a rule change signal, the participant had a 50% chance of choosing an incorrect rule provided that he remembered the previous and thus irrelevant rule. If the participant chose the correct rule, he received positive feedback and had to continue with this rule until the next rule-change signal. The following trials would then be considered trials of the "subblock." If the first choice after the rule-change signal was wrong, negative feedback was given. After having discarded one of the three rules in the 3D-trial, only two rules had to be dealt with on two-dimensional shift trials (2D), i.e., the incorrect rule had to be inhibited and the only remaining option had to be selected and applied until the next change signal. This is a very efficient trial-and-error process in normal subjects, who can

Table 1 | Participant information.

	Switcher	Non-switcher
N	13	16
Gender	11 women, 2 men	10 women, 6 men
Mean age	26 (6.7)	23 (3.1)
Main languages	L1 Russian, L2 German	L1 Russian, L2 German
Age at acquisition of L2	13 (8.1)	11 (4.1)
In Germany since	8 years	10 years
Language proficiency	Same in L1 and L2	Same in L1 and L2
RT Russian non-sw trials	1174 ms (201)	1081 ms (144)
RT German non-sw trials	1164 ms (125)	1093 ms (139)
WAIS-picture completion, correct	14.4 (1.6)	15.1 (1.3)
WAIS-block design, points	35.4 (9.6)	38.1 (7.9)
Language control ability	Weak	Strong
Bilingual interview, no. of CLI	11.4 (10.6)	3.4 (4.7)**
Cognitive control ability	Weak	Strong
ToH moves	43.8 (11.7)	29.3 (12.8)**
ToH errors	515 (684)	119 (172)*

*WAIS, Wechsler Adult Intelligence Scale; RT, reaction time, CLI, cross-language interference; ToH, Tower of Hanoi; significant differences between groups are indicated by an asterisk with **$p < 0.01$, and *$p < 0.05$.*

[1] www.neurobs.com

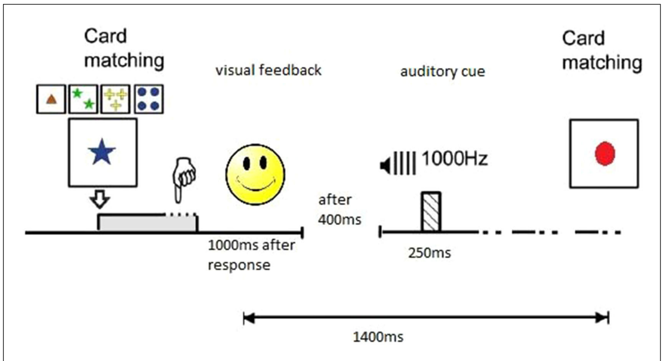


FIGURE 1 | Example trial: After fixation (1000–1700 ms) four cards were presented in one line in the center of a computer screen. Below this display, one larger card (e.g., one blue star) was shown. This critical card had to be sorted according to color, shape, or number and was presented until response was given. Response buttons were assigned in the following way: In the case of the blue-star-card, the number key "1" (depicting one red triangle) should be used to indicate the number rule, "2" (two green stars) to sort it according to the shape rule, "3" (three yellow crosses) would in this case not offer any sorting rule, and "4" (four blue circles) to sort according to the color information. Every critical card differed in color, shape, and number from the previously presented critical cards. One of two different feedback signs (happy or sad "smiley" icons) followed 1000 ms after the response. On all non-rule-change trials an auditory cue of 1000 Hz was presented ("use-the-same-rule-sound"). A new card to be sorted was displayed at 1400 ms after visual feedback. The rule to be applied changed after five to seven trials (e.g., from "sort according to color" to "sort according to shape"). Rule change was indicated by a "change-rule-sound" (500 Hz, 400 ms after the visual feedback, lasting for 250 ms), and participants had to find the correct new rule by trial-and-error.

use past contextual information to optimize set shifting (Barceló et al., 2002).

FLANKER EXPERIMENT

The flanker task (Eriksen and Eriksen, 1974) requires responding to the center letter of a five letter array with either a left-hand (for letter H) or right-hand response (letter S). Additional letters flanking the target letter either favored the target response (congruent trials, HHHHH or SSSSS) or primed the other response (incongruent trials, HHSHH or SSHSS). To increase the number of errors produced 60% of the trials were incongruent. Each stimulus array subtended about 2.5° of visual angle in width, and a fixation cross was presented in the middle of the computer monitor just below the target letter in the array (using Presentation, Neurobehavioral Systems). Each stimulus was presented for 100 ms and a stimulus-onset-asynchrony of 900 ms was used. Letter/hand assignments were counterbalanced between subjects and maintained in both sessions. Prior to the experiment, participants were trained in a short session of 6 blocks of 40 trials each to reach a RT baseline level and were given feedback about their performance. The goal of this procedure was to aim for a reaction time that would yield approximately 10% of errors. The experiment proper consisted of 20 blocks of 200 stimuli each. A 30-s rest period was allowed between blocks. Subjects were required to respond to the stimuli as fast as possible and to correct their errors as fast as possible whenever they detected them.

The electroencephalogram was recorded from 29 tin electrodes mounted in an electro cap against a reference electrode placed on the left mastoid process. Biosignals were re-referenced off-line to the mean of the activity at the two mastoid processes. Blinks and vertical eye movements were monitored with electrodes placed at the sub and supraorbital ridge of the left eye. Lateral eye movements were monitored by a bipolar montage using two electrodes placed on the right and left external canthus. Eye movements were recorded in order to allow for later off-line rejection, which was carried out by a computer program based on an amplitude criterion (75 μV). All electrode impedances were kept below 4 kΩ. Electrophysiological signals were amplified with a band-pass filter of 0.01–50 Hz and digitized at a rate of 250 Hz (4 ms resolution). Data were analyzed using the Event-Related Potential Software System (ERPSS)[2].

RESULTS

WISCONSIN CARD SORTING TEST

Response times showed a gradual speed-up from 3D to 2D trials to subblock trials (see **Figure 2**). Group differences were significant on all trial positions with non-switchers being faster (for 3D $U = 48.0$, $p = 0.006$; for 2D $U = 65.0$, $p = 0.0425$; for subbl. $U = 50.0$, $p = 0.008$).

Switchers committed significantly more errors than non-switchers (**Table 2**) with the majority of their errors occurring for subblock stimuli.

FLANKER TASK

The typical general pattern of results was obtained with more errors [$t(28) = -6.631$, $p < 0.0001$], and slower responses [$t(28) = -7.983$, $p < 0.0001$] on incongruent trials. Importantly, switchers showed a greater interference effect for incongruent stimuli, evidenced by a slower response time for the type of stimulus but not for congruent stimuli (**Table 3**). Moreover, switchers corrected their errors significantly less often than non-switchers, indicating worse self-monitoring abilities. There were no group differences for overall accuracy.

Response-locked averages revealed a typical ERN for the error trials peaking at about 70 ms (**Figure 3**), which was larger in switchers.

The ERN was quantified by a mean amplitude measure (time-window 0–100 ms) which was analyzed by ANOVA. A highly significant main effect was obtained for response type [$F(1, 24) = 31.5$, $p < 0.001$] as well as an interaction between group and response type [$F(1, 24) = 11.6$, $p < 0.001$]. *Post hoc* analyses revealed that this interaction was driven by the smaller ERN to errors in the non-switchers ($p < 0.05$).

DISCUSSION

WISCONSIN CARD SORTING TASK

While both groups had an overall good performance, switchers were slower, and committed more errors with most errors occurring for subblock trials. This suggests that switchers have

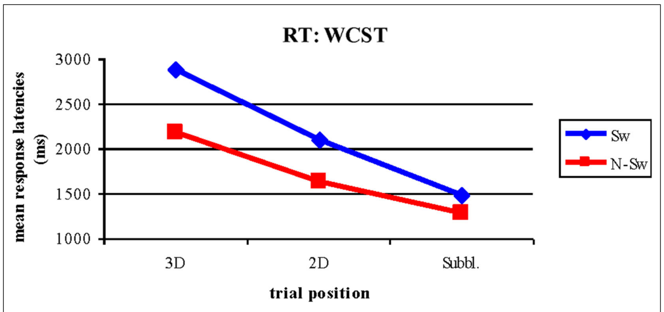


FIGURE 2 | Response latencies: non-switchers were faster for all trial types, i.e., on shift (3D and 2D) as well as on subblock trials (subbl.).

Table 2 | Wisconsin Card Sorting Test: error rate in percent.

	Switcher	Non-switcher	Statistics
All errors	5.6% (18.4)	2.0% (3.8)	$U = 66.5$; $p < 0.05$
Subblock errors	3.8% (14.1)	1.0% (2.2)	$U = 62.0$; $p < 0.05$

Table 3 | Behavioral data of the Flanker experiment (SDs in brackets).

	Switcher	Non-switcher	Statistics
Correct responses	77% (11.7)	80% (9.5)	$T(27) = -0.775$; $p = 0.225$
Correction rate	65% (11.6)	78% (7.8)	$t(27) = 1.805$; $p < 0.05$
RT congr trials	708 ms (79)	707 ms (85)	$t(27) = -0.020$; $p = 0.984$
RT incongr trials	995 ms (107)	863 ms (118)	$t(27) = -3.104$; $p < 0.01$

[2]http://sdepl.ucsd.edu/erpss/

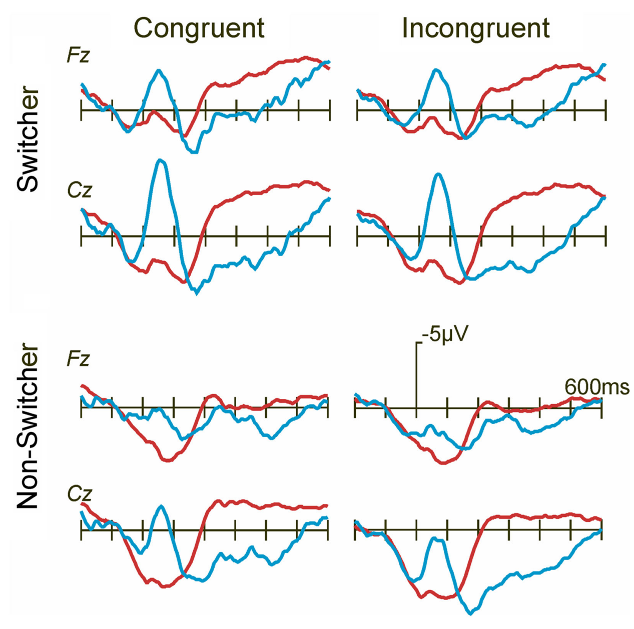


FIGURE 3 | Response-locked grand average ERPs. The red line represents the correct responses, the blue line the errors.

difficulties in maintaining a task set and inadvertently switched to a currently invalid rule.

Bialystok and Martin (2004) and Martin-Rhee and Bialystok (2008) found that bilingual children were better able to shift from one rule (dimension) to the other than monolingual peers on the dimensional change task, which is similar to the WCST. These bilingual children demonstrated an advantage in shifting mental sets, whereas the difference between switchers and non-switchers of the current study emerged mainly during the maintenance of task set which was impaired in the switchers. Vega and Fernandez (2011) recently reported that more balanced bilingual children (with respect to language use) made significantly fewer perseveration errors on the WCST than less balanced children, indicating an advantage in set shifting for the former group. While this suggests a disadvantage of monolingual or less balanced bilingual participants in the inhibition of irrelevant task sets, data from Linck et al. (2008) in L2 Spanish adult learners and proficient Spanish–English and Japanese–English bilinguals revealed that higher language proficiency did not correlate with superior inhibitory abilities as measured by the Simon task.

In contrast to these studies, we measured performance of high proficiency balanced adult bilinguals. The inadvertent rule change in switchers suggests a general problem in this group. While the switcher group has no difficulty in changing from one task set (or one language) to the other, problems emerge in the form of unintentional switches. Thus, we suggest that the switcher group has an increased susceptibility to interference in general.

ERIKSEN FLANKER TASK

Switchers showed a greater susceptibility to interfering information, evidenced in slower response times for incongruent trials, as well as less efficient performance monitoring, evidenced by lower error correction rates.

The ERN is often viewed as an index of the activity of a response monitoring system either in the sense of an error detection mechanism (Falkenstein et al., 1991) or in the sense of a response conflict-monitoring mechanism (Botvinick et al., 2001; Yeung et al., 2004). The latter account proposes that it reflects the degree of conflict between simultaneously activated response tendencies. Following this reasoning, the reduced ERN in the non-switcher group should correspond to less response conflict. This, however, is at odds with the superior correction performance which must be based on a more efficient detection of errors in the non-switchers. Previous research in other cognitive domains has suggested that more efficient processes might be associated with less neural activity (e.g., Hund-Georgiadis and von Cramon, 1999). Thus, one might speculate that the reduced ERN amplitude in non-switchers might reflect more efficient response monitoring mechanisms. This notion, while in line with the behavioral pattern, needs to be substantiated by further research.

GENERAL DISCUSSION

In this paper we asked whether late bilinguals differing in language control also show differences in tasks taxing executive control functions. Such differences would support the notion that language control in bilinguals, in particular with respect to unintentional switching between languages, is related to generic executive control capabilities. Indeed, robust group differences were revealed in the WCST and the flanker task.

Non-switchers demonstrated superior inhibitory control and better set shifting abilities (faster performance, fewer perseveration errors) in the WCST. Switchers had difficulties with continuous rule application, providing evidence for a deficient shielding of the appropriate task set against interfering task sets. This is very similar to their deficient shielding of the appropriate language against interference from the non-target language.

In the flanker task non-switchers revealed superior interference control with faster performance on incongruent trials as well as a higher correction rate following an error. Together with the reduced ERN component in the ERP this is evidence for more efficient conflict and self-monitoring.

Relevant to our study is Friedman and Miyake's (2004) distinction between *Resistance to Distractor Interference* and *Resistance to Proactive Interference.* The first denotes the ability to resist interference of information from the environment irrelevant for the task (such as flanking letters in the flanker task), whereas the second describes the ability to resist intrusions of information which was relevant for a previous task/trial, but is irrelevant on the current task/trial (such as the former sorting rules in the WCST). Interestingly, our non-switcher group showed advantages in both types of resistance to interference.

The contrast between *transient and sustained control processes* (e.g., Wang et al., 2009) has been used to interpret the findings of a bilingual task switching study by Prior and MacWhinney (2010). Transient control processes are relevant for controlling single trials or stimuli, whereas sustained control processes are engaged for a longer period of time to provide state-related activation

(Braver et al., 2003). Prior and MacWhinney observed that a group of bilingual students showed advantages in transient, but not in sustained control processes when compared to monolingual students performing on a shape-and-color-decision task switching paradigm. Bilinguals showed enhanced efficiency in the executive function of shifting between mental sets such as shape decision and color decision. The WCST data of the current study showed that switchers had difficulties in resisting distractor interference, i.e., to keep on using the same rule, most likely explained by their susceptibility to interference in general (Festman et al., 2010). In the framework of Wang et al. (2009) this points to difficulties in sustained control processes. At the same time, switchers were slower and less accurate in choosing the correct rule, implying difficulties with mental shift as well, which, in more general terms, might be interpreted as a difficulty in transient control processes.

Colzato et al. (2008) made a difference between *active inhibition* (in order to exclude particular information from processing) and *reactive inhibition* (in order to exclude particular information after it has been already activated) and reasoned that bilinguals outperform monolinguals by building up and maintaining goal representations more efficiently. Also they seem to map these representations more efficiently onto top-down mechanisms of goal-relevant processes. Whereas Christoffels et al. (2007) did not find the expected asymmetric switch costs for low proficient L2 speakers in German–Dutch bilinguals, the differences between switchers and non-switchers in the current study might well be explained in terms of the active–reactive inhibition distinction. In this framework, switchers might engage in a reactive-inhibition-approach, while non-switchers might rely more on active-inhibition processing. More research on group comparisons of multilinguals and on individual differences with respect to language control are needed to further pinpoint the relationship between executive control and language control in bilingualism.

ACKNOWLEDGMENTS

We are grateful to Antoni Rodriguez-Fornells, Wido Nager, Jan-Ole Schümann, Anne Neubauer, Hanna Rudolph, Daniel Wiswede, Marcus Heldmann, and Toemme Noesselt. The German Research Foundation (DFG) supported this project with a grant to Julia Festman (FE 941/1-1).

REFERENCES

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275.

Barceló, F., Reriáñez, R. T., and Knight, R. T. (2002). Think differently: a brain orienting response to task novelty. *Neuroreport* 13, 1887–1892.

Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.

Braver, T. S., Reynolds, J. R., and Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron* 39, 713–726.

Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., and Gabrieli, J. D. E. (2002). Development of frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron* 33, 301–311.

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 8, 595–604.

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

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

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

Eriksen, B. A., and Eriksen, C. F. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components: II. Error processing in choice reaction tasks. *Electroencephalogr. Clin. Neurophysiol.* 78, 447–455.

Festman, J. (2011). Language control abilities of late bilinguals. *Biling. (Camb. Engl.)* doi: 10.1017/S1366728911000344. [Epub ahead of print].

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

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135.

Garavan, H., Hester, R., Murphy, K., Fassbender, C., and Kelly, C. (2006). Individual differences in the functional neuroanatomy of inhibitory control. *Brain Res.* 1105, 130–142.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. (Camb. Engl.)* 1, 67–81.

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

Hund-Georgiadis, M., and von Cramon, D. Y. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp. Brain Res.* 125, 417–425.

Linck, J. A., Hoshino, N., and Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment. Lex.* 3, 349–374.

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. (Camb. Engl.)* 11, 81–93.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their distribution to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. (Camb. Engl.)* 13, 253–262.

Reiterer, S. M. (2009). "Brain and language talent: a synopsis," in *Language Talent and Brain Activity*, eds G. Dogil and S. M. Reiterer (Berlin: Mouton DeGruyter), 155–192.

Spreen, O., and Strauss, E. (1998). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary.* New York: Oxford University Press.

Swick, D., and Turken, A. D. (2002). Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proc. Natl. Acad. Sci.* 99, 16354–16359.

Vega, C., and Fernandez, M. (2011). Errors on the WCST correlate with language proficiency scores in Spanish-English bilingual children. *Arch. Clin. Neuropsychol.* 26, 158–164.

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

Ye, Z., and Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: evidence from ERPs. *Brain Res.* 1203, 103–115.

Ye, Z., and Zhou, X. (2009). Executive control in language processing. *Neurosci. Biobehav. Rev.* 33, 1168–1177.

Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 April 2011; accepted: 30 March 2012; published online: 18 April 2012.

Citation: Festman J and Münte TF (2012) Cognitive control in Russian–German bilinguals. Front. Psychology ***3****:115. doi: 10.3389/fpsyg.2012.00115*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 10 January 2012
doi: 10.3389/fpsyg.2011.00388

Self-assessment of individual differences in language switching

Antoni Rodriguez-Fornells[1,2,3]*, Ulrike M. Krämer[4,5], Urbano Lorenzo-Seva[6], Julia Festman[7] and Thomas F. Münte[5]

[1] Cognition and Brain Plasticity Group, Bellvitge Biomedical Research Institute, L'Hospitalet de Llobregat, Barcelona, Spain
[2] Department of Basic Psychology, L'Hospitalet de Llobregat, University of Barcelona, Barcelona, Spain
[3] Catalan Institution for Research and Advanced Studies, Barcelona, Spain
[4] Helen Wills Neuroscience Institute, University of California at Berkeley, Berkeley, CA, USA
[5] Department of Neurology, University of Lübeck, Lübeck, Germany
[6] Deparment of Psychology, Universitat Rovira i Virgili, Tarragona, Spain
[7] Institute of Psychology, University of Magdeburg, Magdeburg, Germany

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
Tamar H. Gollan, University of California, USA
Mirjam Broersma, Max Planck Institute for Psycholinguistics, Netherlands

***Correspondence:**
Antoni Rodriguez-Fornells, Department of Basic Psychology, L'Hospitalet de Llobregat, University of Barcelona, Campus Bellvitge, Barcelona 08097, Spain.
e-mail: antoni.rodriguez@icrea.es

Language switching is omnipresent in bilingual individuals. In fact, the ability to switch languages (code switching) is a very fast, efficient, and flexible process that seems to be a fundamental aspect of bilingual language processing. In this study, we aimed to characterize psychometrically self-perceived individual differences in language switching and to create a reliable measure of this behavioral pattern by introducing a bilingual switching questionnaire. As a working hypothesis based on the previous literature about code switching, we decomposed language switching into four constructs: (i) L1 switching tendencies (the tendency to switch to L1; L1-switch); (ii) L2 switching tendencies (L2-switch); (iii) contextual switch, which indexes the frequency of switches usually triggered by a particular situation, topic, or environment; and (iv) unintended switch, which measures the lack of intention and awareness of the language switches. A total of 582 Spanish–Catalan bilingual university students were studied. Twelve items were selected (three for each construct). The correlation matrix was factor-analyzed using minimum rank factor analysis followed by oblique direct oblimin rotation. The overall proportion of common variance explained by the four extracted factors was 0.86. Finally, to assess the external validity of the individual differences scored with the new questionnaire, we evaluated the correlations between these measures and several psychometric (language proficiency) and behavioral measures related to cognitive and attentional control. The present study highlights the importance of evaluating individual differences in language switching using self-assessment instruments when studying the interface between cognitive control and bilingualism.

Keywords: bilingualism, natural language switching, cognitive control, psychometric

INTRODUCTION

Language switching is an omnipresent behavior that characterizes bilingual individuals and communities. This process occurs when bilinguals alternate between two languages while talking to others. In fact, the ability to switch languages or code switch[1] is a very fast, efficient, and flexible process seen in a wide range of bilingual language processing situations. Code switching consists of alternating between or mixing two languages within a single discourse or episode, sentence or constituent, often with no change of interlocutor or topic. Although some authors and theoretical approaches (from foreign language teaching) have previously deemed code switching to be problematic and an index of poor linguistic competence, the current prevailing view is that code switching is a natural and positive aspect of the bilingual individual's linguistic experience and discourse (Zentella, 1997). For example, Poplack (1980), who collected data on natural conversations in Spanish–English Puerto Rican bilinguals in New York, concluded that code switching is a language skill that reflects a high degree of competence and proficiency in bilinguals and that it is more frequent when both languages were learned in early childhood (Miccio et al., 2009). Moreover, code switching may also occur voluntarily and represent a communication strategy for bilinguals to achieve a specific communication goal.

Although a plethora of studies have been devoted to code switching, an important aspect that has been neglected in psycholinguistic, linguistic, and sociolinguistic approaches to this phenomenon is the role of individual differences in language switching (however, see Weinreich, 1953). From a psycholinguistic point of view, it is important to understand what causes natural

[1]In the present manuscript, we did not make any distinction between language mixing/switching and code switching. It is important to mention that the distinction between code switching and language switching, code mixing and borrowings, among others, is controversial. At times, code switching refers to the use of various linguistic units across sentence boundaries, whereas code mixing refers to mixing linguistic units within a sentence (Hatch, 1976; McLaughlin, 1984). However, following other authors (see extended discussion in Bhatia and Ritchie, 1996), we preferred to use these terms indiscriminately to reflect the language switches observed in bilinguals.

language switches, how an utterance is prepared and produced during a natural code switch, how bilinguals understand and comprehend a mixed language input, how bilinguals are aware of the appropriate language to use and how they control code switching tendencies. Code switching tendencies may depend on many elements, including cognitive factors (proficiency in the languages in use, cognitive control functions, cognitive flexibility, general level of cognitive abilities, and personality traits) and other sociopsychological factors (Grosjean, 1982). From a social perspective, any language switch is embedded in a specific social context. Therefore, some authors have proposed that the following factors contribute to language switches: (i) social roles (socioeconomic status, educational background, and relationships between the participants), (ii) situational factors (discourse topic and language suitability in specific contexts), (iii) message-intrinsic considerations (e.g., repetitions, clarifications, emphasis, quotations, and message qualification), and (iv) language attitudes (social dominance, group membership, and security; Ritchie and Bhatia, 2006). In the following paragraphs, we will briefly review the main factors affecting language switching as they relate to linguistic needs (competence, proficiency, and language borrowings) and pragmatic – contextual aspects. Further, we discuss the possible role of interindividual differences in language switching.

LINGUISTIC AND PSYCHOLINGUISTIC FACTORS AFFECTING LANGUAGE SWITCHING

Language switching might be caused by linguistic factors such as proficiency, word semantics, or language similarity (Grosjean, 1982; Poulisse and Bongaerts, 1994; Genesee et al., 2004). Regarding proficiency, language switching frequently occurs due to a lack of knowledge of words in the language being used (L2; Grosjean, 1982). This pattern of switching has been also observed in bilingual children when their linguistic competencies in their less frequently used language are not fully developed and they need to fill their lexical gaps with a word from the other language (Genesee et al., 2004). In addition, in some situations, bilinguals may be faster in accessing the word in their dominant language and tend to produce this utterance instead of the one in the target language. These switching patterns could also be favored in mixed language environments in which both languages are kept highly activated and are further reinforced by the interference that bilinguals experience in language production (Poulisse and Bongaerts, 1994; Colomé, 2001; Rodriguez-Fornells et al., 2005).

Indeed, Gollan and Ferreira (2009) recently reported that balanced Spanish–English bilinguals tend to switch languages more often than unbalanced bilinguals when language switching was measured by the number of times that participants voluntarily switched languages in a naming task. The authors of this study concluded that language switching could be considered to be beneficial in some circumstances and that these switches might be driven by the lexical accessibility of specific words in each language (Poplack, 1980; Clyne, 2003; Owens, 2005). Although it occurs less frequently, the inverse situation (i.e., faster recruitment of L2 than L1 words) may also occur. In some cases, this phenomenon can be observed in the early stages of learning a new language, likely because of the intensive and repetitive practice of novel words. In addition, the lack of use of the L1 due to the immersion in exclusively L2 environments may result in L1 being deactivated (see Grosjean, 1998), resulting in the faster access to L2 words (see for example, Linck et al., 2009).

Semantic factors also have to be taken into account in language switching. For example, there may be words in one language that describe a concept in a very specific way and have no semantically equivalent words in the other language (Bowerman and Choi, 2001; Ameel et al., 2005; Francis, 2005). In particular, abstract or ambiguous types of concepts might not be mapped directly onto corresponding words (Van Hell and De Groot, 1998; Kroll and Tokowicz, 2001; Dong et al., 2005). In this sense, interference from the other language could be considered an accidental code switch, but in due time, it could lead to conscious lexical borrowing (loanword) from the other language (e.g., the words "fax," "spam," and "mouse" in Spanish or "patio" in English).

In addition, the similarity of the two languages spoken in a community could be reflected in the degree of switching (Odlin, 1989; Marian, 2009). For example, Catalan–Spanish languages are highly similar Romance languages, sharing a large number of *cognate words* with similar forms and meanings ["vender" (Spanish) – "vendre" (Catalan), to sell], and switching is commonly observed in many situations (Rodriguez-Fornells et al., 2006; see also Calsamiglia and Tuson, 1984; Woolard, 1988). Interestingly, early observations of naturalistic conversations have suggested that cognate words could act as a language switching trigger. For example, Clyne (1967) observed that code switches frequently occurred close to the use of a cognate word, and he also noticed that "trigger words" existed in the speech of German–English bilinguals that provoked a more or less unconscious switch from one language to the other (see Clyne, 1972). Broersma and De Bot (2006) extended these observations to show that words spoken directly after a cognate word or in the same basic clause were significantly more often code-switched than other words. These results regarding the cognate-related triggering of code switching fit quite well with the idea that both languages are interfering at the production level in bilinguals (Rodriguez-Fornells et al., 2005). If that is the case, the activation of a cognate word in the non-target language may spread activation to associated non-target lexical candidates, and therefore, it would increase the chance of observing a language change. Broersma and De Bot (2006) also found that producing a cognate word is a much more powerful trigger of language switching than hearing a cognate word (Broersma et al., 2009).

SOCIOPRAGMATIC FACTORS AFFECTING LANGUAGE SWITCHING

From a pragmatic point of view, bilinguals choose their language of interaction instantaneously and smoothly, in most of the cases even unconsciously, as a function of whom they are talking to (participants, backgrounds, relationships), what they are talking about (topic, content) and when and where the interaction is taking place. Even small children show language switching abilities and adapt their language to the context and the interlocutor (Petitto et al., 2001). For example, Comeau et al. (2003) showed that six French–English bilingual children (2–4 years old) adjusted their rates of language mixing to the rates of code switching used by the experimenter. Thus, accommodating the language use of the interlocutor appears to be important in bilingual language switching (Petitto et al., 2001; Genesee et al., 2004). If both speakers are able to understand both languages in everyday conversation, code switching might seem natural and acceptable. Code switching is often observed in bilingual families, in particular between siblings being raised abroad from their parent countries. Zentella (1997) observed that bilinguals tend to code switch more in familiar informal settings, implying that extralinguistic social factors such as group identity, age, or gender might modulate language mixing tendencies (Milroy and Gordon, 2003). This code switching pattern observed with friends and family members is very interesting and suggests that it may be easier or more economical in general to mix languages than to keep languages separate. This phenomenon could even imply that some extra effort must be expended to maintain single language production in this type of familiar, informal setting.

Corroborating the role of extralinguistic factors, a particular language is often viewed as more suitable for certain groups, settings, and topics in bilingual societies (Ritchie and Bhatia, 2006). Indeed, many bilinguals use different languages for their public and private "worlds." To this point, Timm (1975) reported that Mexican–American Spanish–English bilingual speakers switched to Spanish to convey personal feelings or to converse about aspects of their culture but switched to English to convey more objective information. The advantage, at the semantic level, of having different ways to express the same ideas permits balanced bilinguals to creatively switch between languages, allowing for special effects in their communication that are not available to monolingual speakers (Zentella, 1997; Auer, 1998). Importantly, such creative use of switching is seen as positive in some societies, favoring conscious, and unconscious switching (e.g., Puerto Rican English–Spanish bilinguals in New York City, English–Hindi mixing in India, and Arabic–French–English mixing in Lebanon; Grosjean, 1982; Poplack, 1985). In contrast, other societies show negative attitudes toward language switching, often due to cultural–historical and linguistic conflicts (e.g., Flemish–French bilinguals in Brussels; Spanish–Catalan bilinguals during the Franco era, Woolard, 1988). Thus, sociolinguistic factors seem to be a prime source of variation in language mixing tendencies in real life and likely will influence and interact with the inherent dynamics of the cognitive systems that support this ability.

INDIVIDUAL DIFFERENCES IN LANGUAGE SWITCHING AND COGNITIVE CONTROL

An additional important factor in language switching is individual differences. Marked interindividual variability has been reported in children raised in bilingual regions (e.g., ranging from 2 to 10% mixing within a single utterance in French–English children from Montreal; see Genesee et al., 2004). At times, speakers show a lack of awareness of the language switch (lack of "metalinguistic awareness").[2] Weinreich (1953) distinguished those who have control over their switches from one language to the other (according to the changes in the interlocutor, topics, and other contextual related switching factors) and those who have difficulty in maintaining or switching codes as required (p. 73). Similarly, Gumperz (1982), in his work on conversational code switching, noted that participants immersed in their conversation are often unaware of which language is used at any one time. He therefore suggested that language selection may be automatic and not readily subject to conscious recall. Similarly, Poulisse and Bongaerts (1994) defined "intentional" and "unintentional" switches in second language production, with the latter characteristically occurring without signs of hesitation or marked intonation. The study by Poulisse and Bongaerts (1994) was motivated by earlier findings by Giesbers (1989; cited in Poulisse and Bongaerts, 1994), who found that unintentional switches can be due to language interference. Similarly, Poplack (1985) distinguished "fluent" switches with smooth transitions and no hesitation from "flagged" switches, which are characterized by repetition, hesitation, into national highlighting and even metalinguistic commentaries.

The existence of unintended (involuntary) switches raises the question of to what degree these switches are related to individual differences in cognitive control and performance monitoring or problems in the control of activation of the non-target language. Explaining these switching patterns and their interindividual variability pose interesting problems for models of bilingual speech production (see review in La Heij, 2005; Kroll et al., 2006). Although this issue has not yet been studied, it is a promising research venue (see Prior and Gollan, 2011; Mas-Herrero et al., submitted; Soveri et al., 2011). Recent evidence indicates that bilinguals show better cognitive control than monolinguals in executive tasks, such as the Simon task (see for a recent review, see Bialystok et al., 2009). This advantage may be due to the more extensive engagement of executive functions since the early stages of language acquisition in infancy. In fact, cognitive control is likely required in bilinguals to switch appropriately between languages, avoiding interference and intrusions from the non-target language (Rodriguez-Fornells et al., 2002a, 2005, 2006; Abutalebi and Green, 2007; Moreno et al., 2008; Ye and Zhou, 2009). For example, Green's Inhibitory Control model (Green, 1986) features an inhibitory mechanism controlling the activation of the dominant language when using the weaker language. Thus, these demands faced by young bilinguals could drastically alter the cognitive control structure of bilingual speakers. However, interindividual differences in executive function may also lead to differences in switching behavior (Festman et al., 2010).

To summarize, there is a need to characterize the individual differences in language switching in multilingual communities through the use of different research approaches, ranging from sociolinguistic and educational to cognitive, psychological, and neuroscientific strategies (for a similar proposal, see Green, 2011).

[2]Metalinguistic awareness is defined in the present study as the lack of knowledge about when it is appropriate to keep both languages separated and when it is appropriate to mix languages; this knowledge depends on the context of the language in use and the history of language socialization. Moreover, an important function of metalinguistic awareness may be to assist the speaker to find and correct an utterance when a mistake occurs, monitoring when communication breakdowns and by analyzing which particular parts of an utterance should be targeted for revision, correction, or improvement (Marshall and Morton, 1978). Notice that this last definition directly involves monitoring and cognitive control processes, which may or may not be language specific.

THE PRESENT PROJECT

The present study sought to develop a reliable self-assessment psychometric instrument to characterize individual differences in language switching, termed the bilingual switching questionnaire (BSWQ). Based on previous results (see above), we hypothesized that language switching could be decomposed into four constructs: (i) first-language (L1) switching tendencies (the tendency to switch from L2 to L1; L1-switch); (ii) second language (L2) switching tendencies (L2-switch), (iii) contextual switch (CS), which assesses the frequency of switches in particular situations or environments; and (iv) unintended switch (US), measuring the lack of awareness of the language switches. The first two factors (L1-switch and L2-switch) were intended to measure switching behavior related to linguistic factors (competence and proficiency in the target and non-target languages and semantic differences across languages; Grosjean, 1982; Poulisse and Bongaerts, 1994; Genesee et al., 2004). The CS construct was designed to measure switching patterns influenced by sociolinguistic factors related to specific situations, people, or topics in which the bilingual speaker usually switches (see above). The last factor, US, aimed to assess unintended language switching not explained by sociolinguistic or linguistic factors (Weinreich, 1953; Gumperz, 1982; Poplack, 1985; Giesbers, 1989; Poulisse and Bongaerts, 1994). In certain situations, however, different factors may simultaneously contribute to eliciting a particular language switch.

Finally, in order to assess the external validity of the individual differences scored with the new questionnaire, we evaluated the correlations between these measures and several psychometric and behavioral measures related first to the language history of the speaker (proficiency, onset of acquisition of each language, and language use) and, second, to cognitive control measures. More specifically, we evaluated the percentage of variance shared between the BSWQ and its factors with (i) quality of inhibitory function (assessed by the stop-signal reaction time (SSRT) in the stop-signal paradigm, Logan, 1995), (ii) stimulus–response interference (using the flanker Eriksen and Eriksen, 1974 task and the Stroop task), and (iii) verbal fluency, which is a well-known measure of executive function.

MATERIALS AND METHODS

PARTICIPANTS AND PROCEDURES

A total of 582 Spanish–Catalan bilingual university students (75.1% women) with a mean age of 21.7 (3.5) years participated after providing their informed consent. The factorial analysis was performed on a final sample of 566 participants (16 participants were discarded due to missing data).

Most of the students were born in Catalonia (surrounding the Barcelona area) or had been raised there since early childhood). All participants had used Spanish and Catalan at home and/or at school during childhood. While newer developments have increased the use of Catalan in the educational system to approximately 90%, the use of Catalan amounted to approximately 50% during the education of the present sample. In addition, exposure to Catalan television and radio programs guarantees enough exposure to Catalan even when Spanish was the only language spoken in the home environment. The results of a bilingual questionnaire of language use and self-assessment of proficiency (Weber-Fox and Neville, 1996; used also in Rodriguez-Fornells et al., 2002a, 2005) revealed a rather balanced use of Spanish and Catalan (mean overall value 3.9 ± 1.6, rated on a seven-point scale with 1 = Catalan only; 2 = Catalan frequently, Spanish rarely; 3 = Catalan majority with Spanish at least 1/4 of the time; 4 = Equal use of Catalan and Spanish; 5 = Spanish majority with Catalan at least 1/4 of the time; 6 = Spanish frequently; Catalan rarely; and 7 = Spanish only). The corresponding ratings for different life stages were infancy, 4.2 ± 2.4; childhood, 3.9 ± 1.8; adolescence, 3.9 ± 1.6; and adulthood, 3.8 ± 1.5. The statistical analysis showed a significant increase in the use of Catalan from infancy to adulthood [F (3,1590) = 22.6, $P < 0.001$].

Regarding the use of Catalan and Spanish in different environments, approximately two-thirds of the participants indicated that they use both languages at the university (63.50%) and in places other than home (66.35%). The remaining third mainly or almost exclusively used one language, with a tendency toward more frequent use of Catalan.

The situation was found to be different at home, where approximately one third spoke solely Catalan (29.36%) and another third solely Spanish (31.38%). In the remaining households, both languages were spoken. In conclusion, in public places, the great majority of participants use both languages, whereas in private environments, one of the two languages is primarily used by most of the participants.

Self-assessment of language skills (4 = perfect, 3 = good, 2 = sufficient, 1 = meager) showed very proficient values, although participants rated themselves significantly better in Spanish than Catalan [overall mean proficiency values: Spanish, 3.87 ± 0.25, Catalan, 3.73 ± 0.46, t (543) = 6.1, $P < 0.001$]. For all language skills tested, better ratings were obtained for Spanish than for Catalan (all $P < 0.001$): *comprehension*, Spanish 3.98 ± 0.14, Catalan 3.91 ± 0.32 [t (543) = 4.5]; *reading*, Spanish 3.97 ± 1.8, Catalan 3.90 ± 0.33 [t (543) = 4.3]; *speaking*, Spanish 3.76 ± 0.49, Catalan 3.60 ± 0.69 [t (543) = 4.4]; and *writing*, Spanish 3.78 ± 0.47, Catalan 3.54 ± 0.76 [t (543) = 6.8].

Finally, participants were asked to estimate the age of their initial exposure to each language. This estimation showed that the mean age of exposure to Spanish was earlier than for Catalan [2.7 ± 1.7 versus 4.1 ± 4.6, t (540) = 6.6, $P < 0.001$].

In summary, the present bilingual sample is very well-balanced in terms of language use and shows high proficiency levels in both languages. However, a small advantage is observed in Spanish proficiency, which is likely because exposure to Spanish occurred earlier than to Catalan in this sample. Although previous early studies of Spanish–Catalan code switching in the Barcelona region have shown that language mixing is not frequent (most likely for political reasons, see Calsamiglia and Tuson, 1984; Woolard, 1988), this tendency might have changed recently because of the efforts of the Catalan autonomic government to increase the presence of Catalan in the schools and media after 1975 (i.e., with the end of the dictatorship period).

PSYCHOMETRIC AND BEHAVIORAL MEASURES

Initially, a pool of 27 items was created to measure the following four constructs: *Spanish-Switch* (L1-Switch), *Catalan-Switch* (L2-Switch), *contextual switch* (CS), and US. In the present study, we systematically used *L*1 to refer to the *Spanish* language and *L*2 for the *Catalan* language when discussing the constructed factors and items. Based upon the initial assessment of the items, their psychometric properties and initial exploratory factorial analysis, 12 of these items were ultimately selected (three for each construct). It is important to bear in mind that the main objective was to measure the four constructs mentioned (see final questionnaire in BSWQ Spanish Version in Appendix and its translation in BSWQ English Translation in Appendix). The entire analysis presented here is based on these 12 selected items.

The participants were required to evaluate the degree to which a behavior characterized his/her language switching habits. A five-point scale (1–5) was used, which quantified the frequency of the behavior described: never (1), rarely (2), occasionally (3), frequently (4), or always (5). Notice that the larger values on the index indicate more frequent switching behavior.

OTHER PSYCHOMETRIC AND BEHAVIORAL MEASURES:

Flanker-Stop combined task

We used a modified variant of the Eriksen flanker task (Eriksen and Eriksen, 1974; adapted from Krämer et al., 2007) that required participants to respond to the central arrow in an array of five arrows (right/left hand response for right/left-directed arrow) and included an inhibitory-stop condition (adapted from Marco-Pallares et al., 2008). The four surrounding arrows were either compatible or incompatible with the central arrow, favoring performance errors. We presented 38.5% compatible and 38.5% incompatible trials. In 11.5% of the trials, we included a stop manipulation similar to a typical stop-signal paradigm (Band et al., 2003). In these trials, the central green arrow changed to red after a variable delay, signaling participants to inhibit their responses in these trials. Two different fixed stop-signal delays were employed (with equal probability), one yielding a low inhibitory rate (180 ms) and one yielding a high inhibitory rate (70 ms; Logan, 1995). The remaining 11.5% of the trials were change trials in which the central arrow changed its direction after 50 ms, indicating to the subject that he/she should react with the other hand. Each stimulus array was presented in the middle of the screen. The stimulus duration was 300 ms, and the stimulus onset asynchrony (SOA) was between 900 and 1100 ms (rectangular distribution; see Marco-Pallares et al., 2008).

The participants received several training trials to become acquainted with the task. They were encouraged to correct their errors in the go-trials as quickly as possible. The experiment was divided into three blocks, each composed of 208 trials, resulting in a total of 624 trials. We were able to extract several measures from this task that reflect inhibitory function, stimulus–response interference, and performance monitoring: the effect of incongruency on reaction time (the reaction time for correct responses in incompatible trials minus compatible trials), percentage of errors (errors in incompatible trials minus compatible trials), percentage of inhibited trials, and SSRT (see Band et al., 2003; we used the easy stop trials for the computation of the SSRT and the percentage of correctly changed trials). For the calculation of the SSRT, the reaction times of the correct trials during which a no stop-signal occurred were collapsed into a single distribution. The RTs were rank ordered, and the mean of the fastest N trials was computed, where N is the number of RTs in the distribution (m) multiplied by the probability of responding at a given delay. This nth RT estimates the time at which the stop process finishes relative to the onset of the go-signal trials. For the estimation of the SSRT relative to the stop-signal onset, the stop-signal delay must be subtracted from this nth value (see Rodríguez-Fornells et al., 2002b).

Stroop task

We used a computerized version of the Stroop task (Stroop, 1935) that presented the words “blue,” “green,” and “red” in either a congruent or incongruent color, requiring the participant to press the button that was associated with the color in which the word was written. A total of 120 trials were presented (50% incongruent), with 10 training trials at the beginning. Stimulus duration was 500 ms, and the SOA varied randomly between 1500 and 2500 ms. We computed the effect of incongruency on reaction time (reaction time for correct responses in incongruent trials minus congruent trials) and the percentage of errors (errors in incongruent trials minus congruent trials).

Fluency task

We used a phonological verbal fluency task in Spanish. The participants were required to write down as many words beginning with the letter *F* as possible within 2 min. The dependent variable was the number of written words. Only non-repeated words were scored as correct. Words that contained minor orthographic errors were considered to be correct.

RESULTS

PRELIMINARY ANALYSES

We computed univariate and multivariate descriptive statistics for the 12 items. The univariate descriptive statistics are shown in **Table 1**. The means ranged from 1.79 to 3.31, whereas variance ranged from 0.7 to 1.57. Polychoric correlation is advised when the distributions of ordinal items are asymmetric or show excessive kurtosis. If both indices are lower than one in absolute value, then Pearson correlation is advised (Muthén and Kaplan, 1985,

Table 1 | Univariate descriptive statistics for the items.

Item	Mean	95% Confidence interval	Variance	Skewness	Kurtosis*
1	2.50	2.38–2.54	0.87	0.47	−0.14
2	2.28	2.21–2.35	0.70	0.17	−0.46
3	2.51	2.43–2.60	1.07	0.15	−0.79
4	3.31	3.22–3.40	1.35	−0.37	−0.75
5	3.16	3.07–3.26	1.38	−0.32	−0.84
6	1.79	1.72–1.86	0.77	1.00	0.36
7	2.44	2.33–2.54	1.57	0.46	−0.92
8	1.97	1.89–2.05	0.87	0.81	0.06
9	2.41	2.33–2.50	1.02	0.45	−0.44
10	2.45	2.37–2.53	0.98	0.33	−0.49
11	2.47	2.39–2.56	1.10	0.43	−0.40
12	2.08	2.00–2.16	0.92	0.78	0.25

**Zero centered. Item 7 has been reversed.*

1992). In our data, the skewness and kurtosis indices were in the range of −1 to 1.

The multivariate kurtosis coefficient was 188.129, and the corresponding significance test ($Z = 13.063$; $P < 0.001$) indicated that the multivariate distribution significantly deviated from a normal multivariate distribution. In this situation, a factor analysis method that assumes normal multivariate distribution is not advisable.

We computed the correlation matrix for the 12 items. The Kaiser–Meyer–Olkin values index was 0.716, rendering the correlation matrix suitable for factor analysis.

EXPLORATORY FACTOR ANALYSIS

We used the FACTOR program to compute the exploratory factor (Lorenzo-Seva and Ferrando, 2006). The correlation matrix was factor-analyzed using minimum rank factor analysis (MRFA; Ten Berge and Kiers, 1991) followed by oblique direct oblimin rotation ($\gamma = 0$). The MRFA allows for computation of the proportion of common variance explained by each of the extracted factors. Because the test was developed to measure four dimensions, this was the number of factors that we extracted.

The proportion of common variance explained was 0.23, 0.25, 0.24, and 0.15 for each factor, respectively, and the overall proportion of common variance explained was 0.86. The root mean square of residuals (RMSR) was 0.0474, whereas following Kelly's criterion (1935), the expected mean value of this index for an acceptable model was 0.0421. Finally, the largest positive standardized residual was 2.69. These results allowed us to conclude that the proposed number of factors we wished to retain was, in fact, acceptable.

To assess the factor simplicity (Kaiser, 1974) of the rotated solution, we computed Bentler's (1977) Simplicity Index (S). This index assesses factor simplicity (with a value of 1 indicating maximal factor simplicity). The values for index S (0.89) suggested high factor simplicity. Finally, the simplicity of the factor structure enabled us to identify the four factors as Switch to Spanish (L1S), Switch to Catalan (L2S), contextual switch (CS), and US. The oblique pattern matrix and the corresponding inter-factor correlation matrix are shown in **Table 2**. Items 1, 4, and 9 were related to L1S. Items 2, 5, and 10 were related to L2S. Items 3, 11, and 12 were related to CS. Finally, items 6, 7, and 8 were related to US.[3]

[3]A confirmatory Factor Analysis (CFA) instead of an Exploratory Factor Analysis (EFA) was conducted to test our assumptions about the underlying factorial structure and to assess whether our results and conclusions were due to the methodological approach (LISREL, Jöreskog and Sörbom, 2001). Multiple indices of fit were examined to evaluate the adequacy of the model. The Comparative Fit Index (CFI)

Table 2 | Oblique exploratory factor solution.

Items		L1S	L2S	CS	US	Communality
4	When I cannot recall a word in Catalan, I tend to immediately produce it in Spanish	**0.78**	0.11	0.14	−0.24	0.74
1	I do not remember or I cannot recall some Catalan words when I am speaking in this language	**0.76**	−0.20	−0.15	0.15	0.80
9	Without intending to, I sometimes produce the Spanish word faster when I am speaking in Catalan	**0.73**	−0.04	−0.01	0.23	0.76
5	When I cannot recall a word in Spanish, I tend to immediately produce it in Catalan	0.16	**0.88**	0.18	−0.29	0.92
10	Without intending to, I sometimes produce the Catalan word faster when I am speaking in Spanish	−0.09	**0.75**	−0.13	0.26	0.80
2	I do not remember or I cannot recall some Spanish words when I am speaking in this language	−0.15	**0.68**	−0.05	0.10	0.65
11	There are situations in which I always switch between the two languages	−0.10	−0.08	**0.92**	−0.04	0.80
12	There are certain topics or issues for which I normally switch between the two languages	0.09	0.00	**0.61**	0.09	0.50
3	I tend to switch languages during a conversation (for example, I switch from Spanish to Catalan or vice versa)	0.10	0.12	**0.54**	0.26	0.62
7	When I switch languages, I do it consciously	−0.05	0.00	−0.02	**−0.51**	0.42
8	It is difficult for me to control the language switches I introduce during a conversation (e.g., from Catalan to Spanish)	0.10	0.14	0.28	**0.49**	0.55
6	I do not realize when I switch the language during a conversation (e.g., from Catalan to Spanish) or when I mix the two languages; I often realize it only if I am informed of the switch by another person	0.01	0.14	0.26	**0.49**	0.52
	Proportion of common explained variance	0.23	0.25	0.24	0.15	–
	Inter-factor correlation matrix	**L1S**	**L2S**	**CS**		
	L2S	−0.14				
	CS	0.27	0.33			
	US	0.13	0.20	0.28	–	

L1S, switch to Spanish; L2S, switch to Catalan; CS, contextual switching; US, unintended switching. Loadings larger than the absolute value of 0.40 are printed in bold face.

The inter-factor correlation matrix showed moderate correlations (from 0.14 to 0.33 in absolute values). Only the correlation between L1S and L2S was negative, indicating that a greater tendency to switch language in one direction was associated with a diminished tendency to switch in the opposite direction. In our sample, responders seemed to be slightly more dominant in one or the other language.

DESCRIPTIVE STATISTICS OF SCALES

The scores for each individual ($N = 566$) on scales L1S, L2S, CS, and US were then computed by raw addition of the corresponding item scores. Note that item 7 was conveniently reversed. Because all of the factors were correlated, an overall score could be obtained by the raw addition of all item scores. Even if factors L1S and L2S were slightly negatively correlated, the score in each indicates the degree of switching from one language to the other, so the addition of both scores for an overall score is still meaningful. This overall score was named overall switching (OS).

The descriptive statistics for the scales and the internal consistency (alpha coefficient) of scales L1S, L2S, CS, US, and OS were then computed. The values of these coefficients, printed in **Table 3**, show that the scale reliabilities are acceptable, with the exception of scale US. In reality, because each scale contains a small number of items, high reliabilities were not expected. Factor scores were estimated using Barrett's factor scores, and the corresponding reliabilities are also printed in **Table 3**. Note that three reliabilities (corresponding to factors L1S, L2S, and CS) are larger than 0.80, suggesting that factor scores should be preferred over raw scale scores. In addition, the reliability of factor US is larger than 0.70.

The mean and SD of each raw scale is also printed in **Table 3**. For the overall raw score (OS), the Kolmogorov–Smirnov test ($Z = 1.25$ and $P = 0.087$) indicated that the distribution of scores did not significantly differ from a normal distribution. Thus, the participants' scores were normally distributed. However, it is important to note that because a large sample was used ($N = 566$), differences between the distribution of participants' scores and a normal distribution were not expected to be statistically significant.

Finally, to graphically represent the switching patterns for a participant or to compare samples, in **Figure 1A**, we have plotted the mean values of the present sample on four axes, each representing a specific factor (with scores ranging from 3 to 15 in this particular case). This graphic can be useful for describing different types of samples. In **Figure 1B**, we depict the raw scores of one of the participants.

CORRELATIONS BETWEEN THE BSWQ AND L1/L2 PROFICIENCY AND LANGUAGE USE MEASURES

The pattern of correlations between language proficiency measures, age of acquisition (AOA) of both languages and language use is presented in **Table 4**. For the onset of language acquisition and proficiency, a very congruent pattern emerged. If Spanish (L1) was acquired later, an increase in switches to Catalan is observed. The reverse pattern was found if the age of Catalan acquisition was later. An analogous pattern is observed for proficiency (averaging comprehension, reading, speaking, and writing scores). A greater proficiency in a target language is correlated with fewer switches to the other language.

The greatest correlations are observed for the language use evaluation of Spanish and Catalan across the lifespan (subtraction score). The predominant use of Spanish is correlated with increased switching to L1 and decreased switching to L2.

In summary, the present pattern of correlations clearly reflects the validity of factors L1S and L2S as measures of language switching for linguistic reasons (competence and proficiency). Importantly, notice that the OS measure did not correlate with proficiency measures, AOA or language use (likely because the effects of L2S and L1S cancel each other out).

CORRELATIONS BETWEEN THE BSWQ AND COGNITIVE CONTROL VARIABLES

The pattern of correlations observed between the BSWQ factors and the overall score is illustrated in **Table 5**. It is important to note that a large sample was used in the present study, and thus, despite being significant, the magnitude of some correlations (and variance shared) is very small.

The significant correlation observed between L1S and Fluency indicates that a larger number of switches is expected if fluency in L1 is greater. This result is important because fluency was evaluated in Spanish. Thus, more switches into L1 are expected if the vocabulary is better developed in L1 than in L2. Indeed,

assesses the lack of fit as estimated by the non-central chi-square distribution of a target model compared to a baseline model. The Goodness-of-Fit Index (GFI) is an index of absolute fit that is related to the relative amount of the observed variances and covariances accounted for by the hypothesized model. Hu and Bentler (1999) recommended a cutoff value close to 0.95 for these fit indices. The root mean square error of approximation (RMSEA) is based on the analysis of residuals and compensates for the effects of model complexity. Hu and Bentler (1999) recommended a cutoff close to 0.06. The values obtained for these indices in our study were CFI = 0.99, GFI = 0.99, and RMSEA = 0.045. Thus, we can conclude that the data perfectly fit the hypothesized four-factor model. In addition, our results and conclusions are independent of the methodological approach used (i.e., EFA versus CFA). However, as this is the first analysis to be published with the SWQ tests, we think that is it more coherent to use an EFA approach. If other researchers aim to replicate our results in other samples, then a CFA would be better justified.

Table 3 | Descriptive statistics for scale and factor scores.

Scales	Mean	SD	Reliability	
			Alpha	CI 90%
L1S	7.3	1.7	0.75	(0.72–0.78)
L2S	8.3	2.2	0.74	(0.71–0.77)
CS	6.8	2.3	0.75	(0.72–0.78)
US	7.0	2.2	0.58	(0.52–0.63)
OS	29.3	6.3	0.74	(0.71–0.77)
FACTOR SCORES				
L1S			0.84	(0.82–0.86)
L2S			0.92	(0.88–0.90)
CS			0.84	(0.82–0.86)
US			0.72	(0.69–0.75)

L1S, switch to Spanish; L2S, switch to Catalan; CS, contextual switch; US, unintended switch; OS, overall switch.

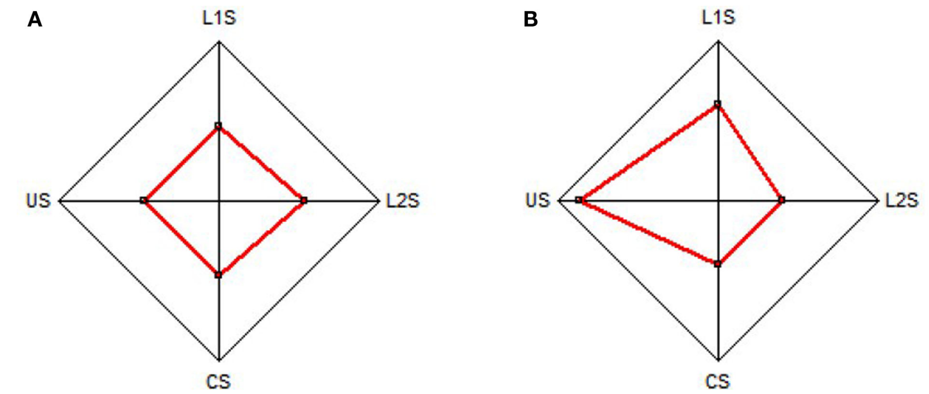


FIGURE 1 | (A) Representation of the mean values observed in the overall Catalan–Spanish sample for the switching tendencies in each factor (larger values represent greater switching). Larger L1–L2 switching was observed, although it was essentially equal for L1 and L2 in the overall sample. Each axis represents a value between 3 and 15. **(B)** The diagram represents an actual participant with strong US switching into L1. Essentially, no switching is observed in the L1 or contextual switch.

Table 4 | Correlations between the BSWQ global scores and proficiency/language use self-assessment scores.

	OS	L1S	L2S	CS	US
Age onset of L1		-0.214^*	0.254^*		
Age onset of L2		0.345^*	-0.286^*		
L1 proficiency		0.207^*	-0.347^*		-0.099^+
L2 proficiency		-0.471^*	0.357^*		
Language use (L2–L1)		0.622^*	-0.570^*		

N = 536. *S, overall switch; L1S, switch to Spanish; L2S, switch to Catalan; CS, contextual switch; US, unintended switch.* $^+P < 0.05$; $*P < 0.001$. *L1 and L2 proficiency: averaging the global scores of rated skills in comprehension, reading, speaking, and writing. Language use: overall language use across different life periods assessed on a seven-point scale; low scores indicate predominance of Catalan (L2) use, and high scores indicate L1 predominance. Empty cells show non-significant effects.*

Table 5 | Correlations between the BSWQ global score and its factors with external variables.

	OS	L1S	L2S	CS	US
FLANKER					
RT incongruence					
STROOP					
RT incongruence			-0.112^+		
% errors incongruence			-0.115^+		
STOP-SIGNAL					
(SSRT)					0.101^+
FLUENCY					
Number of words		0.105^+	-0.157^*		-0.125^*

OS, overall switch; L1S, switch to Spanish; L2S, switch to Catalan; CS, contextual switch; US, unintended switch; Flanker (N = *525): reaction time incongruence effect; Stroop (N* = *491): reaction time incongruence effect and percentage of errors incongruence effect; stop-signal (N* = *491): stop-signal reaction time for the easiest delay (SSRT); fluency (N* = *533): number of words produced that begin with a specific letter;* $^+P < 0.05$; $*P < 0.001$. *Empty cells show non-significant effects.*

Fluency positively correlates with the tendency to use Spanish [$r(529) = 0.14, P < 0.001$]. In summary, a portion of this effect could be due to the dominance of L1 Spanish.

As expected, however, the US factor showed a negative correlation with Fluency, indicating that the production of more words is associated with less switching and most likely better cognitive control. The positive correlation with the stop-signal task (SSRT) indicates that participants with less US tendencies needed less time to inhibit responses on stop trials. Because US may reflect the uncontrolled activation of lexical candidates from the non-target language, its relation to inhibitory abilities is expected in order to increase the suppression of erroneous responses.

A similar interpretation can be made for the correlations between L2S (switching to Catalan) and Stroop interference and L2S and Fluency. Greater L2 switching is associated with lower fluency scores and more interference in the Stroop task (reaction time and erroneous incongruence measures). It is also important to highlight that no differences were observed between the Flanker task measures and any of the BSWQ factors.

Because of the possible differences in the bilingual population, we explored the previous pattern of correlations according to the AOA onset for both languages. First, we excluded from the analysis bilinguals who had learned one of the languages after the age of 10 years to reduce the variability of the sample. It is important to note that in most of the cases, in Catalonia, both languages are acquired during the schooling period due to bilingual educational policies. Second, in the remaining sample, we created three groups of bilinguals according to the AOA for each language: *Catalan–Spanish simultaneous bilinguals* (AOA Spanish ≤ 3 years; AOA Catalan ≤ 3 years), *Catalan–Spanish bilinguals* (AOA Catalan ≤ 3 years; AOA Spanish: ≥ 5 years) and *Spanish–Catalan bilinguals* (AOA Spanish ≤ 3 years, AOA Catalan ≥ 5 years; see **Table 6** for the characterization of the language use variables and proficiency of the sample). As can be seen in **Table 6**, significant

Table 6 | Correlation analysis for the BSWQ and executive tasks according to three different bilingual groups and based on the age of acquisition for L1/L2.

	Simultaneous bilinguals $N = 268$	Catalan–Spanish bilinguals $N = 75$	Spanish–Catalan bilinguals $N = 111$	Overall sample (from Table 4)
PROFICIENCY AND LANGUAGE USE				
Age onset of Spanish	2.1 ± 0.2	5.4 ± 1.1	$2.0 \pm 0.09^{+++}$	
Age onset of Catalan	2.2 ± 0.4	2.0 ± 0.11	$5.5 \pm 1.2^{+++}$	
Spanish proficiency	3.9 ± 0.2	3.7 ± 0.33	$3.9 \pm 0.25^{+++}$	
Catalan proficiency	3.9 ± 0.2	3.9 ± 0.21	$3.7 \pm 0.43^{+++}$	
Language use (1 = Catalan; 7 = Spanish)	3.7 ± 1.3	2.1 ± 0.99	$5.0 \pm 1.3^{+++}$	
BSWQ				
L1S	8.1 ± 2.3	6.4 ± 2.2	$9.0 \pm 2.4^{+++}$	
L2S	7.9 ± 2.2	9.9 ± 2.0	$7.2 \pm 2.5^{+++}$	
CS	7.1 ± 2.6	6.9 ± 2.1	7.1 ± 2.7	
US	6.0 ± 2.3	6.4 ± 2.3	6.5 ± 2.6	
CORRELATIONS BSWQ AND EXECUTIVE TASKS				
Stroop RT incongruence – L2S	−0.09	−0.21	−0.00	-0.11^{+}
Stroop% errors incongruence – L2S	−0.06	−0.05	−0.16	-0.12^{+}
Stop-signal (SSRT) – US	0.08	0.08	0.07	0.10^{+}
Fluency – L1S	−0.06	0.22^{*}	0.14	0.11^{+}
Fluency – L2S	-0.14^{+}	-0.31^{++}	−0.06	-0.16^{+++}
Fluency – US	-0.17^{++}	-0.22^{*}	−0.05	-0.13^{++}

In the proficiency/language use and BSWQ measures, an ANOVA has been conducted with language group as a between-subjects factor (superscript $^{+++}$ indicates significant differences between groups at $P < 0.001$). L1S, switch to Spanish; L2S, switch to Catalan; CS, contextual switch; US, unintended switch; Stroop, reaction time incongruence effect and percentage of errors incongruence effect; stop-signal, stop-signal reaction time for the easiest delay (SSRT); fluency, number of words produced that begin with a specific letter; correlation superscripts: $^{*}P < 0.07$; $^{+}P < 0.05$; $^{++}P < 0.01$; $^{+++}P < 0.001$.

differences were observed for AOA Spanish and Catalan across the three groups of proficiency in Catalan and Spanish and history of language use. Notice, however, that the mean self-assessed proficiency levels are close to the maximum value in all groups (maximum score 4), even in the non-simultaneous groups. Although the groups were created according to the AOA for both languages, very similar groups could be created on the basis of other variables such as proficiency level and language use. Notice that the correlation between these variables is highly significant: AOA Catalan and Catalan proficiency ($r = -0.4$, $P < 0.001$, later exposition to Catalan, decreased proficiency in Catalan) and AOA Catalan and Language use ($r = 0.53$, $P < 0.002$, delayed onset of Catalan exposition, greater use of Spanish).

Moreover, in **Table 6** we can see the group differences in the BSWQ variables. For the L1–L2 switches, large differences were observed across groups. The Catalan–Spanish group tended to switch to Catalan more often, whereas the Spanish dominant group switched more often to Spanish. No differences were observed for the mean number of contextual switches. Instead, the Spanish–Catalan and the Catalan–Spanish groups showed a tendency for larger amount US switches. When the simultaneous bilingual group was compared directly to the Spanish–Catalan group, a marginal statistical trend was observed [$t(372) = -1.74$, $P < 0.08$]. After pooling both non-simultaneous groups (Spanish–Catalan and Catalan–Spanish groups), a significant effect appeared compared to the simultaneous bilingual group [$t(445) = 1.98$, $P < 0.05$]. Thus, the bilingual simultaneous group showed fewer US.

In **Table 6**, we compare the correlations between the selected cognitive control variables highlighted in **Table 5** across the different proficiency groups. First, the most reliable effects across groups were observed for Fluency (number of words) and, specifically, for the relation between fluency L2S and US. However, these effects were significant only in the Catalan–Spanish and simultaneous bilingual groups. No significant effects of fluency were observed for the Spanish–Catalan group. This result is important because it clarifies the correlation that was observed between L2S and Fluency in the full sample. The negative correlation indicates that greater cognitive control (measured by fluency) prevents switches to the L2 in simultaneous and L2-dominant bilinguals. This result is not observed in the Spanish dominant group because the tendency to switch to Catalan is largely reduced.

For the relationship between SSRT and US, the same correlation was observed across the three groups (range, 0.07–0.08), even though they were non-significant due to the reduced sample size. The relationships between the two Stroop measures and the BSWQ seemed less reliable and more variable across the different samples.

DISCUSSION

In the present study, we aimed to psychometrically characterize individual differences in language switching patterns observed in bilinguals. To our knowledge, this is the first attempt to create a self-assessment measure to evaluate individual differences in language switching. Although large differences in language switching have been previously reported between individuals and in bilingual communities, a measure such as the present one helps to systematize these differences. Four factors were validated and assessed using the BSWQ: (i) L1-Switch, which measures the tendency to switch to Spanish (L1); (ii) L2-Switch, which measures the tendency to switch to L2 (Catalan); (iii) contextual switch (CS), which indexes the frequency of switches introduced usually in a particular situation or environment; and (iv) US, which measures the lack of awareness for language switches.

As we expected, the first two factors, L1S and L2S, were associated with switching behavior induced by linguistic needs. The large and robust correlations observed between language use, L1S and L2S reflect the fact that these types of language switches are mainly due to linguistic needs (i.e., to fill a lexical gap in a conversation with a word from the language not in use or to find a better word to convey the message in the other language; Grosjean, 1982; Poulisse and Bongaerts, 1994; Genesee et al., 2004). The correlations of L1S and L2S with proficiency and the onset of language acquisition also clearly point in this direction (see **Table 4**).

The factor CS was intended to reflect switching patterns influenced by external sociolinguistic/pragmatic factors. In contrast to the externally triggered nature of the CS, the factor US measures US that cannot be explained by sociolinguistic or merely linguistic factors (Weinreich, 1953; Gumperz, 1982; Poplack, 1985; Giesbers, 1989; Poulisse and Bongaerts, 1994).

INDIVIDUAL DIFFERENCES IN LANGUAGE SWITCHING AND COGNITIVE CONTROL

The present results provide initial support for a relationship between self-perceived language switching tendencies and cognitive control, thus echoing a theme introduced by Bialystok et al. (2009) on the basis of laboratory experiments. Some small but significant and reliable relationships for the full sample (see **Table 5**) and across bilingual groups (see **Table 6**) were encountered between factor US, the latency of inhibitory processes and verbal fluency scores. This result clearly indicates a link between cognitive control and individual differences in factor US. In addition, when the sample was divided into simultaneous and non-simultaneous bilinguals, fewer US were observed in the simultaneous group; this result once again points to increased cognitive control for bilinguals who are exposed to both languages very early in life.

In a similar fashion, Soveri et al. (2011) used a multiple regression approach in a group of high-fluency Finnish–Swedish bilinguals to study the effects of individual differences in language switching (using the BSWQ) on several executive tasks. Interestingly, the amount of language switching predicted mixing costs in a set-shifting task: greater everyday switching was associated with reduced mixing costs, especially in the number of erroneous responses. Because this executive measure is supposed to reflect the sustained, top-down regulation, and monitoring of alternative or competing task-schemas in order to efficiently react to changes in the task, the authors considered that the relationship could be associated with the long-term effects of language switching on executive function.

However, these cited findings are contradictory to those reported in Prior and MacWhinney (2010), who found a bilingual advantage in switching costs but not in mixing costs in a study with young adults (see also Prior and Gollan, 2011 for similar findings). Moreover, Mas-Herrero et al. (submitted) recently compared a fluent Spanish–Catalan group of bilinguals with a monolingual group in a language switching paradigm. Interestingly, the authors found that increased contextual everyday switching (measured using the BSWQ) predicted reduced switching costs in the language switching paradigm. Thus, this result converges with the previous results from Soveri et al. (2011) and strengthens the relationship between habitual language switching and cognitive control processes (see also Prior and Gollan, 2011).

An important caveat, given that the magnitude of the relationship between language switching and cognitive control was small in the present investigation, is that the present findings should be interpreted with caution. The small magnitude of the effects may explain why recent literature on the relationship between cognitive control and bilingualism has been divergent (Morton and Harper, 2007; Carlson and Meltzoff, 2008; Festman et al., 2010). In a recent study, Hilchey and Klein (2011) reviewed several investigations that compared the performances of monolingual and bilingual groups on non-linguistic interference tasks to evaluate the validity of the claim that bilinguals have advantages in inhibitory control. The authors' conclusions shed serious doubts about the previous findings regarding bilingual advantages in cognitive control, especially for interference tasks (e.g., the flanker or Simon paradigms). Similarly, Gollan et al. (2011) showed a rather small contribution of cognitive control in bilingual language processing. Two groups of younger and older bilinguals were evaluated using cognitive control (the flanker task) and fluency verbal tasks, which specifically focused on scoring cross-language intrusions. Interestingly, only the older bilingual group showed a relationship between cognitive control (as measured by the incongruency error effect in the flanker task) and the number of cross-language errors, as this effect is absent in the younger sample. Although this finding is very interesting because it links declines in cognitive control during aging and cross-language bilingual interference, it also indicates that younger bilinguals likely did not show this effect because they were performing at ceiling levels. This reason may also be why we did not encounter a clear relationship between the flanker incongruency effect and language switching across languages in the present study.

Another important consideration is that in the present study, we used a flanker-stop task with a large amount of trials in each condition and a very fast SOA (900–1100 ms; including the stop task, a total of 624 trials were administered per participant; Krämer et al., 2007). In other studies, a much smaller number of trials was used, emphasizing individual differences during the early stages of performing a new task instead of stable differences in cognitive control after task-practice or habituation to the experimental setup (see Hilchey and Klein, 2011 for a review). As the authors note, the bilingual advantage to conflict resolution in adults and the elderly tends to vanish as a function of the number of trials to which the participants have been exposed (Hilchey and Klein, 2011).

AWARENESS AND ERROR MONITORING IN LANGUAGE SWITCHING

Interestingly, two of the factors identified, Contextual and US, appear to share a common facet, namely, that in many cases, contextual switches might also occur without explicit awareness of the language switch. Although some contextual switches appear to be under the speaker's conscious control (Kroll et al., 2006) and may be driven by pragmatic and social considerations, in other cases, such as when a switch is triggered by a cognate word (Clyne, 1967, 1972; Broersma, 2009), the switch will not be consciously planned. It is also of note that the language selection and on-line adjustments shown by high proficiency bilinguals when interacting with a stranger who has no knowledge of his/her language proceeds very smoothly and flexibly, and in some cases even unconsciously (Gumperz, 1982; Petitto et al., 2001; Comeau et al., 2003; Genesee et al., 2004). Although the subtle and probably subliminal cues that drive this intriguing process are far from being understood, it is important to distinguish this type of contextually triggered switch from the unintended type that is captured by the US factor. The latter refers to unintended and inappropriate switches, reflecting a lack of metalinguistic awareness and similarities with accidental speech errors (Weinreich, 1953; Poplack, 1985; Giesbers, 1989; Poulisse and Bongaerts, 1994). These errors might be related to cognitive control abilities and their interactions with language functions rather than to contextual–situational factors. Indeed, this factor was better correlated with the cognitive control variables and was more reliable across groups in the present study (see **Table 6**).

The differences between contextual and US might involve different executive control processes because one process is triggered externally (CS), whereas the other has internal origins (US). Thus, these psychometric factors reflect the important distinction between behaviors guided by internal processes and behaviors stimulated by the environment (see discussion in Rodríguez-Fornells et al., 2002b). Contextual switches might appear because subtle contextual cues from the environment impact the activation of specific language-based schemas that immediately trigger the activation of lexical items in the non-target language, thereby increasing competition (Green, 1986; Norman and Shallice, 1986; Cooper and Shallice, 2006). For example, typically, bilinguals immediately change the language of their conversation if a third person joins the conversation and the speaker knows that the joiner does not speak that language. Although this process seems to proceed automatically and in a very effortless way, further studies in naturalistic environments are needed to investigate its automaticity when triggered by an external or internal cue, the degree of awareness about this process and the resources that are needed (switch costs).

In this regard, the present distinction between CS and US may be similar to the dichotomy of voluntary/involuntary switches recently described by Gollan and Ferreira (2009). These authors expanded on previous work regarding voluntary switching costs (Arrington and Logan, 2004) to demonstrate the differences between voluntary and involuntary language switches under laboratory conditions (see also Yeung, 2010). The involuntary switches were triggered by specific task instructions to switch the language (cued-switches, which are similar to a method used by Meuter and Allport, 1999), while the voluntary switches were spontaneously induced by requesting that the participants answer in the language of their choice, either Spanish or English. Although the involuntary/cued-switches in Gollan and Ferreira (2009) and the CS factor in the present investigation show some resemblance, it is important to bear in mind that cued language switching tasks used in the laboratory are very unnatural and probably do not reflect the complex dynamics of bilingual communication in mixing contexts. On the other hand, US cannot be equated with the voluntary switches described by Gollan and Ferreira (2009). The best characterization of natural language switching for the US factor is probably a lack of explicit intention to switch. Indeed, a language switch could be internally elicited, but it could also be absolutely involuntary (e.g., a problem with language interference in which the non-target lexical candidate is selected). Rather, "voluntary switches" as defined by Gollan and Ferreira (2009) might be similar to the switches related to linguistic needs (i.e., L1S and L2S), which in some cases could be voluntarily driven and triggered because of differences in proficiency levels.

To better understand unintended and involuntary language switches, ecologically valid and natural situations in which switches can be categorized and separately studied need to be created. Kootstra et al., 2009; see also Kootstra et al., 2010) recently introduced an interesting method to investigate the natural language alignment of the interactions of two or three participants engaged in natural conversation. Their method presents a very promising venue in which to study the switching costs associated with different types of language switches in bilinguals. An interesting experiment would be to create interactions between participants with different language switching tendencies and to study the dynamics of the resulting conversations.

When distinguishing Contextual and US within the context of cognitive control models (e.g., Norman and Shallice, 1986; Shallice, 2004), it is important to consider how bilinguals switch languages and why in some cases these switches bypass awareness. The implementation of a monitoring device in speech production and bilingual models has been postulated by different authors (e.g., a language switching on/off mechanism, McNamara and Kushnir, 1971; a monitoring system, Albert and Obler, 1978; Comparator system, Lipski, 1978 and Sridhar and Sridhar, 1980; see for an interesting and recent review, Nozari et al., 2011 and neuroscientific evidences, Möller et al., 2007). The question remains, how can this type of error monitoring system be implemented, and how do bilinguals evaluate the occurrence of Contextual and US? Cognitive control in this particular situation may also depend on the degree of separation or segregation of target and non-target language representations in the brain (for different proposals, see Green, 1998; Grosjean, 1998; De Bot, 2004) and how top-down and local activation and inhibition mechanisms impact the overall activation level of the target and non-target languages in use (see Green, 1986; Li and Farkas, 2002; Rodriguez-Fornells et al., 2006). Thus, additional work in this area is needed to better understand the cognitive control architecture involved in bilingual language switching.

LIMITATIONS OF THE PRESENT STUDY

Finally, it is important to comment on several limitations of the present study, which was exclusively aimed at developing a self-assessment measure of individual differences in language switching patterns. One of the main caveats is that we did not provide an external, independent variable for language switching or code switching behavior (see Gullberg et al., 2009). For example, it would have been important to gather information on real language switching behavior in conversations between bilinguals (e.g., using the procedure described in Kootstra et al., 2009, 2010), the number of voluntary language switches used when naming simple pictures (see Gollan and Ferreira, 2009) or administrating the BSWQ questionnaire to a third person (relatives or friends of the evaluated person) in order to correlate self- and informant-assessments of language switching. Further research in this direction will be needed to validate the present findings.

However, at least three recent studies may provide some validity to the use of self-reported measures of language switching in daily life such as the BSWQ. First, in an interesting study, Prior and Gollan (2011) used a self-reported measure of language switching in two groups of bilinguals living in the United States, Spanish–English and Mandarin–English speakers. The groups indeed differed in the amount of language switching that occurred in their daily conversations, with a greater switching tendency in the Spanish–English group. This pattern was expected, considering that Spanish is a language that is more common and accessible than Mandarin. This finding validates the use of self-report measures to characterize bilingual switching patterns in communities (Ritchie and Bhatia, 2006). Moreover, Prior and Gollan (2011) used a non-linguistic language switching task to provide a direct measure of non-language and language switching and mixing costs. When compared to a monolingual group, only the Spanish–English bilingual group, which reported larger everyday language switching tendencies, showed a reduced switching cost (either for the non-linguistic or the linguistic task). This result is interesting because it is the first one to suggest a specific link between self-reported individual differences on language switching in daily life and a specific advantage to cognitive control, which improved general switching abilities. However, it is important to highlight that in this study, the Spanish–English group was more balanced in terms of proficiency in both languages, making it more difficult to rule out the possible contribution of proficiency in the effects encountered. Indeed, disentangling the effects of proficiency and language switching tendencies will be difficult because in some populations, these measures are correlated (Gollan and Ferreira, 2009).

Second, as noted above, Soveri et al. (2011) also found that everyday language switching (using the BSWQ) predicted mixing costs in a set-shifting task. Third, in a recent article, Festman et al. (2010) divided a sample of Russian–German bilinguals into language switchers and non-switchers based on the amount of switches observed in a bilingual picture naming task. Interestingly, those who were classified as switchers using this task also showed more involuntary language switches in a simple verbal fluency task. This task was conducted under more natural circumstances and used a bilingual interview in which the target language was altered every 5 min. Importantly, the switch group obtained worse scores on several neuropsychological tests of executive function. These results also indicate that natural individual differences in language switching can be observed and can be related to a reduced cognitive control that likely diminishes the ability of certain bilinguals to prevent cross-language interference. Because the language switches in that study were in most cases involuntary, it may be interesting to relate these findings to individual differences in the US factor identified in the BSWQ. This result is consistent with one of the reliable findings in the present study, the significant correlation observed between cognitive control measures and the US factor. Better cognitive control (based on Fluency or SSRT) was associated with a diminished number of US (see **Table 6**) in the simultaneous and Catalan-dominant bilinguals.

Future studies will be needed on this topic to better characterize individual differences in code switching across different groups of bilinguals (Green, 2011) and to determine which factors in bilingualism are critical for explaining the long-term effects in cognitive control that are observed in some bilingual groups (Bialystok et al., 2009).

ACKNOWLEDGMENTS

We would like to thank E. Camara, T. Cunillera, D. Cucurell, and J. Marco-Pallares for their help during the acquisition of the present data, which were gathered in the context of a large neurogenetic study funded by a grant from the Volkswagen-Foundation (awarded to Thomas F. Münte and Antoni Rodriguez-Fornells). In addition, Antoni Rodriguez-Fornells has been supported by a grant from the Spanish (PSI2008–03901) and the Catalan Governments (SGR 2009 SGR 93). Ulrike M. Krämer and Thomas F. Münte have been supported by the Deutsche Forschungsgemeinschaft (DFG). Thomas F. Münte is also supported by German Federal Ministry of Science (BMBF). The present project benefited from the discussions and support from E. Mas-Herrero, R. de Diego Balaguer, and M. Laine.

REFERENCES

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

Albert, M., and Obler, L. (1978). *The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism*. London: Academic Press.

Ameel, E., Storms, G., Malt, B. C., and Sloman, S. A. (2005). How bilinguals solve the naming problem. *J. Mem. Lang.* 53, 60–80

Arrington, C. M., and Logan, G. D. (2004). The cost of a voluntary task switch. *Psychol. Sci.* 15, 610–615.

Auer, P. (1998). *Code-Switching in Conversation: Language, Interaction and Identity*. New York: Routledge.

Band, G. P., van der Molen, M. W., and Logan, G. D. (2003). Horse-race model simulations of the stop-signal procedure. *Acta Psychol. (Amst.)* 112, 105–142.

Bentler, P. M. (1977). Factor simplicity index and transformations. *Psychometrika* 59, 567–579.

Bhatia, T. K., and Ritchie, W. C. (1996). "Bilingual language mixing, universal grammar, and second language acquisition," in *Handbook of Second Language Acquisition*, eds W. C. Ritchie and T. K. Bhatia (San Diego: Academic Press), 627–688.

Bialystok, E., Craik, F. I. M., Green, D., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Public Interest* 10, 89–129.

Bowerman, M., and Choi, S. (2001). "Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories," in *Language Acquisition and Conceptual Development*, eds M. Bowerman and S. C. Levinson (Cambridge: Cambridge University Press), 475–511.

Broersma, M. (2009). Triggered code switching between cognate languages. *Biling. (Camb. Engl.)* 12, 447–462.

Broersma, M., and De Bot, K. (2006). Triggered code switching: a corpus-based evaluation of the original triggering hypothesis and a new alternative. *Biling. (Camb. Engl.)* 9, 1–13.

Broersma, M., Isurin, L., Bultena, S., and De Bot, K. (2009). "Triggered code switching: evidence from Dutch–English and Russian–English bilinguals," in *Multidisciplinary Approaches to Code Switching*, eds L. Isurin, D. Winford, and K. De Bot (Amsterdam: John Benjamins), 103–128.

Calsamiglia, H., and Tuson, A. (1984). Use of languages and code switching in groups of youths in a Barri of Barcelona. *Int. J. Soc. Lang.* 47, 105–121.

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298.

Clyne, M. (1967). *Transference and Trigerring: Observations on the Language Assimilation of Post-War German-Speaking Migrants in Australia.* The Hague: Martinus Nijhoff.

Clyne, M. (1972). *Perspectives on Language Contact.* Melbourne: The Hawthorn Press.

Clyne, M. (2003). *Dynamics of Language Contact.* Cambridge: Cambridge University Press.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *J. Mem. Lang.* 45, 721–736.

Comeau, L., Genesee, F., and Lapaquette, L. (2003). The modeling hypothesis and child bilingual codemixing. *Int. J. Biling.* 7, 113–126.

Cooper, R. P., and Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychol. Rev.* 113, 887–916.

De Bot, K. (2004). The multilingual lexicon: modeling selection and control. *Int. J. Multiling.* 1, 17–32.

Dong, Y., Gui, S. S., and MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Biling. (Camb. Engl.)* 8, 221–238.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept. Psychophys.* 16, 143–149.

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

Francis, W. S. (2005). "Bilingual semantic and conceptual representation," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (New York: Oxford University Press), 251–267.

Genesee, F., Paradis, J., and Crago, M. B. (2004). *Dual Language Development and Disorders. A Handbook on Bilingualism and Second Language Learning.* Baltimore: Paul H. Brookes Publishing.

Giesbers, H. (1989). *Code-Switching Tussen Dialect en Standaardtaal.* Amsterdam: P.J. Meertens-Instituut.

Gollan, T. H., and Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 640–665.

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

Green, D. (1998). Schemas, tags and inhibition. *Biling. (Camb. Engl.)* 1, 100–104.

Green, D. W. (1986). Control, activation, and resource: a framework and a model for the control of speech in bilinguals. *Brain Lang.* 27, 210–223.

Green, D. W. (2011). Language control in different contexts: the behavioral ecology of bilingual speakers. *Front. Psychol.* 2:103. doi:10.3389/fpsyg.2011.00103

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

Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Biling. (Camb. Engl.)* 1, 131–149.

Gullberg, M., Indefrey, P., and Muysken, P. (2009). "Research techniques for the study of code-switching," in *The Cambridge Handbook of Linguistic Code-Switching*, eds B. E. Bullock and A. J. Toribio (Cambridge: Cambridge University Press), 21–39.

Gumperz, J. J. (1982). *Discourse Strategies.* Cambridge: Cambridge University Press.

Hatch, E. (1976). "Studies in language switching and mixing," in *Language and Man: Anthropological Issues,* eds W. C. McCormack and S. A. Wurm (The Hague: Mouton), 201–214.

Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658.

Hu, L., and Bentler, P. M. (1999). Cutoff criteria for the fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Modeling* 6, 1–55.

Jöreskog, K., and Sörbom, D. (2001). *Lisrel 8: Structural Equation Modeling with the SIMPLIS Command Language.* Hillsdale, NJ: Lawence Erlbaum.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika* 39, 31–36.

Kootstra, G. J., Van Hell, J. G., and Dijkstra, T. (2009). "Two speakers, one dialogue: an interactive alignment perspective on code-switching in bilingual speakers," in *Multidisciplinary Approaches to Code Switching*, eds L. Isurin, D. Winford, and K. De Bot (Amsterdam: John Benjamins), 129–160.

Kootstra, G. J., Van Hell, J. G., and Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: evidence from bilingual monologue and dialogue. *J. Mem. Lang.* 63, 210–231.

Krämer, U., Cunillera, T., Càmara, E., Marco-Pallarés, E., Cucurell, D., Nager, W., Bauer, P., Schüle-Freyer, R., Schöls, L., Rodriguez-Fornells, A., and Münte, T. F. (2007). The impact of COMT and DRD4 genotypes on neurophysiological markers of performance monitoring. *J. Neurosci.* 27, 14190–14198.

Kroll, J. F., Bobb, S. C., and Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Biling. (Camb. Engl.)* 9, 119–135.

Kroll, J. F., and Tokowicz, N. (2001). "The development of conceptual representation for words in a second language," in *One Mind, Two Languages: Bilingual Language Processing*, ed. J. L. Nicol (Malden, MA: Blackwell), 49–71.

La Heij, W. (2005). "Selection processes in monolingual and bilingual lexical access," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York: Oxford University Press), 289–307.

Li, P., and Farkas, I. (2002). "A self-organizing connectionist model of bilingual processing," in *Bilingual Sentence Processing*, eds R. Heredia and J. Altarriba (North-Holland: Elsevier Science Publishers), 59–85.

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language. evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515.

Lipski, J. (1978). "Code-switching and the problem of bilingual competence," in *Aspects of Bilingualism*, ed. M. Paradis (Columbia: Hornbeam Press), 250–264.

Logan, G. D. (1995). "On the ability to inhibit thought and action: a users guide to the stop signal paradigm," in *Inhibitory Processes in Attention, Memory, and Language*, eds D. Dagenbach and T. H. Carr (San Diego: Academic Press), 189–239.

Lorenzo-Seva, U., and Ferrando, P. J. (2006). FACTOR: a computer program to fit the exploratory factor analysis model. *Behav. Res. Methods Instrum. Comput.* 38, 88–91.

Marco-Pallares, J., Camara, E., Münte, T. F., and Rodriguez-Fornells, A. (2008). Neural mechanisms underlying adaptive actions after slips. *J. Cogn. Neurosci.* 20, 1595–1610.

Marian, V. (2009). "Language interaction as a window into bilingual cognitive architecture," in *Multidisciplinary Approaches to Code Switching*, eds L. Isurin, D. Winford, and K. De Bot (Amsterdam: John Benjamins), 161–185.

Marshall, J. C., and Morton, J. (1978). "On the mechanics of EMMA," in *The Child's Conception of Language*, eds A. Sinclair, R. J. Jarvella, and W. J. M. Levelt (New York: Springer-Verlag), 225–239.

McLaughlin, B. (1984). *Second-Language Acquisition in Childhood, Vol. 1, Preschool Children.* Hillsdale, NJ: Lawrence Erlbaum.

McNamara, J., and Kushnir, S. (1971). Linguistic independence of bilinguals: the input switch. *J. Verbal Learn. Verbal Behav.* 10, 480–487.

Meuter, R. F., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40.

Miccio, A. W., Scheffner-Hammer, C., and Rodriguez, B. (2009). "Code-switching and language disorders in bilingual children," in *Cambridge Handbook of Linguistic Code-Switching*, eds B. E. Bullock, A. Toribio, and J. Almeida (Cambridge: Cambridge University Press), 259–270.

Milroy, L., and Gordon, M. J. (2003). *Sociolinguistics: Method and Interpretation.* Malden, MA: Blackwell.

Möller, J., Jansma, B. M., Rodriguez-Fornells, A., and Münte, T. F. (2007). What the brain does before the tongue slips. *Cereb. Cortex* 17, 1173–1178.

Moreno, E. M., Rodriguez-Fornells, A., and Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *J. Neurolinguistics* 21, 477–508.

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726.

Muthén, B., and Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *Br. J. Math. Stat. Psychol.* 38, 171–189.

Muthén, B., and Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: a note on the size of the model. *Br. J. Math. Stat. Psychol.* 45, 19–30.

Norman, D. A., and Shallice, T. (1986). "Attention to action: willed and automatic control of behavior," in *Consciousness and Self-Regulation*, Vol. 4, eds R. J. Davison, G. E. Schwartz, and D. Shapiro (New York: Plenum), 1–18.

Nozari, N., Dell, G. S., and Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cogn. Psychol.* 63, 1–33.

Odlin, T. (1989). *Language Transfer: Crosslinguistic Influence in Language Learning.* Cambridge: Cambridge University Press.

Owens, J. (2005). Bare forms and lexical insertions in code switching: a processing-based account. *Biling. (Camb. Engl.)* 8, 23–38.

Petitto, K. M., Levy, B. G., Gauna, K., Tetreault, K., and Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition. *J. Child Lang.* 28, 453–496.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en Español: toward a typology of code-switching. *Linguistics* 18, 581–618.

Poplack, S. (1985). "Contrasting patterns of code-switching in two communities," in *Methods V: Papers from the V International Conference on Methods in Dialectology*, ed. H. Warkentyne (Victoria, BC: University of Victoria Press), 363–386.

Poulisse, N., and Bongaerts, T. (1994). First language use in second language production. *Appl. Linguist.* 15, 36–57.

Prior, A., and Gollan, T. H. (2011). Good language-switchers are good task-switchers: evidence from Spanish–English and Mandarin–English bilinguals. *J. Int. Neuropsychol. Soc.* 17, 1–10.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. (Camb. Engl.)* 13, 253–262.

Ritchie, W. C., and Bhatia, T. (2006). "Social and psychological factors in language mixing," in *The Handbook of Bilingualism*, eds T. K. Bhatia and W. C. Ritchie (Oxford: Blackwell Publishing), 336–352.

Rodriguez-Fornells, A., de Diego Balaguer, R., and Münte, T. F. (2006). Executive functions in bilingual language processing. *Lang. Learn.* 56, 133–190.

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

Rodríguez-Fornells, A., Lorenzo-Seva, U., and Andrés-Pueyo, A. (2002b). Are high-impulsive and high risk-taking people more motor disinhibited in the presence of incentive? *Pers. Individ. Dif.* 32, 661–683.

Rodriguez-Fornells, A., Van der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., and Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: brain potential and functional imaging evidence. *J. Cogn. Neurosci.* 17, 422–433.

Shallice, T. (2004). "The fractionation of supervisory control," in *The Cognitive Neurosciences*, 3rd Edn, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 943–956.

Soveri, A., Rodriguez-Fornells, A., and Laine, M. (2011). Is there a relationship between bilingual language switching and executive functions? Introducing a within-group analysis approach. *Front. Psychol.* 2:183. doi: 10.3389/fpsyg.2011.00183

Sridhar, S., and Sridhar, K. (1980). The syntax and psycholinguistics of bilingual code mixing. *Can. J. Psychol.* 34, 407–416.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Ten Berge, J. M. F., and Kiers, H. A. L. (1991). A numerical approach to the approximate and the exact minimum rank of a covariance matrix. *Psychometrika* 56, 309–315.

Timm, L. A. (1975). Spanish-English code-switching: El Porqué and how-not-to. *Roman. Philol.* 28. 473–482.

Van Hell, J. G., and De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: effects of concreteness and cognate status in word association. *Biling. (Camb. Engl.)* 3, 193–211.

Weber-Fox, C., and Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *J. Cogn. Neurosci.* 8, 231–256.

Weinreich, U. (1953). *Languages in Contact.* The Hague: Mouton.

Woolard, K. A. (1988). "Code switching and comedy in catalonia," in *Code Switching. Anthropological and Sociolinguistic Perspectives*, ed. M. Heller (New York: Mouton de Gruyter), 53–66.

Ye, Z., and Zhou, X. (2009). Executive control in language processing. *Neurosci. Biobehav. Rev.* 33, 1168–1177.

Yeung, N. (2010). Bottom-up influences on voluntary task switching: the elusive homunculus escapes. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 348–362.

Zentella, A. (1997). *Growing Up Bilingual.* Malden, MA: Blackwell.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 April 2011; accepted: 09 December 2011; published online: 10 January 2012.

*Citation: Rodriguez-Fornells A, Krämer UM, Lorenzo-Seva U, Festman J and Münte TF (2012) Self-assessment of individual differences in language switching. Front. Psychology **2**:388. doi: 10.3389/fpsyg.2011.00388*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDIX

BSWQ SPANISH VERSION

Trate de contestar en que medida las siguientes preguntas representan o se ajustan a su forma de hablar y expresarte en los idiomas que conoce (p. ej., Catalán–Español), en términos generales. Muchas de estas preguntas hacen referencia a si usted cambia o mezcla frecuentemente el catalán y el castellano en sus conversaciones. Cambiar o mezclar lenguajes es una característica muy particular de algunos entornos bilingües, como es el caso en Cataluña. El siguiente cuestionario pretende investigar sobre dichos hábitos de cambio y mezcla de lenguas. Si tiene dudas sobre algunas respuestas, intente comparar su forma de hablar y expresarte con el de la mayoría, o de las personas que conoce bien.

1. Me faltan o no recuerdo algunas palabras en CATALÁN cuando estoy hablando en dicho idioma.
   ☐ nunca ☐ muy raramente ☐ ocasionalmente ☐ frecuentemente ☐ siempre
2. Me faltan o no recuerdo algunas palabras en ESPAÑOL cuando estoy hablando en dicho idioma.
3. Tiendo a mezclar idiomas durante una conversación (por ejemplo, cambio de español a catalán o a la inversa).
4. Cuando no me sale una palabra en CATALÁN, tiendo a producirla inmediatamente en ESPAÑOL.
5. Cuando no me sale una palabra en ESPAÑOL, tiendo a producirla inmediatamente en CATALÁN.
6. Cuando cambio de idioma (p. ej., de catalán a español) o los mezclo, no me doy cuenta de que lo estoy haciendo y suelen ser los otros los que me lo dicen.
7. Cuando mezclo un idioma lo hago conscientemente.
8. Me resulta difícil controlar los cambios de idioma que introduzco (p. ej., de catalán a castellano) a lo largo de una conversación.
9. Sin quererlo, a veces me sale primero la palabra en ESPAÑOL cuando estoy hablando en CATALÁN.
10. Sin quererlo, a veces me sale primero la palabra en CATALÁN cuando estoy hablando en español.
11. Hay situaciones en las cuales siempre mezclo dos idiomas.
12. Hay asuntos o temas sobre los cuales suelo hablar mezclando ambos idiomas.

POR FAVOR, COMPRUEBE SI HA RESPONDIDO A TODAS LAS PREGUNTAS

BSWQ ENGLISH TRANSLATION

Please, try to answer to what degree the following questions are representative of the manner you use to talk or speak in the language you know (e.g., Catalan–Spanish). Many of these questions ask you to report your tendency to switch or mix languages during a conversation. Switching and mixing languages is a characteristic of some bilingual contexts or environments, as for example in Catalonia. The present questionnaire aims to identify the language switching patterns that exist in these languages. If you have doubts about how to rate yourself in the following questions, please try to compare your manner of speaking and talking with that of most people, or those who you know very well.

1. I do not remember or I cannot recall some Catalan words when I am speaking in this language.
   ☐ never ☐ very infrequently ☐ occasionally ☐ frequently ☐ always
2. I do not remember or I cannot recall some Spanish words when I am speaking in this language.
3. I tend to switch languages during a conversation (for example, I switch from Spanish to Catalan or vice versa).
4. When I cannot recall a word in Catalan, I tend to immediately produce it in Spanish.
5. When I cannot recall a word in Spanish, I tend to immediately produce it in Catalan.
6. I do not realize when I switch the language during a conversation (e.g., from Catalan to Spanish) or when I mix the two languages; I often realize it only if I am informed of the switch by another person.
7. When I switch languages, I do it consciously.
8. It is difficult for me to control the language switches I introduce during a conversation (e.g., from Catalan to Spanish).
9. Without intending to, I sometimes produce the Spanish word faster when I am speaking in Catalan.
10. Without intending to, I sometimes produce the Catalan word faster when I am speaking in Spanish.
11. There are situations in which I always switch between the two languages.
12. There are certain topics or issues for which I normally switch between the two languages.

PLEASE, CHECK IF YOU HAVE ANSWERED ALL THE QUESTIONS

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 01 August 2011
doi: 10.3389/fpsyg.2011.00183

Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach

Anna Soveri*[1]*, Antoni Rodriguez-Fornells*[2,3,4] *and Matti Laine*[1]

[1] *Department of Psychology and Logopedics, Abo Akademi University, Turku, Finland*
[2] *Cognition and Brain Plasticity Group, Bellvitge Biomedical Research Institute, L'Hospitalet de Llobregat, Spain*
[3] *Department of Basic Psychology, Campus Bellvitge, University of Barcelona, L'Hospitalet de Llobregat, Spain*
[4] *Catalan Institution for Research and Advanced Studies, Barcelona, Spain*

Edited by:
Ingrid Christoffels, University of Leiden, Netherlands

Reviewed by:
Mireia Hernandez, University Pompeu Fabra, Spain
Lorenza S. Colzato, Leiden University, Netherlands
Zofia Wodniecka, Jagiellonian University, Poland

****Correspondence:***
Anna Soveri, Department of Psychology and Logopedics, Abo Akademi University, 20500 Turku, Finland.
e-mail: anna.soveri@abo.fi

Several studies have suggested a bilingual advantage in executive functions, presumably due to bilinguals' massive practice with language switching that requires executive resources, but the results are still somewhat controversial. Previous studies are also plagued by the inherent limitations of a natural groups design where the participant groups are bound to differ in many ways in addition to the variable used to classify them. In an attempt to introduce a complementary analysis approach, we employed multiple regression to study whether the performance of 30- to 75-year-old Finnish–Swedish bilinguals ($N = 38$) on tasks measuring different executive functions (inhibition, updating, and set shifting) could be predicted by the frequency of language switches in everyday life (as measured by a language switching questionnaire), L2 age of acquisition, or by the self-estimated degree of use of both languages in everyday life. Most consistent effects were found for the set shifting task where a higher rate of everyday language switches was related to a smaller mixing cost in errors. Mixing cost is thought to reflect top-down management of competing task sets, thus resembling the bilingual situation where decisions of which language to use has to be made in each conversation. These findings provide additional support to the idea that some executive functions in bilinguals are affected by a lifelong experience in language switching and, perhaps even more importantly, suggest a complementary approach to the study of this issue.

Keywords: bilingualism, executive functions, inhibition, set shifting, updating, language switching

INTRODUCTION

Executive functions is a broad, still somewhat undefined, concept that involves abilities that make independent, purposive, self-serving, and socially responsible behavior possible (Lezak, 1995). In an attempt to categorize the available concepts and measures in a coherent fashion, Miyake and his colleagues investigated the psychometric relationships between tasks that are commonly used to assess executive functions (Miyake et al., 2000; Friedman and Miyake, 2004; Friedman et al., 2006). Their findings suggest the existence of three major, separable executive functions: the "inhibition" of unwanted responses, the "shifting" between tasks and mental sets (also called "flexibility"), and the "updating" (and monitoring of) working memory (WM) representations. Research during the last three decades has suggested that bilingualism can enhance certain executive functions (for a review see, e.g., Bialystok, 2009).

Several studies comparing groups of monolingual vs. bilingual individuals (both children and adults) have shown a bilingual advantage in executive functions, particularly in the ability to inhibit irrelevant information (Bialystok and Majumder, 1998; Bialystok, 1999; Bialystok and Martin, 2004; Bialystok et al., 2004, 2006b, 2008; Carlson and Meltzoff, 2008; Costa et al., 2008; Bialystok and Viswanathan, 2009; Soveri et al., 2011). Bilingual advantages have also been reported in the ability to efficiently process a mix of different types of trials that either require or do not require inhibition of conflicting information (Bialystok et al., 2006b; Bialystok and Viswanathan, 2009; Costa et al., 2009; see also Martin-Rhee and Bialystok, 2008). Moreover, bilinguals have been reported to excel monolinguals in their ability to store information in WM (Bialystok et al., 2004).

The bilingual advantage in executive functions is thought to stem from the fact that managing two languages requires executive resources in the form of selection of the relevant language and inhibition of the language not in use at that moment (Green, 1998; Meuter and Allport, 1999; Rodriguez-Fornells et al., 2006; Abutalebi and Green, 2007; Moreno et al., 2008; Bialystok et al., 2009; Ye and Zhou, 2009). Since bilinguals have a lifelong experience in controlling their two languages, they should have received more practice than monolinguals in processes that engage executive functions. This idea is supported by previous studies suggesting that earlier second language (L2) acquisition, higher levels of language proficiency in both languages, and a more balanced use of both languages may have positive effects on executive performance in bilinguals (e.g., Bialystok et al., 2006a; Carlson and Meltzoff, 2008). Further, Costa et al. (2009) hypothesized that the bilingual advantage in executive functions may be related to the degree to which the bilingual uses both languages in conversations in everyday life. Bilinguals who tend to mix languages throughout the day may receive more practice in monitoring processes (in terms of selecting which language to use) and therefore show better executive performance than bilinguals from diglossic sociolinguistic environments where the languages are held separate. Albeit speculative, these considerations highlight the need to relate specific aspects of everyday bilingual behavior to performance on executive test measures.

The exact mechanisms underlying the bilingual executive advantage are not clear. Costa et al. (2009) suggested that the bilingual advantage in inhibition tasks may be caused by the bilinguals having to inhibit the language not in use at a given moment, while their more efficient processing of a mix of different types of trials may stem from the fact that bilinguals constantly need to keep track of both languages in order to select the appropriate language for the situation (see also Bialystok et al., 2009). Further, Colzato et al. (2008) suggested that the bilingual advantage is related to reactive inhibition, a process caused by facilitation of the relevant information in a conflict resolution situation, and not to active inhibition, a process in which irrelevant information is actively inhibited. Colzato et al. (2008) proposed that the bilingual advantage in executive functions is not a result of constantly inhibiting the irrelevant language, but of better selection of the relevant language from the competing irrelevant language.

Although the possible bilingual advantage in executive functions has been assessed in several studies, the research field has solely relied on quasi-experimental designs where bilinguals are compared to monolinguals. Such designs lack the key component of experimental designs which is the randomization of participants into the different groups. As a consequence, it is hard to rule out the role of possible confounding factors that may covary with the variable of interest, i.e., language background.

The present study was an attempt to introduce a complementary analysis approach to study the bilingual advantage in executive functions and its underlying mechanisms. We employed multiple regression in a sample of bilingual Finnish–Swedish adults to investigate whether interindividual differences in five bilingualism-related background factors (language switching, contextual switches, unintended switches, use of both languages in everyday life, and age of L2 acquisition) would be related to the participants' performance on tasks measuring three executive functions (inhibition, updating, and set shifting; see Miyake et al., 2000). To measure our bilinguals' everyday language switching tendencies, we employed a Bilingual Switching Questionnaire (BSWQ; Rodriguez-Fornells et al., submitted). We hypothesized that if the proposed bilingual executive advantage indeed stems from practice in language control, i.e., selecting the target language and/or inhibiting the non-target language, the frequency of behaviors calling for such cognitive processes should correlate with executive measures.

MATERIALS AND METHODS

PARTICIPANTS

The present study employed 38 (12 men; 26 women) neurologically healthy, right-handed Finnish–Swedish bilinguals between 30 and 75 years of age ($M = 52.84$, SD = 14.96; **Table 1**). On the average, they were quite highly educated ($M = 15.45$ years of education, SD = 4.14)[1]. All participants were early simultaneous bilinguals, i.e., they had learned both languages before the age of 7 (Swedish: $M = 3.08$ years of age, SD = 1.74, Finnish: $M = 2.78$, SD = 1.56) and since then used both languages throughout their lives. To ensure that they had balanced skills in both of their languages, they were asked to grade their language skills in Finnish and Swedish on a scale from 0 to 6, where 0 corresponded to no skills in that particular language and 6 to skills at a native level (**Table 2**). There was no significant difference between their Finnish and Swedish speaking skills, reading skills, writing skills, or speech comprehension skills [in all cases, $t(37) < 1$].

TASKS AND QUESTIONNAIRES

The Simon task

The first measure of inhibition that we employed was the Simon task (Simon and Rudell, 1967). This task has been suggested to tap both reactive and active inhibition (Colzato et al., 2008) and several studies have shown a bilingual advantage on this task (Bialystok et al., 2004, 2008; Martin-Rhee and Bialystok, 2008; but see Morton and Harper, 2007; Namazi and Thordardottir, 2010). In this task, a blue or a red square appeared on either the left or the right side of the screen. The participants were to push the left button each time a blue square appeared and the right button each time a red square appeared, irrespective of which side the square was presented on. On congruent trials, the response button was on the same side as the square and on incongruent trials, the square was on the opposite side of the response button, i.e., the irrelevant spatial information was conflicting with the correct response.

Table 1 | Demographics and scores on the BSWQ subscales.

	M	SD	Range
Age in years	52.8	15.0	30–75
Years of education	15.5	4.1	8–25
Everyday use of both languages in %	36.5	29.7	0–90
Age of L2 acquisition in years	4.0	1.6	1–6
BSWQ: language switching (6–30 pts)	14.1	3.0	8–19
BSWQ: contextual switches (3–15 pts)	7.8	2.7	3–13
BSWQ: unintended switches (3–15 pts)	6.0	2.0	3–10

Table 2 | Summary of the participants' estimations of their language skills.

Language	*M*	SD
FINNISH		
Speaking	5.68	0.47
Reading	5.74	0.60
Writing	5.39	0.72
Speech comprehension	5.82	0.39
SWEDISH		
Speaking	5.71	0.52
Reading	5.74	0.60
Writing	5.32	0.96
Speech comprehension	5.82	0.39

[1]The participants in the present study were partly the same as in Soveri et al. (2011).

The present task version included 100 trials of which half were congruent and half incongruent. The presentation order of the trials was randomized separately for each subject. The trials were divided into four blocks with a 5-s break in-between. Before starting the actual test, every subject received a practice sequence. Each trial began with a fixation cross at the center of the computer screen. The cross remained on the screen for 800 ms after which it vanished and there was a 250-ms blank interval. The blank interval was followed by a square (either red or blue) which remained on the screen for 1000 ms if no response was given. After the square vanished, the screen was blank for 500 ms. The differences in RTs and error rates between the incongruent and congruent trials (the Simon effect) were used as the dependent measures on this task. These variables reflect the extra processing cost of having to inhibit the incompatible spatial location of the stimulus[2].

The Flanker task

The other measure of inhibition that we used was the Flanker task (adapted from Eriksen and Eriksen, 1974). A bilingual advantage has previously been found on a modified version of this task (Costa et al., 2008, 2009). In the present task version, five black arrows were presented in a horizontal line at the center of the screen. The task was to decide in which direction the arrow in the middle was pointing, irrespective of the direction of the other arrows (the flankers). On congruent trials, all the arrows pointed in the same direction and on incongruent trials, the flankers pointed in a different direction than the arrow in the middle.

The present task consisted of 50 congruent trials and 50 incongruent trials. The presentation order of the trials was randomized separately for each subject. The trials were divided into two blocks with a 5-s break in-between. Before starting the actual test, the participant received a practice sequence. Each trial began with a fixation cross at the center of the screen. The cross vanished after 800 ms and five arrows appeared in a horizontal line. The arrows remained on the screen for 800 ms if no response was given. This was followed by a blank interval of 500 ms. The dependent measures on this task were the differences in RTs and error rates between incongruent and congruent trials (the Flanker effect). The difference variables are measures of the extra processing cost caused by inhibiting the conflicting flanker arrows.

The spatial N-back task

Working memory updating was measured by a visuospatial version of the N-back task (adapted from Carlsson et al., 1998). N-back tasks have not been employed in previous bilingual research, but a study by Bialystok et al. (2004) indicated a smaller WM load effect in bilinguals in a modified Simon task. In the N-back task used in the present study, a white square was presented in one of eight possible locations on the screen. The participant was to remember the location of the previous square (1-back) or the one before the previous square (2-back).

The task used in the present study consisted of 80 one-back trials and 80 two-back trials. The trials were divided into two blocks with 80 trials each and with a 15-s break in-between. Each block consisted of four sequences of 20 trials: two sequences with 1-back trials and two sequences with 2-back trials. Each sequence included 6 targets and 14 non-targets. The order of the sequences was 1-back, 2-back, 2-back, 1-back within the first block, and 2-back, 1-back, 1-back, 2-back within the second block. The presentation order of the trials was pseudorandomized. Before the actual task, the participant was requested to complete a practice sequence.

In the 1-back task, the participant pressed one of the two response buttons: the right one each time the square appeared in the same location as the previous square and the left one each time the location was different. On the 2-back task, the participant was asked to press the right button each time the square was in the same location as the square two trials back and the left button if the location was different. In the beginning of each sequence, the number "1" or "2" appeared at the center of the screen. Number "1" indicated a 1-back sequence and number "2" a 2-back sequence. The number remained on the screen for 5000 ms and was then replaced by a fixation cross in the middle of the screen and a square in one of eight possible locations. The square remained on the screen for 100 ms. A new square appeared 3000 ms after the previous square had disappeared, irrespective of whether a response was given or not. The RT and error rate differences between 2-back and 1-back trials (N-back effect) were used as dependent variables, and reflect the cost of managing the increased demands on updating.

The number–letter task

Shifting abilities were assessed with the Number–letter task (adapted from Rogers and Monsell, 1995). This particular task has not been used in previous bilingualism research. In this task, a number–letter combination (e.g., 3A) appeared in one of two squares at the center of the screen. The task was to either determine if the number was even or odd or if the letter was a vowel or a consonant, depending on in which square the number–letter pair appeared. The squares thus served as cues for which task to perform. Each time the number–letter combination was in the upper box, the task was to determine the number and each time it appeared in the lower box, the task was to determine the letter.

The trials were divided into three different blocks with short breaks in-between. The first two blocks, with 32 trials in each, were single-task blocks, in which the number–letter combination was in the same square on all trials and no task switching was required (Block 1: in the upper square; Block 2: in the lower square). The third block was a mixed-tasks block with 32 switching trials and 48 repetition trials (the task was the same as in the previous trial). The 48 repetition trials included 24 trials in which the participant was asked to decide if the number was even or odd, and 24 trials where the participant was to decide if the letter was a vowel or a consonant. The task switching was unpredictable for the subject, as the number–letter combination appeared in the two squares randomly. The left button was to be pressed each time the number was even or the letter was a vowel, and the right button each time the number was odd or the letter was a consonant. Each block was preceded by a practice sequence.

[2]We also calculated the so-called Gratton effect (Gratton et al., 1992) that reflects the effect of the previous trial type (its compatibility with the current trial type) on performance on the current trial. Two measures were calculated: (a) the difference between incongruent to congruent and congruent to congruent trials, and (b) the difference between congruent to incongruent and incongruent to incongruent trials. The multiple regression models were, however, not significant for either of these variables.

Each trial began with a 150-ms blank interval, after which a fixation cross appeared at the center of the screen. After 300 ms, two small boxes appeared above each other at the center of the screen, with a number–letter combination in one of the boxes. The stimuli remained on the screen for 3000 ms if no response was given. There were two dependent measures for both RTs and error rates on this task. The first one was the switching cost that was defined as the performance difference between the repetition trials and switching trials in the mixed-tasks block. This reflects the cost of a temporary change in task sets. The second dependent variable was the mixing cost that was the performance difference on the single-task trials vs. the repetition trials in the mixed-tasks block. This reflects the cost of maintenance of attentional control in a context where two task sets are active.

The Bilingual Switching Questionnaire

All participants completed a Swedish translation of the BSWQ, a survey instrument developed by Rodriguez-Fornells et al. (submitted) for the study of individual differences in natural language switching. The questionnaire included 12 questions representing four subscales: (a) Tendencies to switch from Swedish to Finnish (e.g., "When I do not find a word in Swedish, I immediately tend to produce it in Finnish"), (b) Tendencies to switch from Finnish to Swedish (e.g., "When I do not find a word in Finnish, I immediately tend to produce it in Swedish"), (c) Contextual switches (e.g., "There are situations in which I always switch between languages"), and (d) Unintended switches (e.g., "It is difficult for me to control the language switches I make during a conversation (e.g., from Swedish to Finnish)"). The participants responded on a 5-point scale varying from *never* (1) to *always* (5). The construction and psychometric assessment of the original BSWQ and its four subscales on a large sample of bilingual Spanish–Catalan speakers is described in Rodriguez-Fornells et al. (submitted). Their paper also includes the original questionnaire and its translation in English.

Three measures from the BSWQ were used in the multiple regression analyses: language switching, contextual switches, and unintended switches. The language switching variable was created by adding up the points on the first two subscales (Tendencies to switch from Swedish to Finnish; Tendencies to switch from Finnish to Swedish).

Our hypotheses concerning the measures from the BSWQ were as follows. Regarding the language switching and contextual switches subscales, we predicted that the more a person switches languages in everyday life (a higher score on a subscale), the better the performance (a smaller processing cost) should be on the executive tasks, if the bilingual advantage in executive functioning stems from a lifelong experience in language switching. In contrast, one would not expect to find such a correlation between executive measures and unintended switches, as they may reflect temporary processes that induce lapses of attention.

Other background tests and questionnaires

All participants were asked to give their written informed consent and to fill out the Edinburgh Handedness Inventory (Oldfield, 1971). They also completed a background information sheet probing their date of birth, education, occupation, vision, hearing, possible reading difficulties, possible neurological and psychiatric illnesses, medication, subjective level of alertness, and possible alcohol intake during the 24-h period preceding the testing. The participants were also asked to fill out a questionnaire concerning their language background and language skills. In this questionnaire, the participants were asked about their age of L2 acquisition, the languages they had used in written and spoken form during the last 3 years, and the frequency (in percent) with which they had used each language in everyday life. In order to obtain a measure of the everyday use of both languages, the percentage of the less frequently used language was subtracted from the percentage of the more frequently used language.

STATISTICAL ANALYSES

Multiple linear regression analyses were conducted for the processing cost in RTs and error rates (**Table 3**), separately for each executive task. Two models of predictors were employed. The first included three background factors, namely participant's age, the age of L2 acquisition, and the percentage of the everyday use of both languages. The second group of predictors included three measures from the BSWQ: the BSWQ language switching measure, the BSWQ contextual switches measure, and the BSWQ unintended switches measure. In both models, the predictors were inserted simultaneously to the analyses.

RESULTS

With regard to the processing cost in RTs (**Tables 4 and 5**), the multiple regression model with age, age of L2 acquisition, and everyday use of both languages was significant for the Simon effect, $F(3,36) = 3.14$, $p = 0.038$, and the mixing cost, $F(3,34) = 3.95$, $p = 0.017$, in the Number–letter task, and the model explained 15% (Adjusted $R^2 = 0.151$) of the variance in the Simon effect and 21% (Adjusted $R^2 = 0.207$) of the variance in the mixing cost. There was a significant association between the predictor age of L2 acquisi-

Table 3 | Performance on the executive tasks.

	RT in ms		Errors in %	
	M	SD	*M*	SD
SIMON TASK				
Congruent	512	90	2.7	3.5
Incongruent	557	90	3.4	3.8
Simon effect	45	37	0.7	4.2
FLANKER TASK				
Congruent	501	69	0.6	1.6
Incongruent	563	81	2.5	2.6
Flanker effect	62	31	1.9	2.5
N-BACK TASK				
1-back	816	170	5.5	4.4
2-back	1017	215	15.0	11.5
N-back effect	201	155	9.5	9.8
NUMBER–LETTER TASK				
Single-task trials	668	120	1.9	2.7
Repetition trials	1000	272	3.6	3.9
Switching trials	1325	322	7.4	7.0
Switching cost	325	139	3.8	6.1
Mixing cost	333	209	1.7	4.0

Table 4 | Summary of the multiple regression analyses: background variables as predictors of processing cost in RTs on the executive tasks.

Variable	The Flanker effect	The Simon effect	The N-back effect	The number–letter task	
				Switching cost	Mixing cost
	B	*B*	*B*	*B*	*B*
Constant	78.57**	–16.53	163.55	145.33	–104.56
Age	0.00	0.31	–0.55	3.23	4.13
Age of L2 acquisition	–0.34	8.95*	7.12	6.64	37.63
Everyday use of both languages	–0.43*	0.27	1.04	–0.49	1.93
R^2	0.16	0.22	0.04	0.15	0.28
F	2.13	3.14*	0.46	1.85	3.95*

*Flanker effect and Simon effect N = 36; N-back effect N = 33; Number–letter task N = 35; Age and everyday use of both languages N = 38; Age of L2 acquisition N = 37. * = p < 0.05, ** = p < 0.01.*

Table 5 | Summary of the multiple regression analyses: background variables as predictors of processing cost in errors on the executive tasks.

Variable	The Flanker effect	The Simon effect	The N-back effect	The number–letter task	
				Switching cost	Mixing cost
	B	*B*	*B*	*B*	*B*
Constant	2.05*	–1.24	–12.90*	0.15	2.06
Age	–0.03	0.02	0.33**	0.01	–0.08
Age of L2 acquisition	0.09	0.19	0.10	0.73	1.03*
Everyday use of both languages	0.00	0.00	0.03	0.00	–0.01
R^2	0.10	0.04	0.33	0.04	0.19
F	1.22	0.46	4.89**	0.42	2.42

*Flanker effect and Simon effect N = 36; N-back effect N = 33; Number–letter task N = 35; Age and everyday use of both languages N = 38; Age of L2 acquisition N = 37. * = p < 0.05, ** = p < 0.01.*

tion and the Simon effect as the outcome variable, indicating that younger age of L2 acquisition resulted in a smaller Simon effect in RTs. Furthermore, all three predictors were marginally significant ($p < 0.10$) in predicting the mixing cost, so that younger age, earlier L2 acquisition, and a more balanced use of both languages in everyday life was associated with a smaller mixing cost. The multiple regression model with the three BSWQ predictors was significant for the mixing cost, $F(3,34) = 2.91$, $p = 0.050$, in the Number–letter task. The model explained 14% (Adjusted $R^2 = 0.144$) of the variance. None of the predictors, however, reached significance in this analysis.

The analyses on the processing cost in error rates (**Tables 6 and 7**) indicated that the multiple regression model with age, age of L2 acquisition, and everyday use of both languages was significant for the N-back effect, $F(3,33) = 4.89$, $p = 0.007$, and the model explained 26% (Adjusted $R^2 = 0.261$) of the variance. There was a significant association between the predictor age and the N-back effect as an outcome variable so that younger age resulted in a smaller N-back effect in errors. The results also showed that the multiple regression model with the three BSWQ predictors was significant for the mixing cost, $F(3,34) = 9.24$, $p < 0.001$, and explained 42% (Adjusted $R^2 = 0.421$) of the variance. Language switching was a significant predictor of the mixing cost in this analysis: the more a participant tended to switch from Swedish to Finnish and *vice versa*, the smaller the mixing cost in errors in the Number–letter task was.

DISCUSSION

Given the somewhat controversial earlier results concerning the bilingual advantage in executive functions, we set out to explore this issue with a new, complementary approach where we sought for relationships between bilinguals' everyday language use and the level of their executive skills. In a sample of 38 Finnish–Swedish early bilinguals, we found that the frequency with which our bilinguals switched between languages in their everyday life significantly predicted the mixing cost (error rate) in our set shifting task (Number-letter task). In broad terms, this result provides support for the assumption that the bilingual advantage stems from a lifelong experience in managing two languages that calls for executive resources (e.g., Green, 1998; Meuter and Allport, 1999; Rodriguez-Fornells et al., 2006; Abutalebi and Green, 2007; Colzato et al., 2008; Moreno et al., 2008; Bialystok et al., 2009; Ye and Zhou, 2009). Not surprisingly, we also found that age was significantly associated with both WM updating and the mixing cost in set shifting, so that younger bilinguals showed smaller processing costs. This is in line with the common finding

Table 6 | Summary of the multiple regression analyses: BSWQ variables as predictors of processing cost in RTs on the executive tasks.

Variable	The Flanker effect	The Simon effect	The N-back effect	The number–letter task	
				Switching cost	Mixing cost
	B	*B*	*B*	*B*	*B*
Constant	50.19*	71.70*	246.71	314.61*	812.66**
Language switching	–1.87	–1.78	7.92	6.08	–21.12
Contextual switches	–0.30	0.42	–8.01	–11.15	–10.09
Unintended switches	6.75*	–0.78	–15.99	1.88	–17.30
R^2	0.17	0.02	0.06	0.04	0.22
F	2.33	0.28	0.61	0.43	2.91*

*Flanker effect and Simon effect N = 36; N-back effect N = 33; Number–letter task N = 35; Language switching, Contextual switches, and Unintended switches N = 38. * = p < 0.05, ** = p < 0.01.*

Table 7 | Summary of the multiple regression analyses: BSWQ variables as predictors of processing cost in errors on the executive tasks.

Variable	The Flanker effect	The Simon effect	The N-back effect	The number–letter task	
				Switching cost	Mixing cost
	B	*B*	*B*	*B*	*B*
Constant	2.27*	–0.38	21.05*	–0.63	15.00**
Language switching	–0.19*	–0.63	–0.13	0.48	–0.62**
Contextual switches	0.07	0.14	–1.19	–0.02	–0.28
Unintended switches	0.14	0.09	–0.06	–0.35	–0.42
R^2	0.16	0.03	0.13	0.05	0.47
F	2.17	0.39	1.49	0.53	9.24**

*Flanker effect and Simon effect N = 36; N-back effect N = 33; Number–letter task N = 35; Language switching, Contextual switches, and Unintended switches N = 38. * = p < 0.05, ** = p < 0.01.*

that the efficiency of executive functions decreases in older age (e.g., Kramer et al., 1999; Kray et al., 2004; Zelazo et al., 2004; Takio et al., 2009).

While the present results are preliminary, they serve to highlight the potential of the complementary methodological approach we are introducing here. Previous studies showing enhanced executive functions in bilinguals have exclusively employed quasi-experimental designs (bilinguals vs. monolinguals) and have thus been unable to rule out all possible confounding factors that could contribute to the observed group differences (see, e.g., Morton and Harper, 2007). However, the present multiple regression approach focuses on the bilinguals and is thus not hampered by the unavoidable methodological problems of naturalistic group designs. Nevertheless, one must keep in mind that regression analyses represent a correlational approach and thus cannot prove causality.

In the present study, it was the mixing cost in the set shifting task that showed sensitivity to the bilingual experience. The underlying cognitive mechanisms of the mixing cost have been under debate. Rogers and Monsell (1995) proposed that the performance difference between single-task blocks and mixed-task blocks is due to an increased WM load, as two different task sets need to be maintained in the mixed-task blocks. However, Rubin and Meiran (2005) showed that the mixing cost is related to a top-down management of competing task sets, and not to WM load. The latter interpretation would fit in the present results: a task-decision process taking place in the mixed-tasks block resembles the bilingual situation where a decision of which language to use has to be made in each conversation.

It is not totally clear as to why we found associations between the bilingual language use and the mixing cost but not the switching cost in the set shifting task. It has, however, been suggested that the mixing cost and switching cost engage different cognitive control processes. The mixing cost may set more demands on sustained control processes, reflecting the constant need to keep different task-sets active or to maintain attentional monitoring processes, in order to efficiently react to changes in the task. The switching cost, on the other hand, may be related to transient control mechanisms, such as reconfiguration of goals or the linking of task cues to their appropriate stimulus–response mappings (Braver et al., 2003). The sustained and transient processes have also been suggested to activate different brain regions (Braver et al., 2003). Furthermore, studies have shown that the mixing cost increases at older age, while the switching cost is less affected by age (for a review, see Mayr and Liebscher, 2001). The switching cost has been defined as a measure of task-set reconfiguration (Rogers and Monsell, 1995), interference from the previous task-set (Allport et al., 1994), or a combination of both (Monsell, 2003; for a review, see Kiesel et al., 2010). The present results may thus give some clues as to exactly which aspects of bilingual language use are important for the executive gains: it might be that language selection and keeping both languages active are more important for the bilingual advantage than inhibition of the non-target language. This is in line also with the scanty associations between the predictors and the inhibition tasks (the single significant model explains only 15% of the variation of the Simon effect), although one should note that the Flanker task and the Simon task may not have been demanding enough for stronger relationships to appear. Contrary to the present findings, however, Prior and MacWhinney (2010) found a bilingual advantage in the switching cost, but not the mixing cost, in a study with young adults (see also Garbin et al., 2010).

One should also note that the present results showed an effect of language switching, but not contextual switches, on the mixing cost in the set shifting task. One possible reason for this may be that the questions in the language switching subscale concern language switching in general, i.e., whether the bilingual typically tends to use a word from the non-target language when the correct word in the target language cannot be retrieved quickly enough. It may be that this type of language switching is related to more sustained control processes, similar to the ones that have been suggested to be involved in the mixing cost. The contextual switches, on the other hand, may be more situation-bound, as the subscale includes questions as to whether there are specific situations and topics where the bilingual tends to mix both languages. This subscale does not give information about the frequency of occurrence for these situations in everyday life. Costa et al. (2009) speculates that those bilinguals who mostly use the two languages in different contexts and do not frequently switch between them, may not show an advantage in monitoring processes, as they end up having less practice on language monitoring. The frequency of unintended switches did not predict executive performance either, probably because they reflect temporary processes that cause fluctuations in attentional control.

In summary, the present results provide some evidence that individual differences in bilingualism-related background factors may predict the mixing cost that bilinguals exhibit in a set shifting task. Our study presents a new, complementary methodological approach that will hopefully shed more light on the important issue of the relationships between bilingual experience and executive functions. There is no doubt that both the measurement of the various aspects of bilingual experience and the cognitive mechanisms of the mixing cost need to be clarified further in future studies. Ultimately, longitudinal data is needed to establish causal connections between bilingualism and enhanced cognition.

ACKNOWLEDGMENTS

This study was funded by a grant from the Joint Committee for Nordic Research Councils in the Humanities and the Social Sciences for a Nordic Center of Excellence (NCoE) in Cognitive Control (Coordinator Professor Lars Nyberg, Umeå University, Sweden). Anna Soveri was also supported by grants from Kommerserådet Otto A. Malms Donationsfond, and the Finnish National Doctoral Program of Psychology. Matti Laine was financially supported by the Academy of Finland. We would like to thank Teemu Laine for technical aid and Maria Pörnull for conducting part of the data collection.

REFERENCES

Abutalebi, J., and Green, D. (2007). Bilingual language production: the neurocognition of language representation and control. *J. Neurolinguistics* 20, 242–275.

Allport, A., Styles, E. A., and Hsieh, S. (1994). "Shifting intentional set: exploring the dynamic control of tasks," in *Attention and Performance XV: Conscious and Nonconscious Information Processing*, eds C. Umilta and M. Moscovitch (Hillsdale, NJ: Erlbaum), 421–452.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Dev.* 70, 636–644.

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Biling. Lang. Cogn.* 12, 3–11.

Bialystok, E., Craik, F. I. M., Green, D. W., and Gollan, T. H. (2009). Bilingual minds. *Psychol. Sci. Public Interest* 10, 89–129.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., Craik, F. I. M., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873.

Bialystok, E., Craik, F. I. M., and Ruocco, A. C. (2006a). Dual-modality monitoring in a classification task: the effects of bilingualism and ageing. *Q. J. Exp. Psychol.* 59, 1968–1983.

Bialystok, E., Craik, F. I. M., and Ryan, J. (2006b). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354.

Bialystok, E., and Majumder, S. (1998). The relationship between bilingualism and the development of cognitive processes in problem solving. *Appl. Psycholinguist.* 19, 69–85.

Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339.

Bialystok, E., and Viswanathan, M. (2009). Components of executive control with advantages for bilingual children in two cultures. *Cognition* 112, 494–500.

Braver, T. S., Reynolds, J. R., and Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron* 39, 713–726.

Carlson, S. M., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298.

Carlsson, S., Martinkauppi, S., Rama, P., Salli, E., Kovenoja, A., and Ahonen, HJ. (1998). The distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. *Cereb. Cortex* 8, 743–752.

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., Heij, W. L., and Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 302–312.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149.

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference cognitive functions: a latent variable analysis. *J. Exp. Psychol. Gen.* 133, 101–135.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., and Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychol. Sci.* 17, 172–179.

Garbin, G., Sanjuan, A., Forn, C., Bustamante, J. C., Rodriguez-Pujadas, A., Belloch, V., Hernandez, M., Costa, A., and Ávila, C. (2010). Bridging language and attention: brain basis of the impact of bilingualism on cognitive control. *Neuroimage* 53, 1272–1278.

Gratton, G., Coles, M. G. H., and Donchin, E. (1992). Optimizing the use of information: strategic control of activation of responses. *J. Exp. Psychol. Gen.* 121, 480–506.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., and Koch, I. (2010). Control and interference in task switching – A review. *Psychol. Bull.* 136, 849–874.

Kramer, A. F., Hahn, S., and Gopher, D. (1999). Task coordination and aging: explorations of executive control processes in the task switching paradigm. *Acta Psychol. (Amst.)* 101, 339–378.

Kray, J., Eber, J., and Lindenberger, U. (2004). Age differences in executive functioning across the lifespan:

the role of verbalization in task preparation. *Acta Psychol. (Amst.)* 115, 143–165.

Lezak, M. D. (1995). *Neuropsychological Assessment*. New York: Oxford University Press, Inc.

Martin-Rhee, M. M., and Bialystok, E. (2008). The development of two types of inhibitory control in monolingual and bilingual children. *Biling. Lang. Cogn.* 11, 81–93.

Mayr, U., and Liebscher, T. (2001). Is there an age deficit in the selection of mental sets? *Eur. J. Cogn. Psychol.* 13, 47–69.

Meuter, R. F. I., and Allport, A. (1999). Bilingual language switching in naming: asymmetrical costs of language selection. *J. Mem. Lang.* 40, 25–40.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., and Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cogn. Psychol.* 41, 49–100.

Monsell, S. (2003). Task switching. *Trends Cogn. Sci. (Regul. Ed.)* 7, 134–140.

Moreno, E. M., Rodriguez-Fornells, A., and Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *J. Neurolinguistics* 21, 477–508.

Morton, J. B., and Harper, S. N. (2007). What did Simon say? Revisiting the bilingual advantage. *Dev. Sci.* 10, 719–726.

Namazi, M., and Thordardottir, E. (2010). A working memory, not bilingual advantage, in controlled attention. *Int. J. Biling. Educ. Biling.* 13, 597–616.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburg Inventory. *Neuropsychologia* 9, 97–113.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. Lang. Cogn.* 13, 253–262.

Rodriguez-Fornells, A., de Diego Balaguer, R., and Münte, TF. (2006). Executive functions in bilingual language processing. *Lang. Learn.* 56, 133–190.

Rogers, R. D., and Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *J. Exp. Psychol. Gen.* 124, 207–231.

Rubin, O., and Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1477–1491.

Simon, J. R., and Rudell, A. P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304.

Soveri, A., Laine, M., Hämäläinen, H., and Hugdahl, K. (2011). Bilingual advantage in focused attention: evidence from the forced-attention dichotic listening paradigm. *Biling. Lang. Cogn.* 14, 371–378.

Takio, F., Koivisto, M., Jokiranta, L., Rashid, F., Kallio, J., Tuominen, T., Laukka, S., and Hämäläinen, H. (2009). The effect of age on attentional modulation in dichotic listening. *Dev. Neuropsychol.* 34, 225–239.

Ye, Z., and Zhou, X. (2009). Executive control in language processing. *Neurosci. Biobehav. Rev.* 33, 1168–1177.

Zelazo, P. D., Craik, F. I. M., and Booth, L. (2004). Executive function across the life span. *Acta Psychol. (Amst.)* 115, 167–183.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 March 2011; accepted: 21 July 2011; published online: 01 August 2011.

*Citation: Soveri A, Rodriguez-Fornells A and Laine M (2011) Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. Front. Psychology **2**:183. doi: 10.3389/fpsyg.2011.00183*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 25 November 2011
doi: 10.3389/fpsyg.2011.00324

Bilingualism and inhibitory control influence statistical learning of novel word forms

James Bartolotti *, Viorica Marian, Scott R. Schroeder and Anthony Shook

Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, USA

Edited by:
Judith F. Kroll, Penn State University, USA

Reviewed by:
Tamar H. Gollan, University of California, USA
Sara Finley, University of Rochester, USA

****Correspondence:***
James Bartolotti, Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA.
e-mail: j-bartolotti@u.northwestern.edu

We examined the influence of bilingual experience and inhibitory control on the ability to learn a novel language. Using a statistical learning paradigm, participants learned words in two novel languages that were based on the International Morse Code. First, participants listened to a continuous stream of words in a Morse code language to test their ability to segment words from continuous speech. Since Morse code does not overlap in form with natural languages, interference from known languages was minimized. Next, participants listened to another Morse code language composed of new words that conflicted with the first Morse code language. Interference in this second language was high due to conflict between languages and due to the presence of two colliding cues (compressed pauses between words and statistical regularities) that competed to define word boundaries. Results suggest that bilingual experience can improve word learning when interference from other languages is low, while inhibitory control ability can improve word learning when interference from other languages is high. We conclude that the ability to extract novel words from continuous speech is a skill that is affected both by linguistic factors, such as bilingual experience, and by cognitive abilities, such as inhibitory control.

Keywords: language acquisition, statistical learning, bilingualism, inhibitory control, Morse code, Simon task

INTRODUCTION

Learning a new language is a complex task comprised of mastering novel phonology, vocabulary, and grammar. Acquisition in adults occurs gradually, and even after years of practice many do not achieve native-like levels of pronunciation (Baker and Trofimovich, 2005; Sebastian-Gallés et al., 2006) or grammatical knowledge (Johnson and Newport, 1989; DeKeyser, 2005; MacWhinney, 2005). Even when adults are able to develop adequate vocabulary skills in a new language (Van Hell and Mahn, 1997; Lotto and De Groot, 1998), they often experience initial difficulty forming strong associations between a novel word's lexical form and its meaning (Kroll and Stewart, 1994; Kroll et al., 2002). An important component of learning success is early acquisition of word form, since focusing on learning isolated word forms first can contribute to subsequent learning of words' meanings (Bogaards, 2001; Graf Estes et al., 2007; Mirman et al., 2008; Fernandes et al., 2009). Acquisition of words and their forms has previously been explored by manipulating the learner's input, and it has been shown that variables such as repeated exposure to specific words (Nation, 2001; de Groot, 2006) and reduced speaking rates (Ferguson, 1975) can improve acquisition. In addition, characteristics of the learner may also contribute to successful acquisition. For example, experience with multiple languages has been associated with improved learning of words' form-meaning links (Cenoz and Valencia, 1994; Sanz, 2000; Cenoz, 2003; Keshavarz and Astaneh, 2004; Kaushanskaya and Marian, 2009b), and this learning advantage may arise in part from better initial acquisition of word form. Similarly, inhibitory control (the ability to suppress competing representations and attend to relevant ones) appears to influence learning (Kaushanskaya and Marian, 2009a) and processing (Bartolotti and Marian, 2010) of novel words, and may affect form acquisition. In the present study, we examined how early learning of word forms is affected by characteristics of the learner, including linguistic experience (in the form of bilingualism) and cognitive ability (in the form of inhibitory control).

Learning the forms of novel words is aided by frequent encounters with those words (Osterhout et al., 2006). McLaughlin et al. (2004) found that the best predictor of word familiarity was how frequently that word had appeared during previous instruction. Language learners who study abroad in immersive second language environments encounter specific novel words during daily exposure more often than students who do not, and as a result show greater gains in proficiency (Freed, 1995). The benefits of language immersion arise both from reduced exposure to the native language (Levy et al., 2007; Linck et al., 2009), and from increased exposure to words in the new language (Kojic-Sabo and Lightbown, 1999; Perani et al., 2003). Increased exposure to the new language can strengthen the representations of recently acquired words and introduce the learner to novel words more frequently. Novel words encountered while listening to speech can be acquired incidentally and can increase vocabulary knowledge considerably.

Incidental learning can be accomplished by using the statistical regularities in speech to determine the boundaries of novel word forms. Sounds that co-occur often are likely to comprise part of a single word, whereas rare sound sequences are likely to mark transitions between words. For example, in the phrase "pretty baby," listeners are sensitive to the fact that "*pre*" followed by "*ty*" is more likely to occur than "*ty*" followed by "*ba*," since "pretty" can be followed by any number of other words. Both infants and adults are able to track this statistical information and use it to identify novel word forms in an unfamiliar language (Saffran et al., 1996, 1999; Ludden and Gupta, 2000; Theissen and Saffran, 2003; Newport and Aslin, 2004; Kovács and Mehler, 2009).

This ability to learn novel forms in a new language via statistical regularities may be indirectly improved by previous bilingual experience. One of the consequences of bilingualism for cognition is improved phonological working memory (Service et al., 2002; Majerus et al., 2008; Adesope et al., 2010), as a result of acquiring and processing a large vocabulary that encompasses multiple languages. High phonological working memory has been associated with gains in statistical learning of word forms (Misyak and Christiansen, 2007), suggesting that bilingualism may improve statistical learning through its influence on phonological working memory. Phonological working memory can be used to maintain large chunks of speech in memory long enough for the transitions between syllables to be compared. In addition, working memory may help to update the relative frequency of different syllable transitions. Based on the transitional probabilities, likely word candidates can be identified and transferred from working memory to long-term memory. Due to gains in phonological working memory, bilinguals should thus outperform monolinguals in statistical learning of word forms in a novel language.

Learning word forms in a new language may also be influenced by level of inhibitory control. When a new language and a known language conflict, interference from the known language may be particularly detrimental to learning, since the two languages are tightly integrated at early stages of learning (Kroll and Stewart, 1994). Since known languages are highly practiced, they can activate more easily than a new language, resulting in learners over-applying transitional probabilities, pronunciations, or rules from their native language to the new language, even when the two are in conflict (Murphy, 2003). For example, the French possessive “de” occurs before many other French words and thus often marks a word boundary, whereas the same syllable in English is commonly used at the beginning of words and rarely indicates a word boundary. An English-speaking learner of French, then, may not attend to novel words following “de,” as this syllable was not a reliable English word boundary cue. By over-applying rules for English word transitions, the learner’s acquisition of French word forms may progress at a slower rate. By using inhibitory control to suppress the non-target language, interference from conflicting native-language constructs can be reduced. With less interference, word boundaries in a new language may be easier to learn and speech segmentation may be improved. Indeed, effective inhibitory control has previously been shown to benefit word segmentation when conflicting information present during learning had to be suppressed (Weiss et al., 2010).

To examine the distinct contributions of bilingual experience and inhibitory control on word segmentation, we tested participants who varied in bilingual experience and level of inhibitory control on their ability to learn languages that were based on the International Morse Code. In Morse code, all information is conveyed rhythmically by changes in duration of pure-tone sequences and silences. A benefit of using Morse code is that it is sufficiently difficult to learn and therefore can discriminate learners from non-learners. An additional benefit of using Morse code is that it does not overlap with any languages participants knew and avoids favoring speakers of one language over another. This low overlap with participants’ known languages enabled us to create an experimental condition in which interference was low and learning required detecting statistical regularities within the Morse stream, but did not require inhibiting competitive interference from known languages. Because the inhibitory demands were reduced, the low-interference condition allowed us to assess whether bilingual experience has an effect on incidental learning of word forms from speech, independent of inhibitory control ability.

In addition to the low-interference condition, we also designed a second, high-interference condition to assess the influence of inhibitory ability on word segmentation. The words in this second, high-interference condition conflicted with the previously learned words in the low-interference condition. Additionally, a colliding cue to word boundaries that conflicted with the transitional probabilities between words was inserted to create interference within the new language itself (Weiss et al., 2010). Weiss et al. (2010) showed that when two sets of word boundary cues were equally salient, participants with strong inhibitory control were able to selectively attend to one set of cues and learn the words. Although the source of the conflict in their study was within the target language instead of across two languages, inhibitory control may similarly improve learning when the locus of interference is between a known language and a new language. Learning in our high-interference condition depended on selectively attending to one of the two sets of word boundary cues (by inhibiting the other), as in Weiss et al. (2010), but also required participants to suppress competing Morse code words that were previously learned in the low-interference condition. This second, high-interference condition therefore enabled us to examine the influence of inhibitory ability on word segmentation in contexts where learners have to reduce interference from conflicting linguistic information both within and across languages.

To summarize, in the present study, we examined the distinct contributions of bilingual experience and inhibitory control on word segmentation. Participants who varied in bilingual experience and level of inhibitory control were taught Morse code words first in a low-interference condition and then in a high-interference condition. The low-interference condition placed few demands on inhibition; in this condition, high bilingual experience was expected to contribute to successful word segmentation. The high-interference condition placed high demands on inhibition; in this condition, inhibitory ability was expected to promote successful word segmentation.

MATERIALS AND METHODS

PARTICIPANTS

Twenty-four Northwestern University students (Mean age = 21.6 years, SD = 2.23) participated for course credit. All participants provided informed consent in accordance with the Northwestern University IRB. Participants completed the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007) to provide information about language proficiency, age of acquisition, and frequency of language use. Languages represented as participants' dominant language included English ($N = 20$), Korean ($N = 1$), and Chinese ($N = 1$). Second languages reported included Spanish ($N = 7$), Chinese ($N = 4$), English ($N = 2$), French ($N = 1$), Gujarati ($N = 1$), Korean ($N = 1$), and Tamil ($N = 1$); five participants reported no meaningful second language experience. A breakdown of language knowledge by group is provided in **Table A1** in the Appendix. Based on participants' reported L2 proficiency, L2 age of acquisition, and L2 frequency of use, a composite score of bilingual experience was computed. L2 proficiency, L2 age of acquisition, and L2 frequency of use were transformed to Z-scores for each participant based on the group mean and SD, and the average of these three scores was used as a composite measure of overall bilingual experience. Participant characteristics are presented in **Table 1**.

A version of the Simon task (Simon and Small, 1969) was used to assess participants' inhibitory control ability. Median splits were used to separate participants into high/low bilingual experience groups based on the bilingual experience composite, and strong/weak inhibitory control based on the size of the Simon effect (median: 33.24 ms). High and low bilingual experience groups did not differ in age, performance IQ (block design and matrix reasoning subtests of the *Weschler Abbreviated Scale of Intelligence*; PsychCorp, 1999), working memory (digit span subtest of the *Comprehensive Test of Phonological Processing*; Wagner et al., 1999), or inhibitory control ability. Strong and weak inhibitory control groups did not differ in age, performance IQ, working memory, L2 proficiency, L2 age of acquisition, or L2 frequency of use.

Inhibitory control as assessed by the Simon task was not correlated with the bilingual experience composite ($p = 0.50$) or any of its components (L2 proficiency, $p = 0.09$; L2 age of acquisition, $p = 0.84$; L2 frequency of use, $p = 0.73$), allowing the effects of bilingual experience and inhibitory control on learning to be considered separately. The lack of a correlation was not unexpected; bilingual advantages in inhibitory control are frequently observed in children (Bialystok and Martin, 2004; Carlson and Meltzoff, 2008) and older adults (Bialystok et al., 2004; Salvatierra and Rosselli, 2011), but results are mixed in younger adults who are in their cognitive prime. In particular, certain tasks of executive functioning, such as Stroop and the Attentional Network Test, commonly reveal bilingual advantages in young adults (Bialystok et al., 2008; Costa et al., 2008; Tao et al., 2011). Other executive functioning tasks, such as the Simon task, are reliable predictors of word segmentation in the presence of conflicting cues (Weiss et al., 2010), but do not appear to be robustly driven by bilingual experience (Bialystok et al., 2005; Bialystok, 2006; Prior and MacWhinney, 2010; Blumenfeld and Marian, 2011; Hilchey and Klein, 2011). In the present study, because variability in bilingual experience was not related to variability in inhibitory control, it was possible to examine the separate effects of these two factors on learning to segment words in a novel language.

MATERIALS

Two artificial languages were created based on the International Morse Code alphabet. In Morse code, letters are composed of combinations of short tones, or "dots" (440 Hz for 100 ms) and long tones, or "dashes" (440 Hz for 300 ms). Two letters (E /./ and T /-/) are made up of a single tone each, and four letters (A /. -/, I /. ./, N /- ./, and M /- -/) are made up of two tones in sequence. When a single letter contained two tones, the tones were separated by a short 100 ms pause. When multiple letters were combined to form a single word, the letters were separated by a longer 300 ms pause, so that the multi-tone letters (i.e., I, A, N, and M) could still be perceived as distinct groupings (without this distinction in pause lengths, the letter sequence "ET" would be indistinguishable from the single letter "A"). By using all six letters (A, E, I, N, M, T), three words were created for each of two languages such that the length of each word was a constant 1100 ms, and no letter was used twice (See **Figure 1**).

Morse code training streams were created for each language with two restrictions: A word could not immediately follow itself, and each word was followed by the other two words an equal number of times. Since the first letter of each word perfectly predicted the second letter, transitional probability within words was

Table 1 | Means and SD (in parentheses) for participant characteristics.

	All participants	Low bilingual experience	High bilingual experience	Weak IC	Strong IC
N	24	11†	11†	12	12
Females	20	10	8	11	9
Age (years)	21.61 (2.23)	22.09 (2.47)	21.18 (2.09)	22.00 (2.61)	21.25 (1.86)
WASI (performance IQ)	110.17 (12.51)	108.91 (12.45)	110.73 (13.46)	110.75 (13.53)	109.55 (11.92)
Digit span (raw score)	16.71 (2.48)	17.00 (2.41)	16.09 (2.66)	17.42 (2.23)	16.00 (2.59)
L2 Proficiency (scale 0–10)	4.50 (3.28)	2.23*** (2.57)	6.77*** (2.16)	4.10 (2.84)	4.91 (3.77)
L2 AoA (years)	7.88 (5.05)	12.33** (1.51)	5.45** (4.61)	9.22 (4.44)	6.38 (5.55)
L2 Frequency of use (%)	12.09 (17.40)	1.00** (2.00)	23.18** (19.01)	14.18 (21.25)	10.00 (13.23)
Simon effect (ms)	33.82 (25.11)	36.48 (28.75)	28.86 (22.43)	53.04*** (16.88)	14.60*** (4.87)

Low and High Bilingual Experience groups were defined by a median split on a composite score of L2 Proficiency, L2 AoA, and L2 Frequency of use. Weak and Strong Inhibitory Control groups were defined based on a median split of the size of the Simon effect.

*IC, inhibitory control; WASI, Weschler abbreviated scale of intelligence; L2, second language; AoA, age of acquisition. †Group sizes are lower in the high/low bilingual experience comparison because language data were not available for two participants. **p < 0.01 between groups, ***p < 0.001 between groups.*

a constant 1.0. Since each word could be followed by either of the two other words, the between-word transitional probability was a constant 0.5. The training stream in the low-interference condition had a 300-ms long pause inserted between words, identical to the long pause that separated letters within a single word. To learn the words, participants would have to attend to the transitional probabilities within and between words. For example, the continuous stream TAEMTANI can be segmented as TA–EM–TA–NI based on the transitional probabilities between letters (see **Figure 2A**). In contrast, in the high-interference condition, the long pause between words was replaced with a 100 ms short pause. The 300 ms long pause that remained within words could be used as a salient grouping cue to identify words in the stream.

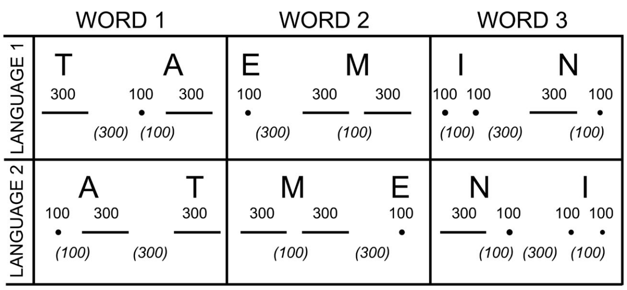


FIGURE 1 | The Morse code words used in the two languages. Long tones, or dashes, are 300 ms long, and short tones, or dots, are 100 ms long. Numbers in parentheses indicate the length of the pause, either 100 or 300 ms. Short pauses separate tones within a letter, and long pauses separate letters within a word.

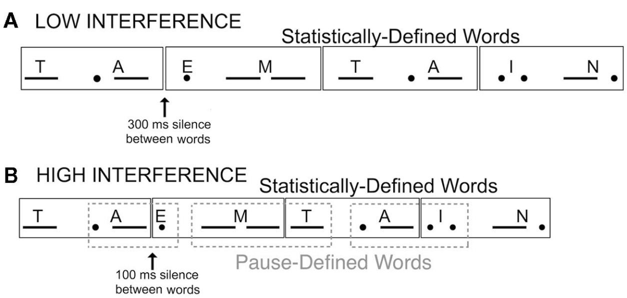


FIGURE 2 | Morse code listening streams in the low and high-interference conditions. The dots and dashes represent short (100 ms) and long (300 ms) tones respectively, while the short and long gaps represent silences of 100 and 300 ms respectively. In the low-interference condition **(A)**, words are marked by statistical probabilities between letters (the transitional probability within words is 100%, e.g., T is always followed by A, but the transitional probability between words is 50%, e.g., A is followed by either E or I). In the high-interference condition **(B)**, the gap between words is reduced, and the statistically defined words (TA, EM, IN) compete with words defined by the long pauses (AE, MT, AI, as well as MI, NT, NE). Each participant was exposed to both the low- and high-interference conditions. The order of the two conditions was fixed, with the low-interference condition always occurring first, but the language used in each condition was counterbalanced across subjects. In the example above, the same language is shown in both the low-interference condition and the high-interference conditions in order to highlight the difference between conditions.

If participants ignored the different pause lengths, they would still be able to learn the words based on the transitional probabilities, as in the low-interference condition. If instead participants used the pause lengths as a cue to word boundaries, they would learn a different set of words than those defined by the transitional probabilities. There were thus two colliding cues to word boundary: the between-word transitional probabilities (as in the low-interference condition), and the pause-based cues (see **Figure 2B**). To learn the words, participants would have to inhibit one of the two word boundary cues and attend to the other.

PROCEDURE

The Morse code language associated with each condition was counterbalanced across participants, so that half of the participants heard Language 1 for the low-interference condition and Language 2 for the high-interference condition, while the other half of the participants heard Language 2 for the low-interference condition and Language 1 for the high-interference condition. The order of the two conditions was fixed, with all participants completing the low-interference condition first, followed by the high-interference condition. This was done to ensure that no previously learned Morse code words could compete with targets during the low-interference condition. Learned words would then have to be inhibited during the following high-interference condition, increasing the inhibitory demands of the high-interference condition.

At the beginning of each learning condition, participants were instructed to listen to a series of tones and were told that they would be tested on information about the tones later. Participants wore headphones and listened to the Morse code stream over three blocks, each 4 min and 12 s long. Participants received a 1-min silent break between blocks.

Immediately after the third training block, participants were tested on their knowledge of the language with a 12-item two-alternative forced-choice task. Participants were instructed to indicate which of two Morse code words was more familiar by pressing the "1" (first word) or "9" (second word) key on a computer keyboard. Word pairs were presented with a 1-s pause between words, and a 4-s pause between trials. Each of the three words was presented in four trials: twice before and twice after two different part-words. Part-words were created by concatenating the second letter from one word with the first letter of another word, and had appeared in the listening stream half as often as the actual words. In the high-interference condition, the part-words were words that could have been learned by using pause-based cues instead of statistical cues. Accuracy scores were obtained and normalized to chance performance, with a score of 0 indicating 6 out of 12 correct (where correct meant selecting the statistically defined word). Positive scores indicated learning of the statistical probabilities. In the high-interference condition, negative scores indicated learning of the pause-based rules.

All participants also completed a visual Simon task (Simon and Small, 1969; Weiss et al., 2010) to index inhibitory control. Participants viewed blue and brown rectangles that appeared on the left, right, or center of a computer screen and selected a response based on the item's color, while ignoring its location. The instructions were to press a blue button on the left side of the keyboard if the rectangle was blue, or to press a brown button on the right side of the keyboard if the rectangle was brown. In Congruent trials, the stimulus and the response were on the same side (e.g., a blue rectangle on the left side of the screen). In Incongruent trials, stimulus and response were on opposite sides (e.g., blue rectangle on the right side of the screen). In Neutral trials, the stimulus appeared in the center of the screen. Congruent, Incongruent, and Neutral trials appeared in an equal ratio (42 trials each, 126 total). A single trial involved (1) a fixation cross for 350 ms, (2) a blank screen for 150 ms, (3) a colored rectangle for 1500 ms, (4) in the event of an error, a red "X" as feedback for 1500 ms, and (5) a blank screen for an 850 ms inter-trial interval. All participants completed a practice session before the actual task. The Simon effect was calculated by subtracting reaction time on Congruent trials from reaction time on Incongruent trials. A small Simon effect indicates better ability to ignore the inconsistent location cue, and strong inhibitory control.

RESULTS

WORD LEARNING IN THE LOW-INTERFERENCE CONDITION

When interference during learning was low, bilingual experience positively influenced word learning ability, whereas level of inhibitory control did not influence learning (**Figure 3**). The high bilingual experience group performed significantly better than chance ($M = 2.41$, $SD = 2.01$), $t(10) = 3.98$, $p < 0.01$, while the low bilingual experience group did not differ from chance ($M = 1.09$, $SD = 2.34$), $p = 0.15$, indicating that the high bilingual experience group was able to learn the Morse code words. The same pattern of results was observed when each factor in the bilingual experience composite was considered separately, that is, when participants were divided into two groups based on median splits in L2 proficiency, L2 age of acquisition, or L2 frequency of use. Both the strong ($M = 1.79$, $SD = 2.46$), $t(11) = 2.52$, $p < 0.05$, and weak ($M = 1.92$, $SD = 1.98$), $t(11) = 3.36$, $p < 0.01$, inhibitory control groups performed above chance, indicating that they were able to learn the Morse code language.

Learning was not correlated with bilingual experience, $p = 0.76$, L2 age of acquisition, $p = 0.40$, or L2 frequency of use, $p = 0.55$, but was marginally correlated with L2 proficiency, $r = 0.40$, $p = 0.06$ (**Figure 4**). Inhibitory control ability was not correlated with learning, $p = 0.37$.

WORD LEARNING IN THE HIGH-INTERFERENCE CONDITION

When interference during learning was high, strong inhibitory control increased word learning, but bilingual experience did not (**Figure 5**). In this condition, positive scores above chance indicate learning of the words based on statistical cues, while negative scores below chance indicate learning of the words based on pause cues. Participants with strong inhibitory control performed significantly below chance, indicating that they learned according to the pause cues ($M = -1.18$, $SD = 1.60$), $t(10) = -2.45$, $p < 0.05$, while participants with weak inhibitory control did not differ from chance ($M = 0.58$, $SD = 1.73$), $p = 0.27$. No consistent pattern of learning was observed when bilingual experience was considered; neither the high bilingual experience group ($M = 0.18$, $SD = 2.14$; $p = 0.78$), nor the low bilingual experience group ($M = -0.55$,

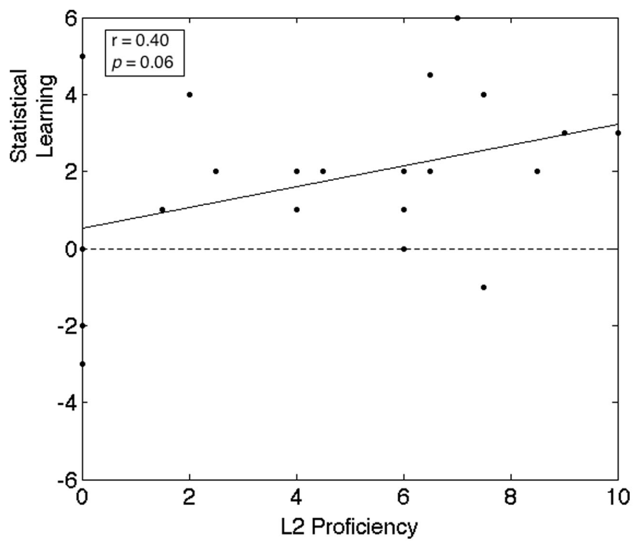


FIGURE 4 | Relationship between second language proficiency and learning of word forms in the low-interference condition.

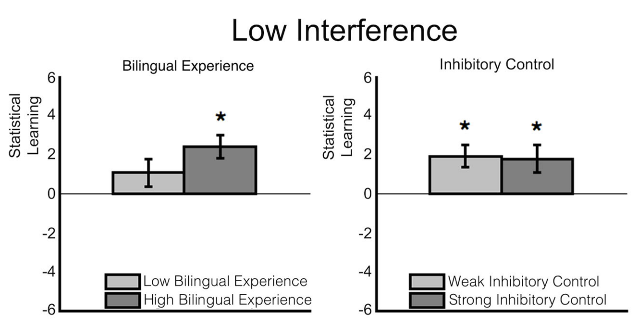


FIGURE 3 | Effects of second language proficiency and inhibitory control on learning the new language in the low-interference condition. (Asterisks indicate a significant difference from chance, alpha of 0.05. Error bars indicate 1 SE).

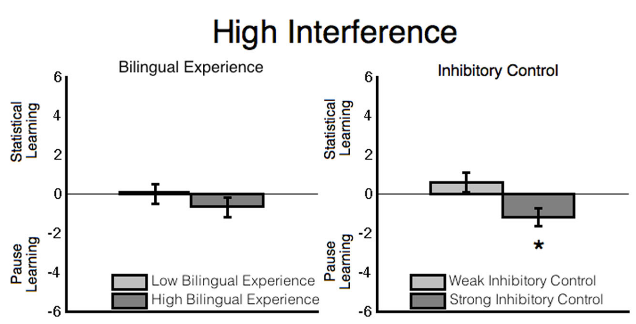


FIGURE 5 | Effects of second language proficiency and inhibitory control on learning the new language in the high-interference condition. Positive scores indicate learning of statistical cues and negative scores indicate learning of pause cues. (Asterisks indicate a significant difference from chance, alpha of 0.05. Error bars indicate 1 SE).

SD = 1.57), $p = 0.28$ differed from chance. In addition, when each factor in the bilingual experience composite was considered separately, no group performed better than chance.

Inhibitory control was correlated with learning success, $r = 0.47$, $p < 0.05$ (**Figure 6**), while bilingual experience, $p = 0.94$, L2 proficiency, $p = 0.99$, L2 age of acquisition, $p = 0.86$, and L2 frequency of use, $p = 0.75$ were not correlated with learning.

DISCUSSION

Learning words in a new language is a multi-step process involving the acquisition of new word forms and of mapping these acquired word forms to meaning. While previous research suggests that bilingualism improves learning of form-meaning mappings in another language (Cenoz and Valencia, 1994; Sanz, 2000; Cenoz, 2003; Keshavarz and Astaneh, 2004; Kaushanskaya and Marian, 2009b), in the current study we found that bilingual experience can improve acquisition of word forms alone when interference between languages is low. In addition, we showed that inhibitory control promoted successful word segmentation when there were competing cues to word boundaries both within and across languages. Successful acquisition of word forms has previously been shown to be one factor that contributes to later stages of word learning (e.g., mapping form to meaning, Graf Estes et al., 2007; Mirman et al., 2008) and increases the rate at which vocabulary is expanded (Bogaards, 2001). During natural language learning, both bilingual experience and inhibitory control may contribute in different degrees to early acquisition of novel word forms (depending on the characteristics of the language to be learned), which may benefit the process of learning a novel language.

High bilingual experience was associated with successful segmentation of Morse code word forms from a continuous auditory stream with consistent cues to word boundaries. In order to learn the word forms in the low-interference condition, participants had to analyze the transitional probabilities between sounds and extract the most commonly occurring sequences. This ability to analyze probabilities may depend in part on phonological working memory, which has previously been associated with improved statistical learning (Misyak and Christiansen, 2007). Extensive bilingual experience has been associated with gains in phonological working memory (Service et al., 2002; Majerus et al., 2008; Adesope et al., 2010), which may have contributed to bilinguals' ability to learn the words. It is possible that bilinguals used working memory more effectively than monolinguals to maintain large chunks of the auditory sequence for statistical analysis. Working memory could also contribute by updating the frequencies of specific transitions over time, and by facilitating the transfer of newly detected words to long-term memory. By effectively applying phonological working memory to the statistical learning task, bilinguals may have been able to better extract and retain novel word forms.

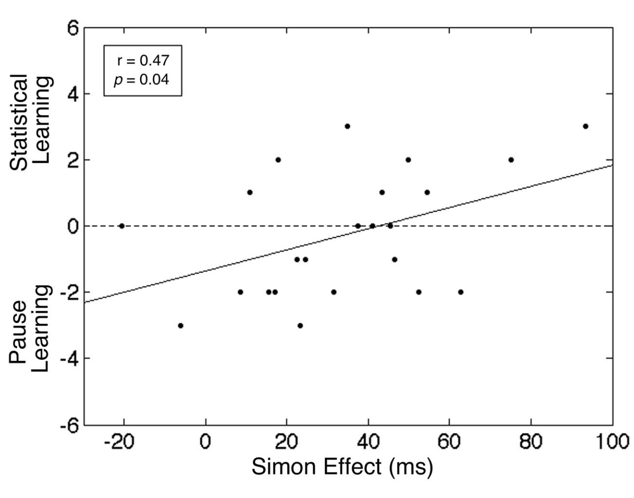


FIGURE 6 | Relationship between inhibitory control (assessed by the Simon task) and learning of word forms in the high-interference condition.

Statistical learning is itself a measure of implicit learning ability, as participants are typically not aware of having consciously learned any of the words. Bilinguals' improved performance on the statistical learning task is thus consistent with observed bilingual advantages on language learning tasks that rely heavily on implicit learning (Klein, 1995; Kovács and Mehler, 2009). For example, Nation and McLaughlin (1986) found that proficiency in multiple languages improved learning of an artificial grammar when participants did not explicitly attend to the grammatical rules, but acquired them implicitly during the course of learning novel words. As a consequence of acquiring the words and grammar of multiple languages, bilinguals may develop a more effective implicit learning mechanism than monolinguals. This increased efficiency could contribute to bilinguals' improved incidental learning of word forms while listening to speech.

In contrast to bilinguals' performance in the low-interference condition, in the high-interference condition, bilingual experience had no effect on word segmentation success. One possibility for the lack of learning is that both those with low and those with high bilingual experience may have been unable to consistently attend to either the statistical or pause-based cues. If participants shifted attention between the two cues during training, then at test neither the statistically defined words nor the pause-defined words would be more familiar and performance would remain at chance. Alternatively, it may be that bilingual experience improves efficiency of integrating multiple cues. Given that most languages use correlated cues to word segmentation and relatively few contrasting cues (Christiansen et al., 2005; Sahni et al., 2010), when cues are not correlated (as was the case in the high-interference condition), this ability to integrate cues may be a drawback. Learning in the high-interference condition required that participants attend to a single set of cues and suppress the other; bilinguals may have either been unable to attend to either cue, or attended to and integrated both cues.

Participants with strong inhibitory control were able to selectively attend to a single set of cues in the high-interference condition, suggesting that inhibitory control can also contribute to word segmentation ability. In the high-interference condition, conflict occurred due to both incongruent word boundaries between the two Morse code languages across blocks, and colliding cues to word boundary within the listening stream. Learning word boundaries in the high-interference condition required one to ignore recently learned transitional probabilities from the low-interference condition, and to selectively attend to one of the two colliding cues to word boundaries in the high-interference condition. Successful learning could be accomplished by inhibiting irrelevant information in memory (previous transitional probabilities) and in the auditory stream (one of the two colliding cues to word boundaries).

The overall pattern of learning in the high-interference condition suggests that participants with strong inhibitory control suppressed the statistical information and relied on pause lengths between letters to segment words. One possibility is that pause lengths were a more salient cue than the transitional probabilities, making them easier to learn. However, if the pause boundaries in the auditory stream had been much more salient than the statistical boundaries, we might have expected all groups to pick up on this cue and learn the pause-defined words. In a previous study using colliding statistical and pause-based cues to word segmentation, when the pause cue was made more salient (by manipulating its length), participants overwhelmingly were able to learn the pause-defined words (Weiss et al., 2010). In our colliding cue condition we saw evidence of learning only in the strong inhibitory control group, which suggests that the pause cues were learnable, but that the statistical cues were close enough in saliency to interfere with learning in the weak inhibitory control group. The tendency of the strong inhibitory control group to segment words according to the pauses may reflect a strategy that minimized sources of interference. Recall that in the high-interference condition, pauses conflicted with transitional probabilities, while the transitional probabilities conflicted with both the pauses and the transitional probabilities from the low-interference condition. The pauses thus directly competed with only one source, while the statistical boundaries directly competed with two sources. The participants with strong inhibitory control may have been sensitive to this difference and applied inhibition in a way that maximized cue salience, by suppressing all statistical cues and engaging learning of the pauses between words.

To summarize, our findings suggest that experience with a second language helped learners identify novel words by attending to statistical regularities in the signal, whereas inhibitory control helped learners identify novel words by suppressing conflicting language knowledge and focusing attention on the meaningful aspects of a novel language. To date, there has often been considerable attention paid to how bilingual experience may impact executive functioning or its subcomponents, including response suppression, inhibitory control, task switching, and task monitoring (Bialystok and Martin, 2004; Bialystok et al., 2004; Bialystok et al., 2008; Carlson and Meltzoff, 2008; Costa et al., 2008, 2009; Hernández et al., 2010; Prior and MacWhinney, 2010; Soveri et al., 2010; Salvatierra and Rosselli, 2011; Tao et al., 2011). While bilingual advantages are typically more robust in young children or older adults (see Hilchey and Klein, 2011), they can be observed in young adults, particularly on tasks that require context monitoring (Costa et al., 2008, 2009; Prior and MacWhinney, 2010; Tao et al., 2011). The link between bilingualism and executive functioning is thought to stem from bilinguals' need to control language access. Both of a bilingual's languages remain active when only one is present in the immediate linguistic context, requiring the bilingual to monitor the language in use, selectively attend to the target language, and inhibit the non-target language. Constant training of the executive functions recruited to direct attention during language processing may improve executive functioning in other domains. However, in young adults, inhibitory control ability appears to be influenced by other factors besides bilingualism as well (Bialystok et al., 2005; Bialystok, 2006; Prior and MacWhinney, 2010; Blumenfeld and Marian, 2011; Hilchey and Klein, 2011). In the present study, we were able to examine the differential effects of bilingual experience and inhibitory control on learning to segment words in two Morse code languages that differed in the strength of conflicting information. By examining acquisition of word forms in these different learning contexts, we have shown that linguistic and cognitive characteristics of the learner can affect success at an early stage of language learning, specifically, during word form acquisition.

One potential limitation of the current study is that the Morse code languages that participants learned were composed of pure tones that do not closely resemble natural speech. The choice to use pure-tone stimuli was made in order to limit transfer of prior knowledge during learning. Bilinguals have been shown to readily transfer words and grammatical structures from languages they already know when it can facilitate learning (Cenoz, 1997; Murphy, 2003), and using Morse code stimuli avoided confounding bilingual experience with increased transfer of prior language knowledge. By using word forms based on Morse code, we were able to control participants' prior experience with the target language, and since language backgrounds were unlikely to confer a benefit, we were able to specifically target the effects of bilingual experience and inhibitory control on sequence learning. It is important that future research extends the findings from the current study to natural language learning, as there is reason to believe that the processes involved in learning the Morse code words and in natural language acquisition overlap. The ability to extract information from a continuous stream through statistical learning mechanisms appears to be a domain-general skill, and has been shown to affect sequence learning of musical tones (Saffran et al., 1999), visual shapes (Kirkham et al., 2002), and tactile stimuli (Conway and Christiansen, 2005), as well as that of non-word syllables (Saffran et al., 1996; Ludden and Gupta, 2000; Theissen and Saffran, 2003; Newport and Aslin, 2004; Kovács and Mehler, 2009). In addition, sequence learning skill has been shown to correlate positively with second language learning success in a classroom setting (Ettlinger et al., 2011), suggesting that word segmentation ability can contribute to natural language learning.

It is likely that previous bilingual experience and inhibitory control ability work simultaneously to promote learning, but their relative influences may depend on the relationship between known languages and the target language. For example, bilingualism may be a more important factor in learning word forms when the target language contains novel, non-overlapping features, such as the distinct writing systems between English and Chinese. Inhibitory control may be more important in promoting learning when the two languages conflict, such as the shared Roman alphabet but contrasting letter to phoneme mappings between English and French. As each case of novel language learning contains non-overlapping and conflicting components, both bilingual experience and inhibitory control are likely to be contributing factors to early acquisition of novel word forms, though their specific influences will depend on the characteristics of the novel language and already known languages.

In conclusion, the present study extends previous research on the role of linguistic experience and inhibitory control in later stages of language learning to early stages of language acquisition. While previous work has shown that linguistic experience and inhibitory control impact acquisition of and access to *form-meaning mappings* (Cenoz and Valencia, 1994; Sanz, 2000; Cenoz, 2003; Keshavarz and Astaneh, 2004; Kaushanskaya and Marian, 2009a,b; Bartolotti and Marian, 2010), we propose that linguistic experience and inhibitory control also influence initial acquisition of *word form*. Moreover, our results suggest that linguistic experience and inhibitory control may affect learning in different ways, depending on the relationship between the language to be learned and prior linguistic knowledge. The current study suggests that internal factors such as linguistic experience and cognitive ability can interact with external factors such as a new language's structure and its conflict with known languages to influence early components of language learning. Future work will need to examine how these interactions influence later stages of language learning. Investigating how internal and external factors interact within the learning process is essential for understanding ultimate language attainment.

ACKNOWLEDGMENTS

This research was funded in part by grants NICHD 1R01HD059858 and NSF BCS-0418495 to the second author. The authors would like to acknowledge Sarah Chabal, Jen Krizman, Natalia Daniel, and Vivian Chang for comments on an earlier draft of this paper.

REFERENCES

Adesope, O. O., Lavin, T., Thompson, T., and Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Rev. Educ. Res.* 80, 207–245.

Baker, W., and Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Lang. Speech* 48(Pt 1), 1–27.

Bartolotti, J. V., and Marian, V. (2010). "Linguistic control in monolingual and bilingual language learners," in *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, eds S. Ohlsson and R. Catrambone (Austin, TX: Cognitive Science Society), 532.

Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79.

Bialystok, E., Craik, F., and Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 859–873.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., and Martin, M. M. (2004). Attention and inhibition in bilingual children: evidence from the dimensional change card sort task. *Dev. Sci.* 7, 325–339.

Bialystok, E., Martin, M. M., and Viswanathan, M. (2005). Bilingualism across the lifespan: the rise and fall of inhibitory control. *Int. J. Biling.* 9, 103–119.

Blumenfeld, H. K., and Marian, V. (2011). "Bilingualism influences perceptual inhibition more than stimulus-response inhibition," in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, eds L. Carlson, C. Hölscher, and T. Shipley (Austin, TX: Cognitive Science Society), 3009.

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Stud. Second Lang. Acquistion* 23, 321–343.

Carlson, S., and Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Dev. Sci.* 11, 282–298.

Cenoz, J. (1997). "The influence of bilingualism on multilingual acquisition: some data from the Basque Country," in *I Simposio Internacional sobre o Bilingüismo: Comunidades e individuos bilingües,* Universidade de Vigo, Vigo, 278–287.

Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: a review. *Int. J. Biling.* 7, 71–87.

Cenoz, J., and Valencia, J. F. (1994). Additive trilingualism: evidence from the Basque country. *Appl. Psycholinguist.* 15, 195–207.

Christiansen, M. H., Conway, C. M., and Curtin, S. (2005). "Multiple-cue integration in language acquisition: a connectionist model of speech segmentation and rule-like behavior," in *Language Acquisition, Change and Emergence: Essay in Evolutionary Linguistics*, eds J. W. Minett and W. S. Y. Wang (Hong Kong: City University of Hong Kong Press), 205–249.

Conway, C. M., and Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 24–39.

Costa, A., Hernández, M., Costa-Faidella, J., and Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: now you see it, now you don't. *Cognition* 113, 135–149.

Costa, A., Hernández, M., and Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: evidence from the ANT task. *Cognition* 106, 59–86.

de Groot, A. M. B. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Lang. Learn.* 56, 463–506.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Lang. Learn.* 55, 1–25.

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., and Wong, P. C. M. (2011). The relationship between artificial and natural language learning. *85th Annual Meeting of the Linguistic. Society of America,* Pittsburgh.

Ferguson, C. (1975). "Sound patterns in language acquisition," in *Developmental Psycholinguistics: Theory and Applications, Georgetown University Round Table on Languages and Linguistics*, ed. D. P. Daniel (Washington, DC: Georgetown University Press), 1–16.

Fernandes, T., Kolinsky, R., and Ventura, P. (2009). The metamorphosis of the statistical segmentation output: lexicalization during artificial language learning. *Cognition* 112, 349–366.

Freed, B. F. (ed.). (1995). *Second Language Acquisition in a Study Abroad Context.* Philadelphia: John Benjamins.

Graf Estes, K., Evans, J. L., Alibali, M. W., and Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychol. Sci.* 18, 254–260.

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., and Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Biling. (Camb. Engl.)* 13, 315–325.

Hilchey, M. D., and Klein, R. M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychon. Bull. Rev.* 18, 625–658.

Johnson, J. S., and Newport, E. L. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cogn. Psychol.* 21, 60–99.

Kaushanskaya, M., and Marian, V. (2009a). Bilingualism reduces native-language interference during novel-word learning. *J. Exp. Psychol. Learn. Mem Cogn.* 35, 829–835.

Kaushanskaya, M., and Marian, V. (2009b). The bilingual advantage in novel word learning. *Psychon. Bull. Rev.* 16, 705–710.

Keshavarz, M. H., and Astaneh, H. (2004). The impact of bilinguality on the learning of English vocabulary as a foreign language (L3). *Int. J. Biling. Educ. Biling.* 7, 295–302.

Kirkham, N. Z., Slemmer, J. A., and Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain-general learning mechanism. *Cognition* 83, B35–B42.

Klein, E. (1995). Second versus third language acquisition: is there a difference? *Lang. Learn.* 45, 419–465.

Kojic-Sabo, I., and Lightbown, P. M. (1999). Students' approaches to vocabulary learning and their relationship to success. *Mod. Lang. J.* 83, 176–192.

Kovács, A. M., and Mehler, J. (2009). Flexible learning of multiple speech structures in bilingual infants. *Science* 325, 611–612.

Kroll, J. F., Michael, E., Tokowicz, N., and Dufour, R. (2002). The development of lexical fluency in a second language. *Second Lang. Res.* 18, 137–171.

Kroll, J. F., and Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *J. Mem. Lang.* 33, 149–174.

Levy, B. J., McVeigh, N. D., Marful, A., and Anderson, M. C. (2007). Inhibiting your native language: the role of retrieval-induced forgetting during second-language acquisition. *Psychol. Sci.* 18, 29–34.

Linck, J. A., Kroll, J. F., and Sunderman, G. (2009). Losing access to the native language while immersed in a second language: evidence for the role of inhibition in second-language learning. *Psychol. Sci.* 20, 1507–1515.

Lotto, L., and De Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Lang. Learn.* 48, 31–69.

Ludden, D., and Gupta, P. (2000). "Zen in the art of language acquisition: statistical learning and the less is more hypothesis," in *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*, eds L. R. Gleitman and A. K. Joshi (Hillsdale, NJ: Erlbaum), 812–817.

MacWhinney, B. (2005). "A unified model of language acquisition," in *Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. De Groot (Oxford: Oxford University Press), 49–67.

Majerus, S., Poncelet, M., Van der Linden, M., and Weekes, B. S. (2008). Lexical learning in bilingual adults: the relative importance of short-term memory for serial order and phonological knowledge. *Cognition* 107, 395–419.

Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967.

McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat. Neurosci.* 7, 703–704.

Mirman, D., Magnuson, J. S., Graf Estes, K., and Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition* 108, 271–280.

Misyak, J. B., and Christiansen, M. H. (2007). Extending statistical learning farther and further: long-distance dependencies, and individual differences in statistical learning and language," in *Proceedings of the 33rd Annual Meeting of the Cognitive Science Society*, eds D. S. McNamara and J. G. Trafton (Austin, TX: Cognitive Science Society), 1307–1312.

Murphy, S. (2003). Second language transfer during third language acquisition. *Work. Pap. TESOL Appl. Linguist.* 3, 1–21.

Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, R., and McLaughlin, B. (1986). Novices and experts: an information processing approach to the "good language learner" problem. *Appl. Psycholinguist.* 7, 41–55.

Newport, E. L., and Aslin, R. N. (2004). Learning at a distance: statistical learning of non-adjacent dependencies. *Cogn. Psychol.* 48, 127–162.

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., and Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Lang. Learn.* 56, 199–230.

Perani, D., Abutalebi, J., Paulesu, E., Brambati, S., Scifo, P., Cappa, S. F., and Fazio, F. (2003). The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum. Brain Mapp.* 19, 170–182.

Prior, A., and MacWhinney, B. (2010). A bilingual advantage in task switching. *Biling. (Camb. Engl.)* 13, 253–262.

PsychCorp. (1999). *Wechsler Abbreviated Scale of Intelligence Manual.* San Antonio, TX: Harcourt Assessment

Saffran, J. R., Aslin, R. N., and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52.

Sahni, S. D., Seidenberg, M. S., and Saffran, J. R. (2010). Connecting cues: overlapping regularities support cue discovery in infancy. *Child Dev.* 81, 727–736.

Salvatierra, J. L., and Rosselli, M. (2011). The effect of bilingualism and age on inhibitory control. *Int. J. Biling.* 15, 26–37.

Sanz, C. (2000). Bilingual education enhances third language acquisition: evidence from Catalonia. *Appl. Psycholinguist.* 21, 23–44.

Sebastian-Gallés, N., Rodríguez-Fornells, A., De Diego-Balaguer, R., and Díaz, B. (2006). First- and second-language phonological representations in the mental lexicon. *J. Cogn. Neurosci.* 18, 1277–1291.

Service, E., Simola, M., Metsänheimo, O., and Maury, S. (2002). Bilingual working memory span is affected by language skill. *Eur. J. Cogn. Psychol.* 14, 383–408.

Simon, J. R., and Small, A. M. (1969). Processing auditory information: interference from an irrelevant cue. *J. Appl. Psychol.* 53, 433–435.

Soveri, A., Laine, M., Hämäläinen, H., and Hugdahl, K. (2010). Bilingual advantage in attentional control: evidence from the forced-attention dichotic listening paradigm. *Biling. (Camb. Engl.)* 14, 371–378.

Tao, L., Marzecová, A., Taft, M., Asanowicz, D., and Wodniecka, Z. (2011). The efficiency of attentional networks in early and late bilinguals: the role of age of acquisition. *Front. Psychol.* 2:123. doi:10.3389/fpsyg.2011.00123

Theissen, E. D., and Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Dev. Psychol.* 39, 706–716.

Van Hell, J. G., and Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: learning concrete and abstract foreign words by experienced and inexperienced learners. *Lang. Learn.* 47, 507–546.

Wagner, R. K., Torgesen, J. K., and Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing: CTOPP*. Austin, TX: Pro-Ed, Inc.

Weiss, D., Gerfen, C., and Mitchel, A. (2010). Colliding cues in word segmentation: the role of cue strength and general cognitive processes. *Lang. Cogn. Process.* 25, 402–422.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 02 March 2011; accepted: 21 October 2011; published online: 25 November 2011.

Citation: Bartolotti J, Marian V, Schroeder SR and Shook A (2011) Bilingualism and inhibitory control influence statistical learning of novel word forms. Front. Psychology 2:324. doi: 10.3389/fpsyg.2011.00324

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDIX

Table A1 | First and second language knowledge by group.

	All participants (*N*)		Low bilingual experience (*N*)		High bilingual experience (*N*)		Weak IC (*N*)		Strong IC (*N*)	
	L1	**L2**	**L1**	**L2**	**L1**	**L2**	**L1**	**L2**	**L1**	**L2**
English	20	2	10	1	10	1	10	1	10	1
Spanish	–	7	–	3	–	4	–	4	–	3
Chinese	1	4	1	1	–	3	1	2	–	2
Korean	1	1	–	–	1	1	–	–	1	1
French	–	1	–	1	–	–	–	1	–	–
Gujarati	–	1	–	–	–	1	–	–	–	1
Tamil	–	1	–	–	–	1	–	1	–	–

IC, inhibitory Control; L1, first language; L2, second language.

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 16 March 2012
doi: 10.3389/fpsyg.2012.00057

Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech

Irina Pivneva*[1,2]*, Caroline Palmer*[1,2] *and Debra Titone*[1,2]**

[1] Department of Psychology, McGill University, Montreal, QC, Canada
[2] Centre for Research on Brain, Language, and Music, McGill University, Montreal, QC, Canada

Edited by:
Judith F. Kroll, Penn State University, USA

Reviewed by:
Miriam Gade, University of Zurich, Switzerland
Noriko Hoshino, Kobe City University of Foreign Studies, Japan
Gerrit Jan Kootstra, Radboud University Nijmegen, Netherlands

****Correspondence:***
Debra Titone, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, QC H3A 1B1, Canada.
e-mail: dtitone@psych.mcgill.ca

Bilingual language production requires that speakers recruit inhibitory control (IC) to optimally balance the activation of more than one linguistic system when they produce speech. Moreover, the amount of IC necessary to maintain an optimal balance is likely to vary across individuals as a function of second language (L2) proficiency and inhibitory capacity, as well as the demands of a particular communicative situation. Here, we investigate how these factors relate to bilingual language production across monologue and dialogue spontaneous speech. In these tasks, 42 English–French and French–English bilinguals produced spontaneous speech in their first language (L1) and their L2, with and without a conversational partner. Participants also completed a separate battery that assessed L2 proficiency and inhibitory capacity. The results showed that L2 vs. L1 production was generally more effortful, as was dialogue vs. monologue speech production although the clarity of what was produced was higher for dialogues vs. monologues. As well, language production effort significantly varied as a function of individual differences in L2 proficiency and inhibitory capacity. Taken together, the overall pattern of findings suggests that both increased L2 proficiency and inhibitory capacity relate to efficient language production during spontaneous monologue and dialogue speech.

Keywords: bilingualism, dialogue, monologue, inhibition, proficiency

INTRODUCTION

Speaking in one's first language (L1) is subjectively effortless, yet speech production involves a complex set of linguistic operations that require cognitive control (Kempen and Hoenkamp, 1987; Levelt, 1989). Speakers first conceptualize a message and then activate words in memory that are semantically and syntactically compatible with the message. Speakers then select from among this set the specific words that best convey the message, plan their articulation, and finally, implement the speech plan and produce their message at a rate of about 150–300 words per minute (Goldman-Eisler, 1968). These processes are incremental in that speakers transfer partially prepared fragments of the message from one stage to the next before completely preparing the message in its entirety. Thus, speakers begin articulating earlier parts of the message before fully activating and planning later parts of the message. The net effect of these cascaded and incremental speech processes is that native language production is quite cognitively demanding, in terms of word finding and word choice, grammatical and phonological realization, and overall fluency (Levelt, 1989; Dell et al., 1999; Griffin and Ferreira, 2006).

The production of fluent speech is likely to require even greater cognitive control for bilingual speakers, who face the challenges just described, as well as demands associated with knowing and using more than one language (Kroll et al., 2008; Colomé and Miozzo, 2010; De Groot, 2011). These added demands include a greater need to manage cross-language competition arising from parallel activation of two languages (Kroll et al., 2006, 2008), less practice using inhibitory control (IC) during L2 speech production (Abutalebi and Green, 2007), and weaker links between conceptual and linguistic representations in the L2 and possibly L1 (Poulisse and Bongaerts, 1994; Gollan et al., 2008). Indeed, recent work suggests that the added demands of bilingual language processing might lead to enhanced non-linguistic cognitive function for processes necessary to reduce cross-language competition, such as inhibitory capacity and selective attention (Bialystok et al., 2004; Bialystok, 2009).

In this study, we investigate how individual differences among bilinguals in L2 proficiency and inhibitory capacity modulate language production during spontaneous monologue and dialogue speech. Our theoretical framework derives from the IC model of bilingual language production, which is depicted in **Figure 1** (Green, 1998). A core assumption of this model is that language production is a communicative action that is analogous to non-linguistic physical actions (Green, 1998; Abutalebi and Green, 2007). Like physical actions, bilingual language production consists of mental task schemas, which are action sequences that are implemented by a conceptualizer (C). These task schemas achieve particular goals (G), which may be routine (L1 production) or non-routine (L2 production). For any given goal, parallel activation of multiple task schemas compete to control output (O). Consequently, the supervisory attentional system (SAS; Shallice and Burgess, 1996) suppresses routine goals via IC operations,

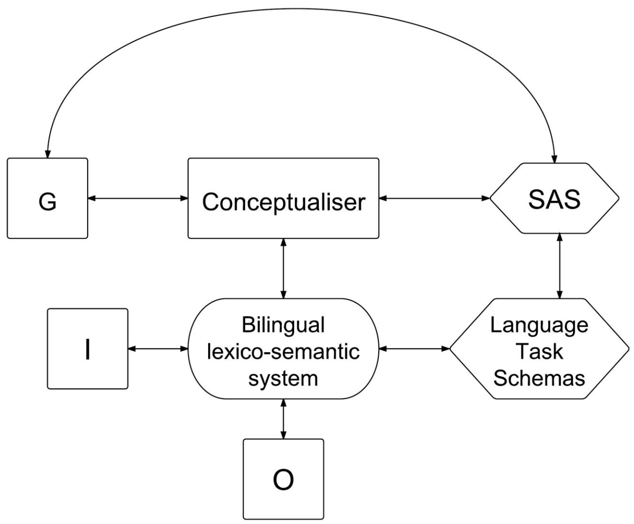


FIGURE 1 | Inhibitory control (IC) model of bilingual language production (Green, 1998).

and monitors the successful implementation of non-routine goals, based on input from the bilingual lexico-semantic system. Accordingly, when a bilingual speaker engages in a dialogue with a monolingual speaker in their L2, the conceptualizer relays input (I) from the bilingual lexico-semantic system to the SAS, which, in turn, implements greater IC to globally suppress the irrelevant but more routine L1 dialogue language schema. As well, within the bilingual lexico-semantic system, IC fine-tunes the relative activation and inhibition of words within each language to select and output appropriate words for the dialogue.

Abutalebi and Green (2007) extended the IC model to incorporate neurocognitive evidence about bilingual language production. They identified a network of cortical regions (prefrontal, inferior parietal, and anterior cingulate cortices) and subcortical structures (basal ganglia, the head of the caudate nucleus in particular) that modulate competition between L1 and L2 knowledge activation during bilingual language production. Within this framework subcortical structures (basal ganglia) modulate the global activation of L1 or L2 task schemas, whereas frontal cortical structures modulate local activation of L1 and L2 lexical activation. Using this framework, the authors also make more specific claims about the role of L2 proficiency. When L2 proficiency is low, L2 language production is more controlled and less automatic (see also Favreau and Segalowitz, 1983; Segalowitz and Hulstijn, 2005; Segalowitz, 2010), thus requiring IC (prefrontal function, in particular; see also Petrides, 1998). In contrast, when L2 proficiency is high, L2 production is automatic and less dependent on IC, although L1 production effort might instead increase due to a weakening of the links between word forms and concepts in the L1 (Bialystok, 2001; Michael and Gollan, 2005; Gollan et al., 2008, 2011; Ivanova and Costa, 2008; Bialystok et al., 2010).

Thus, the IC model (Green, 1998) and its extension (Abutalebi and Green, 2007) make several logical predictions about the role of IC during bilingual language production: (1) L2 language production should require greater IC than L1 production to the extent that L2 proficiency is low (and indeed, L1 language production may become more difficult as L2 proficiency increases); (2) these effects should interact with communicative task demands (i.e., a highly vs. less demanding communicative task should limit the resources available for IC to occur); and (3) bilinguals should successfully produce language insofar that they intrinsically possess IC capacity, after accounting for L2 proficiency.

Bilingual language production studies provide some support for these predictions, although many questions remain. Consistent with the first prediction, many studies show that L2 production (which is usually the less-dominant language) is indeed more effortful than L1 production (which is usually the more dominant language). This pattern of findings arises when bilinguals produce single words in response to pictures (Linck et al., 2008; Gollan and Ferreira, 2009; Hanulová et al., 2010; Sandoval et al., 2010), and also when they produce extended speech (Towell et al., 1996). Moreover, as L2 proficiency increases, language production in a less-dominant L2 improves (Poulisse and Bongaerts, 1994; Kormos, 2006; De Jong and Wempe, 2009). For example, at high L2 proficiency levels picture-naming speed and accuracy become more similar across L2 and L1 (Costa and Caramazza, 1999; Kroll et al., 2002; Costa and Santesteban, 2004). A similar pattern of effects is also seen during spontaneous speech production. For example, increased L2 proficiency is associated with increased articulation rate, longer utterance durations, shorter and less frequent silent pauses, and a greater number of words produced in the L2 when bilinguals narrated a story from a cartoon strip (Kormos and Dénes, 2004).

Increased L2 proficiency also relates to increased L1 processing effort when bilinguals produce single words in response to a picture (Gollan et al., 2005, 2008, 2011; Ivanova and Costa, 2008), overtly name visually presented words (Flege, 1999), or to general measures of functional language ability (i.e., subtractive bilinguals Lambert, 1974). Interestingly, our group recently found that these effects of increased L2 ability on L1 processing extend to eye movement measures of reading (Titone et al., 2011; Whitford and Titone, 2012). Presumably, such effects on L1 language processing arise because bilinguals who are highly proficient in their L2 use their L2 to a great extent, and as a consequence, use their L1 relatively less. Thus, over time and repeated L2 practice and use, L1 representations grow weaker while L2 representations grow stronger.

Returning to the second prediction of the IC model, there is also evidence that L1/L2 differences in language production are sensitive to increased task demands. For example, language production is more effortful during simultaneous interpretation, in which bilinguals must understand the utterance in one language and produce it in another (Christoffels and De Groot, 2004). As well, there is preliminary evidence of task demand effects for spontaneous speech when it is produced with or without a conversational partner. For example, bilinguals produce more disfluencies (e.g., uhs, ums) when answering speculation questions during a dyadic interview (e.g., What makes an ideal friend?) than when producing speech without a conversational partner (e.g., telling a story from a picture; Fehringer and Fry, 2007). This suggests the possibility that a dialogue context may be relatively more effortful than a monologue context, especially during L2 language production. This finding is interesting in light of recent work suggesting that dialogue speech can be less effortful than monologue speech because conversational partners provide additional sources of information that can facilitate speech planning, such as immediate feedback about communication success or lexical and syntactic priming across partners (Garrod and Pickering, 2004; Hartsuiker et al., 2004; Pickering and Garrod, 2004; Costa et al., 2008; Hartsuiker and Pickering, 2008; Kootstra et al., 2010). While such facilitative interactive alignment effects are certainly possible, they are likely offset by other increased task demands of spontaneous dialogue, such as integrating language production and comprehension simultaneously, and making decisions about when to speak or listen, all within the time limits of normal conversational exchange (McFarland, 2001; Wilson and Wilson, 2005).

Finally, there is preliminary evidence consistent with the third prediction of the IC model that individual differences in inhibitory capacity modulate bilingual language production, over and above the effects of L2 proficiency. Linck et al. (2008) found that bilinguals with greater inhibitory capacity vs. those without, as assessed by non-linguistic tasks, inhibited L1 activation during L2 production more efficiently, irrespective of L2 immersion environment, L2 proficiency, or L1/L2 script similarity. However, given that Linck and colleagues investigated single word production, an open question is whether individual differences in inhibitory capacity exert similar effects when producing extended spontaneous speech and in different communicative contexts.

Thus, the purpose of the present study is to investigate several questions about bilingual language production in the domain of spontaneous monologue and dialogue speech. Based on the IC model (Green, 1998) and its extension (Abutalebi and Green, 2007), we predicted that L2 vs. L1 language production would be more effortful overall; however, increased L2 proficiency would reduce this difference (Poulisse and Bongaerts, 1994; Green, 1998; Gollan et al., 2005; Fehringer and Fry, 2007; Ivanova and Costa, 2008; Kroll et al., 2008; Linck et al., 2008). We also predicted that dialogue speech would be more effortful than monologue speech, particularly in the L2 vs. L1 context (Fehringer and Fry, 2007). Finally, we predicted that individual differences in inhibitory capacity, while accounting for L2 proficiency, would interact with the language produced (L1 vs. L2) and task demands (monologue vs. dialogue). For example, it is possible that spontaneous speech produced in the most demanding condition (L2 dialogue) would require greater IC than speech produced in the least demanding condition (L1 monologue).

To test these predictions, we recorded participants as they spontaneously produced L1 and L2 monologue and dialogue speech (each participant performed in every condition). Participants also completed a battery that assessed their L2 proficiency and inhibitory capacity. To elicit spontaneous speech, we used a modified version of the Map task (Anderson et al., 1991), which is frequently used to study spontaneous speech in the context of natural dialogues (Brown and Miller, 1980; Macafee, 1983; Macaulay, 1985). In this task, each of two conversational partners receives a map that the other cannot see. One partner is assigned the role of instruction giver, and the other of instruction follower. Each map contains a starting point and black and white drawings of landmarks, along with their word labels, that occasionally mismatch across the instruction giver and follower's map versions. Of note, the instruction giver's map has a route that must be verbally described so that the instruction follower can reconstruct the route on her own map. Because some of the landmarks mismatch across the maps, conversational partners spend time discussing these discrepancies (see Appendices A and B for examples of maps and speech output).

We modified the Map task procedure in the following ways. First, participants always served as instruction givers, and the same experimental confederate always served as the instruction follower. Second, we implemented a comparable monologue version in which participants instructed a "hypothetical" listener. Finally, all participants performed the task in their L1 and L2, with order counterbalanced across participants.

All speech output was digitally recorded and analyzed with respect to two kinds of measures: global language output measures, which provided information about the content of what was produced, and acoustic–temporal measures, which provided information about how the speech was produced in real time. Global language output measures consisted of the subjective impressions of trained raters regarding the clarity of speaker's instructions (clarity of semantic content), the fluency of the speaker (smoothness of speech, absence of interruptions, hesitations and self-repairs, and changes in speech rate), and the extent to which the speaker sounded native-like.

Acoustic–temporal measures were ascertained using software that we developed to extract from the speech recordings the number of vocalizations and their length, and the silent pause durations preceding each vocalization. We used these two indices to compute a ratio, which consisted of individual vocalization durations over their prior silent pause durations (VD/PPD) across all utterances (see Materials and Methods for further detail). We focused on the ratio between each vocalization duration and its prior silent pause duration, based on prior work suggesting that vocalization durations reflect speech output effort (Henderson et al., 1966; Goldman-Eisler, 1968; Kormos and Dénes, 2004; Kormos, 2006; Segalowitz, 2010), and that prior silent pause durations reflect speech planning effort (Lindsley, 1975; Chaffe, 1980; Levelt, 1983; Ferreira, 1991; Segalowitz, 2010). Given these findings, it stands to reason that a large ratio reflects a situation where a given vocalization is less effortful to plan than a vocalization having a small ratio. As well, examining this ratio, rather than vocalization duration or internal pause duration alone, has an advantage of standardizing any difference in vocalization durations that could arise due to within- or between-monologues or dialogues, participants, or languages.

MATERIALS AND METHODS

PARTICIPANTS

A total of 22 English–French and 20 French–English bilingual adults ($N = 42$, $M = 21.21$, $SD = 2.52$; seven males, 35 females) from McGill University (Montréal, Canada) participated for course credit. Participants were healthy young adults, 18–35 years old, with normal or corrected-to-normal vision, and no self-reported speech or hearing disorders. Originally, we recruited 64 participants (32 English–French and 32 French–English) but we excluded 22 participants (10 English–French and 12 French–English) for the following reasons. Four reported acquiring first language other than English or French (two from each group). Seven reported that L2 was currently their more-dominant language (all French–English). Seven reported on a L2 proficiency questionnaire that they would not choose to speak L2 at all (five English–French and two French–English). Three were excluded because of equipment failure during sound recording (all English–French). One participant did not complete a portion of the speech production task (French–English).

We used an adapted version of the language experience and proficiency questionnaire (LEAP-Q) to assess participants' L2 proficiency (Marian et al., 2007). At the time of testing, French–English bilingual participants reported learning French as their first language, rated it as their dominant language, and reported high proficiency in English. Similarly, English–French bilingual participants reported learning English as their first language, rated it as their dominant language, and reported high proficiency in French. For subsequent analyses, we used the rating sub-scales of the LEAP-Q to calculate a standardized L2 proficiency score, modeled after McMurray et al. (2010). **Table 1** summarizes self-assessed L2 proficiency measures.

MATERIALS

We selected four pairs of maps from the Map task corpus (http://groups.inf.ed.ac.uk/maptask/#maps). Two maps were used to elicit monologue speech for each participant, once in L1 and once in L2, and two additional maps were used to elicit L1 and L2 dialogue speech. Because the Map task corpus was created in English, we translated verbal labels into French and pasted them onto new maps.

PROCEDURE

We randomly assigned participants to one of two counterbalancing streams (see **Figure 2**). As illustrated in **Figure 2**, we counterbalanced whether the Map task was performed first in the L1 or L2 separately for English–French and French–English participants. All participants completed the monologue version of the Map task in one language, followed by the dialogue version of the Map task in the same language. Then, they completed the monologue version of the Map task in the other language, followed by the dialogue version of the Map task in the same language. Half of the participants completed the Map task in L1 first (left panel of **Figure 2**) the other half of participants completed the Map task in L2 first (right panel of **Figure 2**). Following Map task administration, all participants completed a battery that assessed their inhibitory capacity, the vocabulary subtest of the Wechsler abbreviated scale of intelligence (WASI) and a language background questionnaire. Testing session lasted approximately 2 hours.

Across all monologue and dialogue versions, participants always served as the instruction giver. In the English version of

Table 1 | Self-assessed L2 proficiency ratings, language history, and standardized L2 proficiency scores ($n = 42$).

	English–French ($n = 22$)		French–English ($n = 20$)	
	M	**SD**	***M***	**SD**
Rating scales (0–10)				
Speaking ability	7	2	8	2
Reading ability	8	2	8	1
Writing ability	7	2	7	1
Translating ability	7	2	8	2
Listening comprehension	8	2	8	1
Pronunciation	7	2	7	2
Fluency	7	2	7	2
Vocabulary	7	2	7	2
Grammatical ability	7	2	7	1
Overall competence	7	2	8	1
Sum of rating scales (0–100)	71	19	76	13
Standardized L2 proficiency score	−0.03	0.93	0.27	0.67
Age of acquisition (years old)				
Began acquiring L2*	5	2	7	4
Became competent in L2	10	4	12	5
Choose to speak L2 (%)**	17	12	34	21
Degree of L1 interference when speaking in L2 (0–5)**	2	1	3	1
Percent of present time spent functioning in each language				
L1***	82	9	48	17
L2***	14	7	50	17

**Two-tailed independent samples t-test significant at $p < 0.05$.*

***Two-tailed independent samples t-test significant at $p < 0.01$.*

****Two-tailed independent samples t-test significant at $p < 0.001$.*

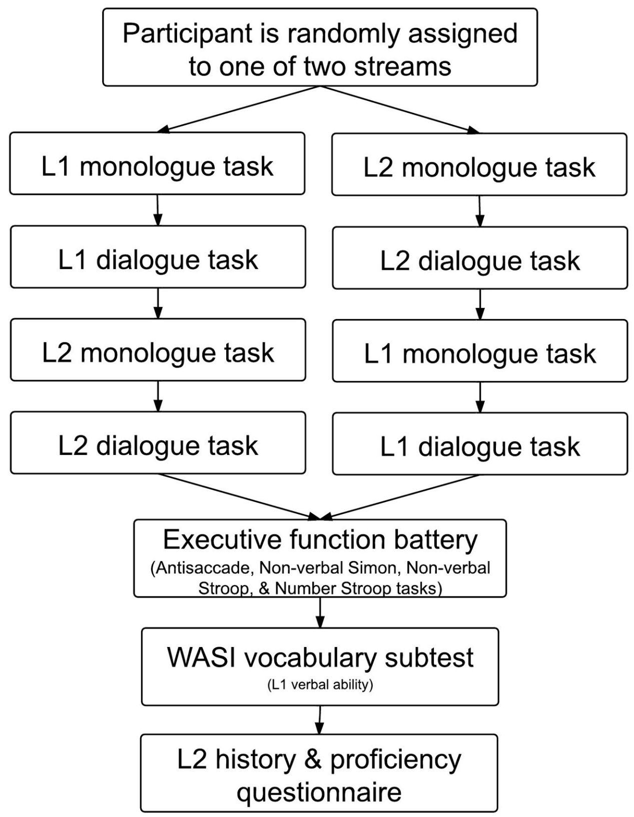


FIGURE 2 | Illustration of the procedure in the current study.

the Map task, we instructed participants in English to verbally guide the instruction follower through a printed route from start to finish, in English. In the French version of the Map task, we instructed participants in French to verbally guide the instruction follower, in French. In the monologue versions, we instructed participants to guide an imaginary person. In the dialogue versions, we instructed participants to guide their conversational partner, a confederate of the experiment. In the dialogue versions, the confederate reproduced the route on her version of the map, based strictly on the instructions of the participants. When participants and the confederate encountered discrepancies in labels across their versions of maps, unknown to the participant, the confederate was required to exclusively refer to the landmarks by the labels printed on her map.

Participants and the confederate performed the Map task in the same room. Participants and the confederate were instructed to speak at a normal rate, and faced away from each other to prevent gaze and posture coordination (Shockley et al., 2009). During monologues and dialogues, participants viewed maps on a 20″ monitor located 71 cm away from where they were seated. Participants wore an AKG C420 PP MicroMic Series III headset microphone, while we used a Zoom H4 Handy Recorder to record their speech at 44 kHz in stereo, such that participants' voice was acoustically isolated to the left channel and the confederate's to the right channel.

INDIVIDUAL DIFFERENCES MEASURES

To assess individual differences in inhibitory capacity, we administered an anti-saccade task (Hallett, 1978), a non-linguistic Simon (Simon and Ruddell, 1967), and Stroop (Stroop, 1935) tasks modeled after Blumenfeld and Marian (2011) and a Number Stroop task. To assess L1 verbal ability, we administered vocabulary subtest of the WASI (Wechsler, 1999).

Anti-saccade task

This task assessed ability to inhibit the pre-potent tendency to look toward a peripherally presented target (Hallett, 1978). We used an Eye-Link 1000 tower mounted system (SR-Research, ON, Canada) with a sampling rate of 1 kHz to monitor and record fixation durations of the right eye. Participants were presented randomly intermixed pro-saccade and anti-saccade trials. At the onset of each trial, participants saw a small black fixation circle in the center of the computer screen, followed by a central fixation square that remained on the screen for 1000, 1250, or 1500 ms. The central fixation square was green to cue participants to engage in a pro-saccade trial, and red to cue participants to engage in an anti-saccade trial. Thus, contingent on the color of the central fixation square, participants looked toward (pro-saccade trials) or away (anti-saccade trials) from peripherally located black square targets. We computed an *Anti-saccade Cost* variable for each participant based on correct trials only (Bialystok et al., 2006), where we subtracted the average reaction time of all pro-saccade trials from the average reaction time of all anti-saccade trials.

Non-linguistic simon and stroop tasks

We adapted these tasks from Blumenfeld and Marian (2011). Participants saw arrows on a screen. In the Simon task, the arrows pointed up or down. When the arrows pointed up, participants used their left hand to press a response button on the left, and when the arrows pointed down, participants used their right hand to press a response button on the right. Trials were congruent when the arrow appeared on the same side of the computer screen as the response and incongruent when the arrow appeared on the opposite side of the computer screen as the response. The Simon effect reflects the finding that participants execute a motor response more quickly and accurately when the left/right spatial location of the stimulus corresponds to the left/right spatial location of the response button (Simon and Ruddell, 1967). In the Stroop task, the arrows pointed left or right. When the arrows pointed left, participants used their left hand to press a response button on the left, and when the arrows pointed right, participants used their right hand to press a response button on the right. Trials were congruent when the arrow appeared on the same side as its pointed direction and incongruent when the arrow appeared on the opposite side as its pointed direction. The Stoop effect reflects the finding that participants execute a motor response more quickly and accurately when the semantic meaning of the stimulus corresponds to the required response (Stroop, 1935). We computed a cost score for the Simon and Stroop tasks separately, in which we subtracted the average reaction time on congruent trials from the average reaction time on incongruent trials. Only correct trials were included in these averages.

Number stroop task

This task also assessed the ability to inhibit a strong automatic cognitive response. We presented a series of numbers ranging from one to four digits on a computer screen. Participants were instructed to use their dominant hand to press one of four response buttons that corresponded to the number of digits appearing on the screen. Trials were congruent when the quantity of digits corresponded to the depicted numbers (22 required response 2) and incongruent when the quantity of digits did not correspond to the depicted numbers (e.g., 222 required response 3). We computed a cost score for the correct reaction times for each participant by subtracting the average reaction time on congruent trials from the average reaction time on incongruent trials.

Descriptive statistics from each task appear in **Table 2**. Two-tailed independent samples t-tests revealed that performance did not significantly differ between English–French and French–English participants on all tasks ($p > 0.05$). Using these measures of inhibitory capacity, we computed a standardized composite inhibition cost score (McMurray et al., 2010).

WASI vocabulary subtest

Participants defined words in L1, which we scored and transformed into scaled score using age-appropriate norms.

RESULTS

We constructed a series of linear mixed effect (LME) models, as implemented in the lme4 library (Bates, 2005) in R Project for Statistical Computing version 2.10.1 (Baayen, 2008; Baayen et al., 2008; R Development Core Team, 2009). The models included as variables of interest the main effects and interactions of language (L1 vs. L2), speech type (monologue vs. dialogue), L2 proficiency (continuous), and inhibitory capacity score (continuous). All models had random intercepts for items (i.e., number of different maps) and participants (Baayen, 2008). All models had language group (English–French vs. French–English) as control variable to account for L2 vs. L1 linguistic differences between two groups. We excluded L1 verbal ability (WASI scaled scores) from the models reported below because there was only one instance where it accounted for a significant amount of variance. This was in the clarity of instructions measure (see below), where increased verbal ability was associated with higher ratings. Our dependent variables consisted of the global output measures and acoustic–temporal measures previously described. We first report the results for the global output measures, followed by results for the acoustic–temporal measures. Within each set of analyses, we first report the analyses that assess the contribution of L2 proficiency, followed by analyses that assess the added contribution of inhibitory capacity.

Table 2 | Minima, maxima, means, and SDs for individual difference measures.

	Min	Max	*M*	SD
L1 VERBAL ABILITY				
WASI score	8	18	14	3
SIMON TASK				
Congruent	376	660	489	65
Incongruent	424	728	527	73
Cost	−48	−68	-38^{***}	
STROOP TASK				
Congruent	275	844	480	108
Incongruent	211	783	510	106
Cost	64	61	-30^{***}	
NUMBER STROOP TASK				
Congruent	446	717	582	72
Incongruent	475	833	652	87
Cost	−29	−116	-70^{***}	
ANTI-SACCADE TASK				
Pro-saccade	230	415	304	47
Anti-saccade	341	528	411	45
Cost	−111	−113	-107^{***}	
INHIBITION COST SCORE				
English–French	−0.0099	0.0087	−0.0002	0.0048
French–English	−0.0070	0.0088	0.0002	0.0037

****Two-tailed paired samples t-test significant at $p < 0.001$.*

GLOBAL OUTPUT MEASURES

Global output measures included the clarity of speaker's instructions and speaker fluency and nativeness. We selected and adapted these measures from the work of (Pinkham and Penn, 2006). To obtain these measures a team of independent raters (two native-English and two native-French) coded participants' speech files separately in monologues and dialogues and in L1 and L2. For each monologue or dialogue recording, the independent raters assigned a score from one to nine on the following dimensions, the clarity of speaker's instructions and speaker fluency and nativeness. Raters were trained on 20 English and French speech samples; however, they coded only speech samples that matched their native language. Interrater reliability on the training samples was high (Cronbach's alpha = 0.93). Descriptive statistics for each dimension are shown in **Table 3**.

L2 proficiency and the clarity of instructions

Table 4 presents the results of LME models for clarity of instructions. The clarity of instructions was lower in L2 ($M = 6.67$) than L1 speech ($M = 7.06$), resulting in a significant main effect of language ($t = -2.02$, $p < 0.05$). As well, the clarity of instructions was lower in monologues ($M = 6.34$) than dialogues ($M = 7.39$), resulting in a significant main effect of speech type ($t = 2.96$, $p < 0.01$). Finally, the clarity of instructions varied with the language of production and L2 proficiency, resulting in a significant two-way interaction between language and L2 proficiency ($t = 2.09$, $p < 0.05$). This interaction is depicted in **Figure 3**. The left panel of **Figure 3** shows that the clarity of instructions in monologues was significantly lower in L2 than in L1 speech for bilinguals with low L2 proficiency. Moreover, the L2 vs. L1 difference in the clarity of instructions decreased as L2 proficiency increased. Finally, the right panel of **Figure 3** shows that the clarity of instructions did not differ between L1 and L2 across all levels of L2 proficiency in dialogues.

Table 3 | Mean ratings and SEs of the mean/or clarity of instructions, speaker fluency, and nativeness for monologues and dialogues in L1 and in L2.

		Language	*M*	SE
Clarity of instructions	Monologue	L1	6.62	0.26
		L2	6.05	0.33
	Dialogue	L1	7.49	0.18
		L2	7.28	0.18
Speaker fluency	Monologue	L1	7.52	0.17
		L2	7.74	0.15
	Dialogue	L1	7.84	0.16
		L2	6.58	0.22
Speaker nativeness	Monologue	L1	8.86	0.05
		L2	6.98	0.25
	Dialogue	L1	8.88	0.05
		L2	6.30	0.30

Table 4 | Linear mixed effects models for global output measures (clarity of instructions, speaker fluency, and nativeness) to illustrate interactions between speech type, language, and L2 proficiency.

	Clarity of instructions			Speaker fluency			Speaker nativeness		
	b	SE	*t*-Value	*b*	SE	*t*-Value	*b*	SE	*t*-Value
Fixed effects									
Intercept	6.34	0.25	25.08***	7.32	0.21	35.77***	8.55	0.22	39.62***
Speech type (monologue, dialogue)[1]	0.95	0.32	2.96**	0.30	0.22	1.32	0.02	0.26	0.08
Language (L1, L2)[2]	−0.65	32	−2.02*	0.19	0.22	0.86	−2.02	0.22	−9.04***
L2 proficiency	0.59	0.28	1.75	0.15	0.22	0.67	−0.06	0.21	−0.26
Language group[3] (English–French vs. French–English)	0.35	0.23	1.53	0.42	0.22	1.91	0.65	0.20	3.22**
Speech type × language	0.41	0.45	0.90	−1.45	0.32	−4.59***	−0.69	0.32	−2. 18*
Speech type × L2 proficiency	−0.30	0.39	−0.77	−0.03	0.27	−0.11	0.11	0.27	0.40
Language × L2 proficiency	0.82	0.39	2.09*	−0.18	0.27	−0.65	1.05	0.27	3.84**
Speech type × language × L2 proficiency	−0.81	0.55	−1.46	0.37	0.39	0.96	−0.08	0.39	−0.20
Random effects		Variance			Variance			Variance	
Subject		0.00			0.23			0.16	
Item		0.00			0.00			0.02	
Residual		2.10			1.03			1.01	

**pMCMC < 0.05 level, **pMCMC < 0.01 level, ***pMCMC < 0.001 level.*

[1]Baseline = monologue.

[2]Baseline = L1.

[3]Baseline = English–French.

L2 proficiency and speaker fluency

Speaker fluency varied as a function of language of production (L1 vs. L2) and speech type (monologue vs. dialogue), resulting in a significant two-way interaction between language and speech type ($t = -4.59$, $p < 0.001$). As shown in **Table 3**, speaker fluency was lowest in L2 dialogues ($M = 6.58$) as compared to L1 monologues ($M = 7.52$), L2 monologues ($M = 7.74$), and L1 dialogues ($M = 7.84$), the latter of which did not differ. This interaction is depicted in **Figure 4**. The left panel of **Figure 4** shows that speaker fluency did not significantly differ between L1 and L2 monologues. The right panel of **Figure 4** shows that speaker fluency was significantly lower in L2 than L1 dialogues. L2 proficiency did not significantly predict speaker fluency for L1 vs. L2 and monologues vs. dialogues.

L2 proficiency and speaker nativeness

Speaker nativeness was lower in L2 ($M = 6.64$) than L1 speech ($M = 8.87$), resulting in a significant main effect of language ($t = -9.04$, $p < 0.001$). Speaker nativeness also varied as a function of the language of production (L1 vs. L2) and speech type (monologue vs. dialogue), resulting in a significant two-way interaction between language and speech type ($t = -2.18$, $p < 0.05$). As shown in **Table 3**, speaker nativeness was lowest in L2 dialogues ($M = 6.30$) followed by L2 monologues ($M = 6.98$) and highest in L1 monologues ($M = 8.86$) and L1 dialogues ($M = 8.88$). Finally, speaker nativeness varied as a function of language of production and L2 proficiency, resulting in a significant two-way interaction between language and L2 proficiency ($t = 3.84$, $p < 0.01$). **Figure 5** shows this interaction across left and right

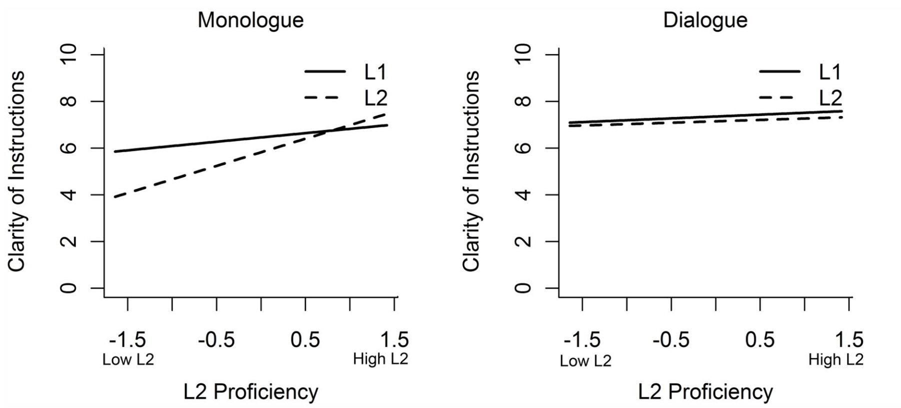


FIGURE 3 | Graphical representation of partialed effects from model fits of L2 proficiency on the clarity of instructions in L1 and in L2 across monologues (left panel) and dialogues (right panel). Clarity of instructions was significantly lower in L2 vs. L1 for low L2 proficient bilinguals in monologues. Clarity of instructions was not different for high L2 proficient bilinguals in monologues and for bilinguals of all L2 proficiency levels in dialogues.

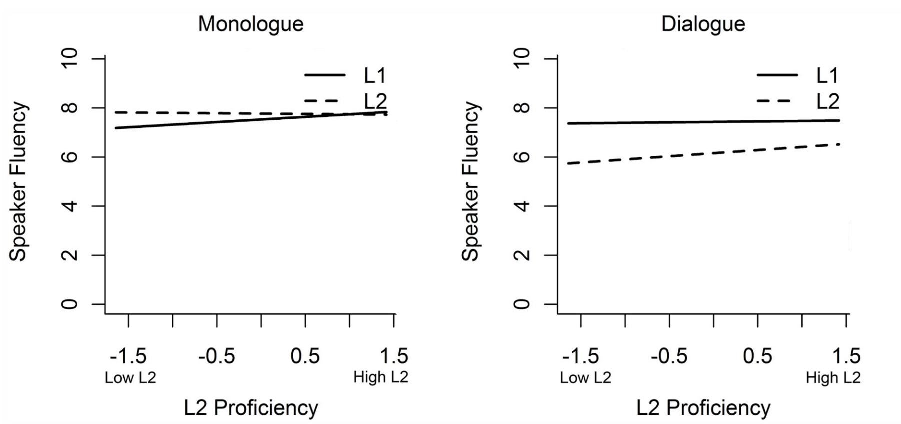


FIGURE 4 | Graphical representation of partialed effects from model fits of L2 proficiency on the speaker fluency in L1 and in L2 across monologues (left panel) and dialogues (right panel). Speaker fluency was lower in L2 than in L1 but only in dialogues. Speaker fluency was not different in L2 vs. L1 in monologues.

panels. Speaker nativeness for L1 monologues and dialogues was high across all levels of L2 proficiency. Conversely, speaker nativeness for L2 monologues and dialogues varied as a function of L2 proficiency. Bilinguals with low L2 proficiency showed lower speaker nativeness than bilinguals with high L2 proficiency.

Inhibitory capacity and clarity of instructions, speaker fluency, and nativeness

To assess whether individual differences in inhibitory capacity modulated global output measures, we included the composite inhibition cost score as a fixed effect to the models previously described. Thus, we constructed models with four-way interactions between language, speech type, L2 proficiency and inhibition cost score for clarity of instructions, speaker fluency, and speaker nativeness. Within these final models, inhibitory capacity did not significantly relate to any of the global output measures, neither as the main effect nor as part of the higher-order interactions (all $ts < 1.53$, $p > 0.05$).

ACOUSTIC–TEMPORAL MEASURES OF SPEECH PRODUCTION

The acoustic–temporal measure of interest was the ratio of individual vocalization durations over their prior silent pause durations (VD/PPD). Again, we assumed greater ratios reflect increased efficiency of speech planning. First, we describe how we processed speech files to compute this measure.

Pre-processing of speech files

To minimize cross-talk between conversational partners, we recorded speech at a relatively low volume. Thus, prior to analysis, we amplified the speech signal by 26 dB and removed inaudible speech below 40 dB. We used Soundforge (version 8.0, Sony Creative Software) to standardize the amplitude of the speech signal across monologues and dialogues, and to remove all instances of coughs and laughs. After this pre-processing stage, we used custom software to distinguish periods of vocalization from periods of silence for each speaker, based on prior work (Alpert et al., 1986; Welkowitz et al., 1990). For the purpose of this study, we

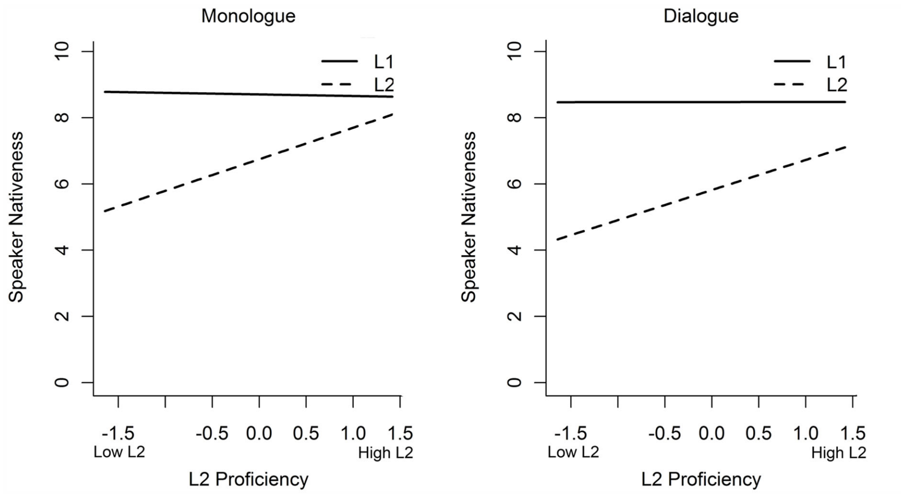


FIGURE 5 | Graphical representation of partialed effects from model fits of L2 proficiency on the speaker nativeness in L1 and in L2 across monologues (left panel) and dialogues (right panel). Speaker nativeness was lower in L2 vs. L1 in dialogues vs. monologues. Speaker nativeness was lowest in L2 vs. L1 for low L2 proficiency bilinguals but L2 vs. L1 difference decreased for high L2 proficient bilinguals.

only selected instances where silent pause preceded a vocalization duration uttered by the participant (see Appendix C). Independent periods of vocalization were registered when the speaker signal exceeded minimum amplitude for at least 250 ms. Periods of silence were registered when the speaker signal remained below minimum amplitude for at least 250 ms. These timing parameter estimates were based on prior work using similar automated speech processing methods and other studies of spontaneous speech (Goldman-Eisler, 1968; Alpert et al., 1986; Welkowitz et al., 1990; Wilson and Wilson, 2005; Kormos, 2006; Segalowitz, 2010). Initial silences (prior to the initial vocalization or following the final vocalization) and silences less than 250 ms were removed from estimates of the mean vocalization durations. Descriptive statistics for vocalization and silent pause durations are shown in **Table 5**.

Of note, our custom software also identifies in the speech signal switching pauses, turn-taking boundaries, and strong and weak interruptions. While these are important features of dialogue speech, we excluded them from the calculations of ratios to enable direct comparison of dialogue and monologue speech, the latter of which lacks these features.

L2 proficiency and the ratio of vocalization durations to prior pause durations (VD/PPD)

VD/PDD ratios were smaller for L2 ($M = 3.33$) vs. L1 speech ($M = 3.86$), resulting in a main effect of language ($t = -3.93$, $p < 0.001$). VD/PPD ratios were smaller in dialogues ($M = 3.33$) than monologues ($M = 3.86$), resulting in a main effect of speech type ($t = -3.48$, $p < 0.001$). Finally, VD/PPD ratios varied as a function of language of production, speech type, and L2 proficiency. This resulted in a significant three-way interaction between speech type, language, and L2 proficiency ($t = -2.60$, $p < 0.05$), shown in **Figure 6**. The left panel of **Figure 6** shows that VD/PPD ratios for monologues were smaller in L2 than L1 for low L2 proficiency bilinguals. However, the L2 vs. L1 difference in VD/PPD ratios for monologues decreased as L2 proficiency increased. In particular, as L2 proficiency increased, it appears that L2 VD/PPD ratios also increased while L1 VD/PPD rations decreased. In contrast to monologues, there was no effect of L2 proficiency for dialogues. The right panel of **Figure 6** shows that VD/PPD ratios were smaller for L2 vs. L1 speech, regardless of L2 proficiency.

Inhibitory capacity and VD/PPD ratios

To investigate whether individual differences in inhibitory capacity relate to monologue and dialogue speech production, we added as a fixed effect the composite inhibition cost score to the three-way interaction (language × speech type × L2 proficiency) of the model just presented. **Table 6** presents the results of this LME model. There was again a significant three-way interaction between language, speech type, and L2 proficiency, but no four-way interaction with inhibitory capacity ($t = -0.85$, $p > 0.05$). However, VD/PPD ratios decreased as inhibitory capacity decreased (inhibition cost increased), resulting in a main effect of inhibitory capacity ($t = -2.20$, $p < 0.05$). As well, inhibitory capacity interacted with speech type and L2 proficiency, resulting in a significant three-way interaction between speech type, L2 proficiency, and inhibitory capacity ($t = 2.70$, $p < 0.01$). This interaction is shown in **Figure 7**. As seen in the upper and lower left panels of **Figure 7**, VD/PPD ratio increased as both L2 proficiency and inhibitory capacity increased. In contrast, as seen in the upper right panel of **Figure 7**, VD/PPD ratios did not significantly vary for L1 dialogues as a function of L2 proficiency or inhibitory capacity. Finally, as seen in the lower right panel of **Figure 7**, VD/PPD ratios again increased as both L2 proficiency and inhibitory capacity increased.

DISCUSSION

We investigated how individual differences in L2 proficiency and inhibitory capacity relate to bilinguals' spontaneous monologue and dialogue language production. There were several key findings pertaining to the role of L2 proficiency, task demands, and inhibitory capacity.

Table 5 | Mean values (ms), SEs of the mean, and mean observation count for vocalization durations, prior silent pause durations, computed VD/PPD ratios, and total speech sample duration for monologues and dialogues in L1 and in L2.

		Language	Mean	SE	Mean observation count
Vocalization duration	Monologue	L1	2272	280	32.30
		L2	1959	274	40.72
	Dialogue	L1	1624	237	52.14
		L2	1422	225	66.33
Prior silent pause duration	Monologue	L1	689	58	
		L2	698	61	
	Dialogue	L1	584	63	
		L2	596	59	
VD/PPD ratios	Monologue	L1	4.14	0.64	
		L2	3.58	0.60	
	Dialogue	L1	3.58	0.61	
		L2	3.08	0.56	
Total sample duration	Monologue	L1	151169	9292	
		L2	154476	10775	
	Dialogue	L1	494494	27188	
		L2	468151	21076	

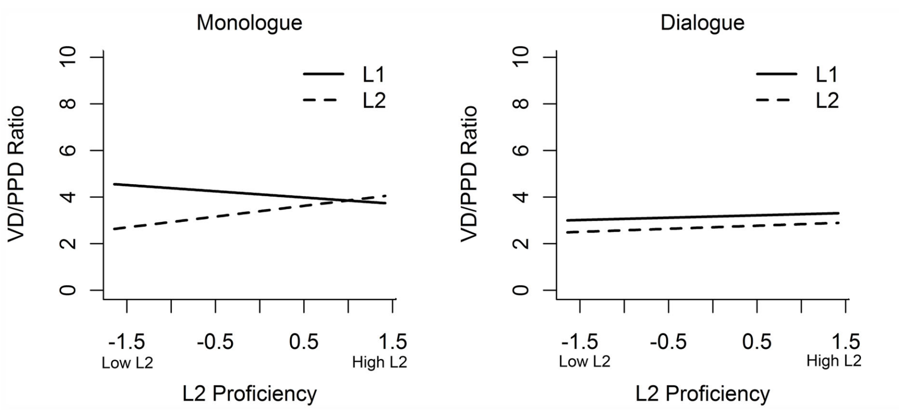


FIGURE 6 | Graphical representation of partialed effects from model fits of L2 proficiency on the VD/PPD ratios in L1 and in L2 across monologues (left panel) and dialogues (right panel). Speech planning and production was lower in L2 vs. L1 for low L2 proficient bilinguals in monologues and across all L2 proficiency levels in dialogues. Speech planning and production was not different in L2 vs. L1 for high L2 proficient bilinguals in monologues.

Consider first the results for the global output measures. The clarity of instructions produced was higher when people spoke in their L1 than their L2, although increased L2 proficiency helped to close the gap between L1 and L2 clarity. Dialogue speech also was rated as clearer in content than monologue speech, which is consistent with recent work suggesting that dialogue speech is easier to produce than monologue speech and also that the goal of dialogue is to relay the message clearly to a conversational partner (Garrod and Pickering, 2004; Hartsuiker et al., 2004; Pickering and Garrod, 2004; Hartsuiker and Pickering, 2008; Kootstra et al., 2010). Specifically, this finding suggests that the presence of a conversational partner was associated with enriched semantic content during language production, presumably because the conversational partner provided the speaker with ongoing feedback about when the content of their output was unclear.

The other two global output measures behaved somewhat differently from the clarity of instructions. Speech fluency (whether people spoke in a fluid or halting way) was generally high except for L2 dialogue speech, which is arguably the most cognitively demanding of the different language production conditions. This effect of speech fluency was unaffected by differences in L2 proficiency. Speaker nativeness, in contrast, was influenced by several factors: L2 proficiency, the language of speech, and whether a monologue or a dialogue was produced. L2 speech was rated as less native-like than L1 speech, and this difference was larger for dialogue than monologue speech. Finally, the difference between L1 and L2 speaker nativeness also decreased as L2 proficiency increased.

Taken together, the global output measures suggest that language knowledge (whether L1 or L2 production is adjusted

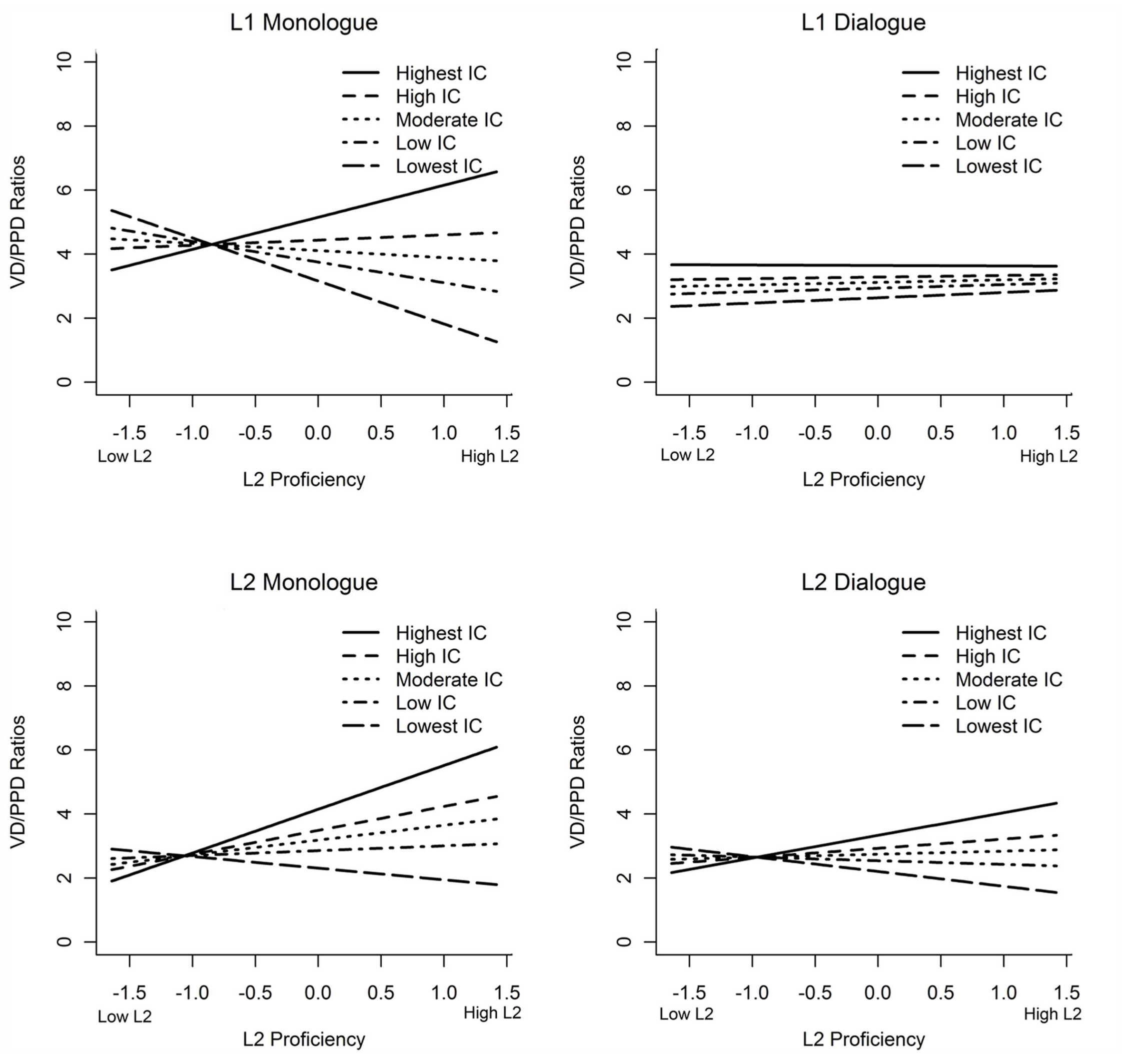


FIGURE 7 | Graphical representation of partialed effects from model fits of an interaction between inhibitory capacity, L2 proficiency, and speech type (monologue vs. dialogue; left vs. right panels) on the VD/PPD ratios in L1 and in L2 (upper vs. lower panels). Speech planning and production was more efficient for bilinguals with high vs. low inhibitory capacity at high L2 proficiency levels in monologues. Speech planning and production is more efficient for bilinguals with high vs. low inhibitory capacity at high L2 proficiency levels in L2 dialogues. Speech planning and production is not different across all inhibitory capacity levels at low L2 proficiency levels in monologues and dialogues.

by individual differences in L2 proficiency) and task demands (whether people produce speech in a monologue or a dialogue) modulate the substance of what is produced during spontaneous monologue or dialogue speech. Absent here are any effects arising from individual differences in inhibitory capacity. This is potentially surprising given the IC model's focus on inhibition as a critical mechanism for bilingual language processing. However, it is possible that global measures of language production output are not the most appropriate level of analysis to observe an effect of inhibitory capacity. Rather, as clearly implied by the IC model, inhibition may have more local effects on the ongoing planning of individual vocalizations.

Indeed, we found clear evidence that the acoustic–temporal measures showed sensitivity to individual differences in inhibitory capacity. Recall, our primary acoustic–temporal measure was the ratio between the duration of each vocalization and the duration of its prior pause (VD/PPD). Prior work suggests that there is a close linkage between the planning that takes place prior to a vocalization, and the nature of what is produced (Lindsley, 1975; Chaffe, 1980; Levelt, 1983; Ferreira, 1991; Segalowitz, 2010). Thus, a large value for this ratio should indicate that a speaker produced a given vocalization with relatively little planning effort. In contrast, a small value for this ratio should indicate that a speaker produced a given vocalization with relatively more planning effort. Consistent with our findings for the global output measures, monologues had higher ratios than dialogues. L1 speech also had higher ratios than L2 speech, although increased L2 proficiency reduced this difference overall. Unlike monologues, dialogues had more uniform ratios, as seen in **Figure 6**.

Table 6 | Linear mixed effects models for the temporal measure.

	VD/PPD ratios		
	b	**SE**	***t*-Value**
Fixed effects			
Intercept	3.85	0.29	13.16**
Speech type (monologue, dialogue)[1]	−0.60	0.17	−3.48***
Language (L1, L2)[2]	−0.72	0.18	−3.93***
L2 proficiency	−0.42	0.28	−1.48
Inhibitory capacity	−119.51	54.29	−2.20*
Language group[3] (English–French vs. French–English)	0.54	0.39	1.39
Speech type × language	0.20	0.23	0.87
Speech type × L2 proficiency	0.50	0.21	2.38*
Language × L2 proficiency	0.71	0.22	3.25**
Speech type × inhibitory capacity	68.83	40.91	1.68
Language × inhibitory capacity	−13.80	42.82	−0.32
L2 proficiency × inhibitory capacity	−120.38	68.69	−1.75
Speech type × language × L2 proficiency	−0.71	0.28	−2.60*
Speech type × language × inhibitory capacity	9.80	52.29	0.19
Speech type × L2 proficiency × inhibitory capacity	136.17	50.54	2.7**
Language × L2 proficiency × inhibitory capacity	−17.54	53.69	−0.33
Speech type × language × L2 proficiency × inhibitory capacity	−55.43	64.88	−0.85
Random effects		**Variance**	
Subject		1.30	
Item		0.00	
Residual		12.12	

VD/PPD ratios (ratio of vocalization durations over their prior silent pause durations) to illustrate interactions between speech type, language, L2 proficiency, and inhibition capacity.

**pMCMC < 0.05 level, **pMCMC < 0.01 level, ***pMCMC < 0.001 level.*

[1]Baseline = monologue.

[2]Baseline = L1.

[3]Baseline = English–French.

However, individual differences in inhibitory capacity also modulated VD/PPD ratios for monologues and dialogues. For monologues, increased inhibitory capacity appears to have blocked for L1 monologues the apparent decline associated with increased L2 proficiency. At the same time, increased inhibitory capacity appears to have enhanced the apparent growth associated with increased L2 proficiency (left panel of **Figure 7**). For dialogues, in contrast, increased inhibitory capacity seems to have facilitated overall VD/PPD ratios when people conversed in their L1. Increased inhibitory capacity also seems to have facilitated VD/PPD ratios when people who are high in L2 proficiency conversed in their L2 (right panel of **Figure 7**).

Thus, it appears that high L2 proficient bilinguals may expend more local effort at each vocalization in their L1 to maintain a high level of L1 global output clarity. In contrast, it appears that high L2 proficient bilinguals may expend more local effort at each vocalization in their L2, and at the same time the global clarity is significantly reduced. Finally, bilinguals who have greater inhibitory capacity produce language more efficiently at the level of individual vocalizations, over and above the effects of L2 proficiency, as a function of communicative task demands.

These results are consistent with prior work showing that speech production is more effortful in a less-dominant language (Hernandez et al., 2000; Kormos and Dénes, 2004; Fehringer and Fry, 2007; Gollan and Ferreira, 2009; Hanulová et al., 2010; Sandoval et al., 2010), and that L2 proficiency is an important determinant of L1 and L2 production performance (Poulisse and Bongaerts, 1994; Costa and Caramazza, 1999; Kormos and Dénes, 2004; Gollan et al., 2005, 2008, 2011; Ivanova and Costa, 2008). Such effects of increased L2 proficiency on both L2 and L1 production are consistent with the IC model, according to which L2 production should be more controlled and effortful than L1 production, especially when L2 proficiency is low (see also Segalowitz, 2010). Presumably, however, as L2 proficiency increases, L2 production becomes relatively more routine and less effortful, while L1 production may become relatively less so (Abutalebi and Green, 2007).

Another key finding was that dialogue speech appeared to be more effortful than monologue speech across several measures, especially during L2 production. Specifically, dialogue speech was less fluent and native-like, and required more effort to produce at the individual vocalization level, consistent with prior work (Fehringer and Fry, 2007). Interestingly, the semantic clarity of what was produced in the L2 was greater for dialogues than monologues, presumably because speakers had the opportunity to better monitor their output through feedback from their conversational partner. In this way, our results are also consistent with prior work suggesting that dialogue speech production may be easier than monologue speech production due to interactive alignment processes (Garrod and Pickering, 2004; Hartsuiker et al., 2004; Pickering and Garrod, 2004; Hartsuiker and Pickering, 2008; Kootstra et al., 2010).

Our final key finding was that individual differences in inhibitory capacity modulated bilingual language production at the level of individual vocalizations and this interacted with communicative task demands. Specifically, bilinguals with higher inhibitory capacity were more efficient in planning and producing individual vocalizations than bilinguals with lower inhibitory capacity, particularly for monologue speech. In contrast, dialogue speech was generally more effortful overall. These findings are consistent with prior work showing that bilinguals with increased inhibitory capacity inhibit L1 during L2 production more efficiently than bilinguals with decreased inhibitory capacity, irrespective of L2 proficiency (Linck et al., 2008). Thus, consistent with the IC model, these findings suggest that increased L2 proficiency and inhibitory capacity are necessary for efficient bilingual speech planning and production.

While the results of this study improve our understanding of bilingual language production, there are several potential limitations that would be important to address in future work.

One potential limitation is that our particular use of the map task, where objects on the maps contained verbal labels, may have created a relatively low-demand communicative situation that underestimated the normal challenges of spontaneous monologue and dialogue production. Thus, the effects of inhibitory capacity observed in this study might have been even more pronounced had we used a more demanding communicative task to elicit spontaneous speech. There are several features of our task that may have made it less demanding than expected: verbal labels on the maps; a single experienced confederate rather than a completely naïve conversational partner; the fact that dialogues always followed monologues may have preferentially advantaged dialogues over monologues. Regarding this latter point, however, there was little evidence of a dialogue advantage for any measure except the clarity of instructions.

In contrast, it is also possible that our dialogue speech condition may have been more demanding than normal because of the following. First, the confederate could interrupt the participant when encountering mismatches in map landmarks in dialogues. While no such mismatches were encountered during monologue speech, future work could assess whether presence vs. absence of mismatches in map landmarks in dialogues contributes to task difficulty. Second, participants and the confederate faced away from each other, thereby blocking any visual cues during conversational interaction. Given that conversational partners communicate more easily when the visual channel is available throughout dialogue speech (Doherty-Sneddon et al., 1997), it is possible that L1 and L2 dialogue speech may become less effortful when conversational partners can see each other as they speak. Thus, the results here for dialogue only generalize to auditory-only dialogue processes, such as when two people converse by telephone.

Another potential limitation is that it is possible that the dialogue speech condition had smaller VD/PPD ratios because of a higher likelihood of dialogues having shorter vocalizations than monologues. While it is possible that the ratio measure is compressed for dialogues vs. monologues because of the higher likelihood of shorter vocalizations for dialogues, we believe that the ratio measure has information to offer regarding the ease of language production in our study for several reasons. First, the conversation task used is one where longer turns are entirely appropriate to the extent that the content of what is produced is useful (i.e., having one person describe to another person where to go on a map). In this way, our communication task differs from normal conversation where there may not be as concrete a goal or topic, and interchanges may be more rapid and short. Second, the behavior of the ratio for dialogues alone shows that it responds in expected ways as a function of our independent variables, and in a similar way to monologues. Indeed, when we perform LME analyses on the dialogues alone, we find a significant three-way interaction (language × L2 proficiency × inhibitory capacity interaction, $t = -2.20$, $p < 0.05$), suggesting that greater inhibitory capacity is associated with higher ratios for high L2 proficient bilinguals during L2 dialogues (see right panels of **Figure 7**). This effect is compatible with the monologue data where ratios were also higher as inhibitory capacity and L2 proficiency increased.

A final potential limitation concerns the independence of L2 proficiency and inhibitory capacity. Given prior work suggesting that bilinguals have better inhibitory capacity than monolinguals (reviewed in Bialystok, 2010), it is possible that bilinguals with high L2 proficiency might have greater inhibitory capacity than bilinguals with low L2 proficiency, by definition. This, in turn, would complicate our interpretation of the results for each variable individually. However, contrary to this hypothesis, the correlation between L2 proficiency and inhibitory capacity in our sample was not significant ($r = -0.16$, $p = 0.31$), perhaps due to the fact that all of the bilinguals tested here had some minimal high level of L2 proficiency to be able to produce spontaneous speech in an L2 monologue or dialogue context. As well, even presuming a statistically reliable relationship between L2 proficiency and IC, the LME approach would have allowed us to statistically disentangle the relative contributions of each to some extent, as these two variables are not likely to be perfectly correlated.

To conclude, the findings reported here suggest that individual differences among bilinguals in L2 proficiency and inhibitory capacity significantly modulate bilingual language production in monologues and dialogues, consistent with predictions of the IC model (Green, 1998; Abutalebi and Green, 2007) and prior work using other production tasks (Linck et al., 2008). Thus, our results establish a link between inhibitory capacity and bilingual language production among bilinguals, which is consistent with recent views suggesting that being bilingual enhances cognitive function (Bialystok et al., 2004; Bialystok, 2009). Finally, this study represents a first attempt at developing semi-automated methods to investigate the temporal dynamics of bilingual language production during more naturalistic conditions, such as during spontaneous monologue and dialogue speech.

ACKNOWLEDGMENTS

This research was supported by an NSERC Discovery Award (PI: Titone). The authors gratefully acknowledge additional support from CIHR for the Frederick Banting and Charles Best CGS, Master's Award (awarded to Pivneva), and the Canada Research Chairs Program (awarded to Palmer and Titone).

REFERENCES

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

Alpert, M., Homel, P., Merewether, F., Martz, J., and Lomask, M. (1986). Voxcom: a system for analyzing natural speech in real time. *Behav. Res. Methods* 18, 267–272.

Anderson, A. H., Bader, M., Bard, E. G., Boyle, E. H., Doherty, G. M., Garrod, S. C., Isard, S. D., Kowtko, J. C., McAllister, J. M., Miller, J., Sotillo, C. F., Thompson, H. S., and Weinert, R. (1991). The HCRC map task corpus. *Lang. Speech* 34, 351–366.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412.

Bates, D. M. (2005). Fitting linear models in R: Using the lme4 package. *R News* 5, 27–30.

Bialystok, E. (2001). *Bilingualism in Development: Language, Literacy, and Cognition*. New York: Cambridge University Press.

Bialystok, E. (2009). "Effects of bilingualism on cognitive and linguistic performance across the lifespan," in *Streitfall Zweisprachigkeit – The Bilingualism Controversy,* eds I. Gogolin and U. Neumann (Wiesbaden: VS Verlag für Sozialwissenschaften), 53–67.

Bialystok, E. (2010). Bilingualism. *Wiley Interdiscip. Rev. Cogn. Sci.* 1, 559–572.

Bialystok, E., Craik, F. I. M., Klein, R., and Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: evidence from the Simon task. *Psychol. Aging* 19, 290–303.

Bialystok, E., Craik, F. I. M., and Ryan, J. (2006). Executive control in a modified antisaccade task: effects of aging and bilingualism. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 1341–1354.

Bialystok, E., Luk, G., Peets, K. F., and Yang, S. J. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Biling. Lang. Cogn.* 13, 525–531.

Blumenfeld, H. K., and Marian, V. (2011). Bilingualism influences inhibitory control in auditory comprehension. *Cognition* 118, 245–257.

Brown, E. K., and Miller, J. E. (1980). *The Syntax of Scottish English.* Final Report to SSRC (UK), Project No. 5152, London.

Chaffe, W. (1980). "Some reasons for hesitating," in *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*, eds H. W. Dechert and M. Raupach (The Hague: Mouton), 85–90.

Christoffels, I. K., and De Groot, A. M. B. (2004). Components of simultaneous interpreting: a comparison with shadowing and paraphrasing. *Biling. Lang. Cogn.* 7, 1–14.

Colomé, À., and Miozzo, M. (2010). Which words are activated during bilingual word production? *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 96–109.

Costa, A., and Caramazza, A. (1999). Is lexical selection language specific? Further evidence from Spanish-English bilinguals. *Biling. Lang. Cogn.* 2, 231–244.

Costa, A., Pickering, M. J., and Sorace, A. (2008). Alignment in second language dialogue. *Lang. Cogn. Process.* 23, 528–556.

Costa, A., and Santesteban, M. (2004). Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *J. Mem. Lang.* 50, 491–511.

De Groot, A. M. B. (2011). *Language and Cognition in Bilinguals and Multilinguals: An Introduction.* New York: Psychology Press.

De Jong, N. H., and Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behav. Res. Methods* 41, 385–390.

Dell, G. S., Chang, F., and Griffin, Z. M. (1999). Connectionist models of language production: lexical access and grammatical encoding. *Cogn. Sci.* 23, 517–542.

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., and Bruce, V. (1997). Face-to-face and video-mediated communication: a comparison of dialogue structure and task performance. *J. Exp. Psychol. Appl.* 3, 105–125.

Favreau, M., and Segalowitz, N. S. (1983). Automatic and controlled processes in the 1st-language and 2nd-language reading of fluent bilinguals. *Mem. Cognit.* 11, 565–574.

Fehringer, C., and Fry, C. (2007). Hesitation phenomena in the language production of bilingual speakers. *Folia Linguistica* 41, 37–72.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *J. Mem. Lang.* 30, 210–233.

Flege, J. (1999). *Age of Learning and Second-Language Speech.* Hillsdale, NJ: Lawrence Erlbaum.

Garrod, S., and Pickering, M. J. (2004). Why is conversation so easy? *Trends Cogn. Sci. (Regul. Ed.)* 8, 8–11.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in Spontaneous speech/Frieda Goldman-Eisler.* London, NY: Academic Press.

Gollan, T. H., Bonanni, M. P., and Montoya, R. I. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology* 19, 278–287.

Gollan, T. H., and Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. *J. Exp. Psychol. Learn. Mem. Cogn.* 35, 640–665.

Gollan, T. H., Montoya, R. I., Cera, C., and Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: aging, bilingualism, and the weaker links hypothesis. *J. Mem. Lang.* 58, 787–814.

Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., and Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: the frequency-lag hypothesis. *J. Exp. Psychol. Gen.* 140, 186–209.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Biling. Lang. Cogn.* 1, 67–81.

Griffin, Z. M., and Ferreira, V. S. (2006). "Properties of spoken language production," in *Handbook of Psycholinguistics*, 2 Edn, ed. M. J. T. M. A. Gernsbacher (London: Elsevier), 21–59.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Res.* 18, 1279–1296.

Hanulová, J., Davidson, D. J., and Indefrey, P. (2010). Where does the delay in L2 picture naming come from? Psycholinguistic and neurocognitive evidence on second language word production. *Lang. Cogn. Process* 26, 902–934.

Hartsuiker, R. J., and Pickering, M. J. (2008). Language integration in bilingual sentence production. *Acta Psychol. (Amst.)* 128, 479–489.

Hartsuiker, R. J., Pickering, M. J., and Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychol. Sci.* 15, 409–414.

Henderson, A., Goldman-Eisler, F., and Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Lang. Speech* 9, 207–216.

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

Ivanova, I., and Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychol. (Amst.)* 127, 277–288.

Kempen, G., and Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cogn. Sci.* 11, 201–258.

Kootstra, G. J., van Hell, J. G., and Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: evidence from bilingual monologue and dialogue. *J. Mem. Lang.* 63, 210–231.

Kormos, J. (2006). *Speech Production and Second Language Acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Kormos, J., and Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System* 32, 145–164.

Kroll, J. F., Bobb, S. C., Misra, M., and Guo, T. (2008). Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol. (Amst.)* 128, 416–430.

Kroll, J. F., Bobb, S. C., and Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Biling. Lang. Cogn.* 9, 119–135.

Kroll, J. F., Michael, E., Tokowicz, N., and Dufour, R. (2002). The development of lexical fluency in a second language. *Second Lang. Res.* 18, 137–171.

Lambert, W. (1974). *Culture and Language as Factors in Learning and Education.* Bellingham, WA: Western Washington State College.

Levelt, W. J. M. (1983). Monitoring and self-repair in speech. *Cognition* 14, 41–104.

Levelt, W. J. M. (1989). *Speaking. Cambridge.* MA: The MIT Press.

Linck, J. A., Hoshino, N., and Kroll, J. F. (2008). Cross-language lexical processes and inhibitory control. *Ment. Lex.* 3, 349–374.

Lindsley, J. R. (1975). Producing simple utterances – how far ahead do we plan. *Cogn. Psychol.* 7, 1–19.

Macafee, C. (1983). *Glasgow.* Amsterdam: John Benjamin.

Macaulay, R. (1985). *The Narrative Skills of a Scottish Coal Miner.* Amsterdam: John Benjamin.

Marian, V., Blumenfeld, K. H., and Kaushanskaya, M. (2007). The language proficiency and experience questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967.

McFarland, D. H. (2001). Respiratory markers of conversational interaction. *J. Speech Lang. Hear. Res.* 44, 128–143.

McMurray, B., Samelson, V. M., Lee, S. H., and Tomblin, J. B. (2010). Individual differences in online spoken word recognition: implications for SLI. *Cogn. Psychol.* 60, 1–39.

Michael, E., and Gollan, T. H. (2005). "Being and becoming bilingual: individual differences and consequences for language production," in *The Handbook of Bilingualism: Psycholinguistic Approaches*, eds J. F. Kroll and A. M. B. de Groot (New York: Oxford University Press), 389–407.

Petrides, M. (1998). *Specialised Systems for the Processing of Mnemonic Information Within the Primate Frontal Cortex.* Oxford: Oxford University Press.

Pickering, M. J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* 27, 169–190.

Pinkham, A. E., and Penn, D. L. (2006). Neurocognitive and social cognitive predictors of interpersonal skill in schizophrenia. *Psychiatry Res.* 143, 167–178.

Poulisse, N., and Bongaerts, T. (1994). 1st language use in 2nd-language production. *Appl. Linguist.* 15, 36–57.

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing (Version 2.10.1): R Foundation for Statistical Computing.* Available at: http://www.r-project.org

Sandoval, T. C., Gollan, T. H., Ferreira, V. S., and Salmon, D. P. (2010).

What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Biling. Lang. Cogn.* 13, 231–252.

Segalowitz, N. (2010). *The Cognitive Bases of Second Language Fluency*. New York: Routledge.

Segalowitz, N., and Hulstijn, J. (2005). "Automaticity in bilingualism and second language learning," in *Handbook of Bilingualism: Psycholinguistic Approaches,* eds J. F. Kroll and A. B. B. de Groot (New York: Oxford University Press), 371–388.

Shallice, T., and Burgess, P. (1996). The domain of supervisory processes and temporal organization of behaviour. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351, 1405–1412.

Shockley, K., Richardson, D. C., and Dale, R. (2009). Conversation and coordinative structures. *Top. Cogn. Sci.* 1, 305–319.

Simon, J. R., and Ruddell, A. P. (1967). Auditory S-R compatibility: the effect of an irrelevant cue on information processing. *J. Appl. Psychol.* 51, 300–304.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662.

Titone, D., Libben, M., Mercier, J., Whitford, V., and Pivneva, I. (2011). Bilingual lexical access during L1 sentence reading: the effects of L2 knowledge, semantic constraint, and L1–L2 intermixing. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 1412–1431.

Towell, R., Hawkins, R., and Bazergui, N. (1996). The development of fluency in advanced learners of French. *Appl. Ling.* 17, 84–119.

Wechsler, D. (1999). Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation.

Welkowitz, J., Bond, R. N., and Zelano, J. (1990). An automated-system for the analyses of temporal speech patterns – description of the hardware and software. *J. Commun. Disord.* 23, 347–364.

Whitford, V., and Titone, D. (2012). Second-(language)experience modulates first- and second-language word frequency effects: evidence from eye movement measures of natural paragraph reading. *Psychon. Bull. Rev.* 19, 73–80.

Wilson, M., and Wilson, T. (2005). An oscillator model of the timing of turn-taking. *Psychon. Bull. Rev.* 12, 957–968.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

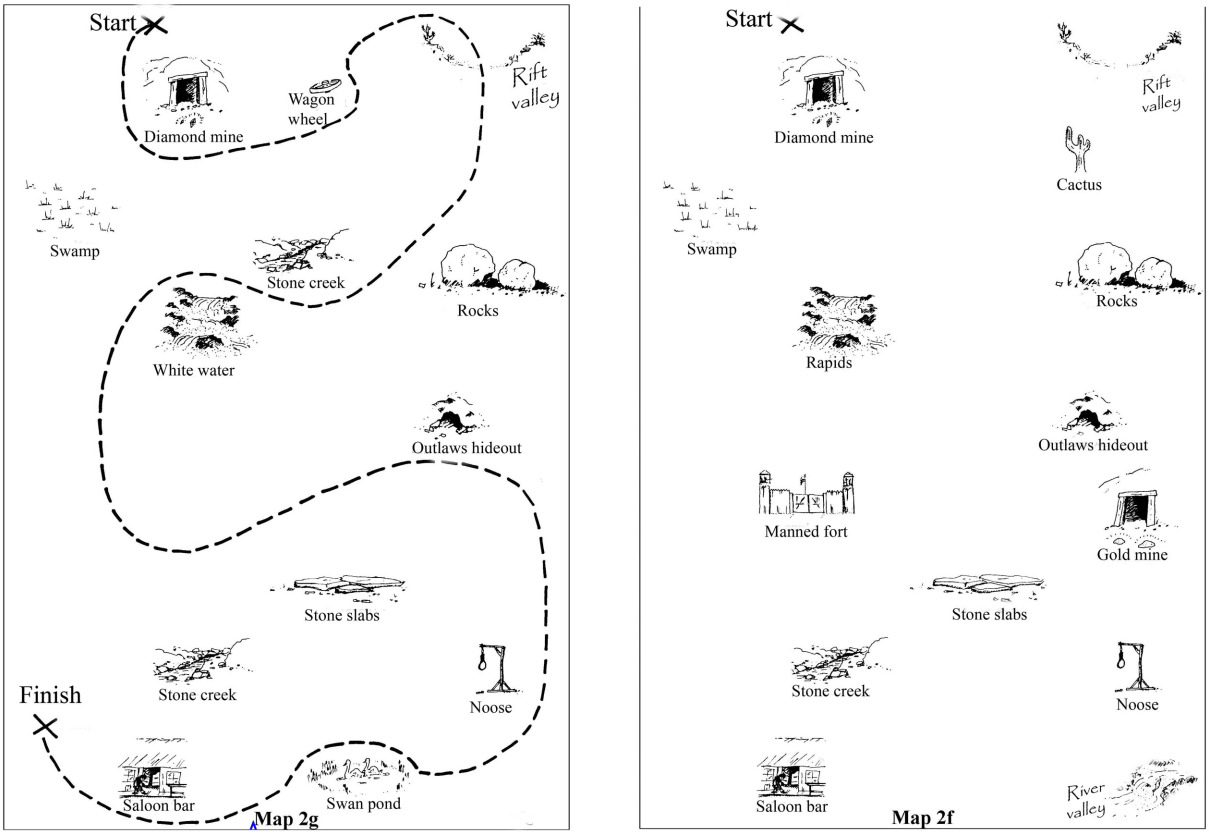
Received: 09 August 2011; accepted: 13 February 2012; published online: 16 March 2012.

*Citation: Pivneva I, Palmer C and Titone D (2012) Inhibitory control and L2 proficiency modulate bilingual language production: evidence from spontaneous monologue and dialogue speech. Front. Psychology **3**:57. doi: 10.3389/fpsyg.2012.00057*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.

APPENDICES

APPENDIX A



APPENDIX B

FE participant – L1 dialogue

Participant: On va vers, euh... Vers la gauche, un peu. Donc en ligne droite, euh...A l'horizontale. Après, moi il y a l'eau vive. C'est une chute...
Confederate: Moi je touche les rapides. Est-que c'est correct? Moi j'ai les rapides ici.
Participant: Euh...Ben ça ressemble à des rapides sur ma photo mais ça s'appelle «eau vive», donc peut-être que c'est la même chose. Euh...On les contourne. Donc on passe en haut pour descendre le plus vers la gauche de la carte.
Confederate: Vers la gauche de la carte... OK.
Participant: On le contourne, oui. Vers le haut des...de l'eau vive pour descendre après ça. À gauche. Donc on descend quand même un peu, pour se rendre dans le dernier tiers de la carte, disons. Donc on descend à la verticale.
Confederate: A la verticale...OK.
(Translation)
Participant: We go towards, um... towards the left a little. So in a straight line....um... horizontally. After, I have white water. It's a fall.
Confederate: I touch the rapids. Is that OK? I have the rapids here.
Participant: Um... Well it looks like rapids on my picture but it's called "white water" So maybe it's the same thing. Um... so we go around them. So we pass above to go down to the left of the map.
Confederate: To the left of the map...OK.
Participant: We go around them, yes. Towards the top of the... the white water... To go down after that. On the left. So we still go down a little. To get to the last third of the map, let's say. So we go down vertically.
Confederate: Vertically...OK.

EF participant – L1 dialogue

Participant: And then do you have stone creek?
Confederate: Um... Yes, at the bottom of the page.
Participant: Um...No, OK. There is another one.
Confederate: OK.
Participant: Um....It's not far from the rocks but it's... It's basically, like, right at the center of the page. Where the rocks are but, like, towards the center. So you... You go under the stone creek...
Confederate: OK.
Participant: ...After the rocks. And then there is white water.
Confederate: OK. I have rapids...
Participant: OK. So... it's probably the same thing. And so you go over it and then...

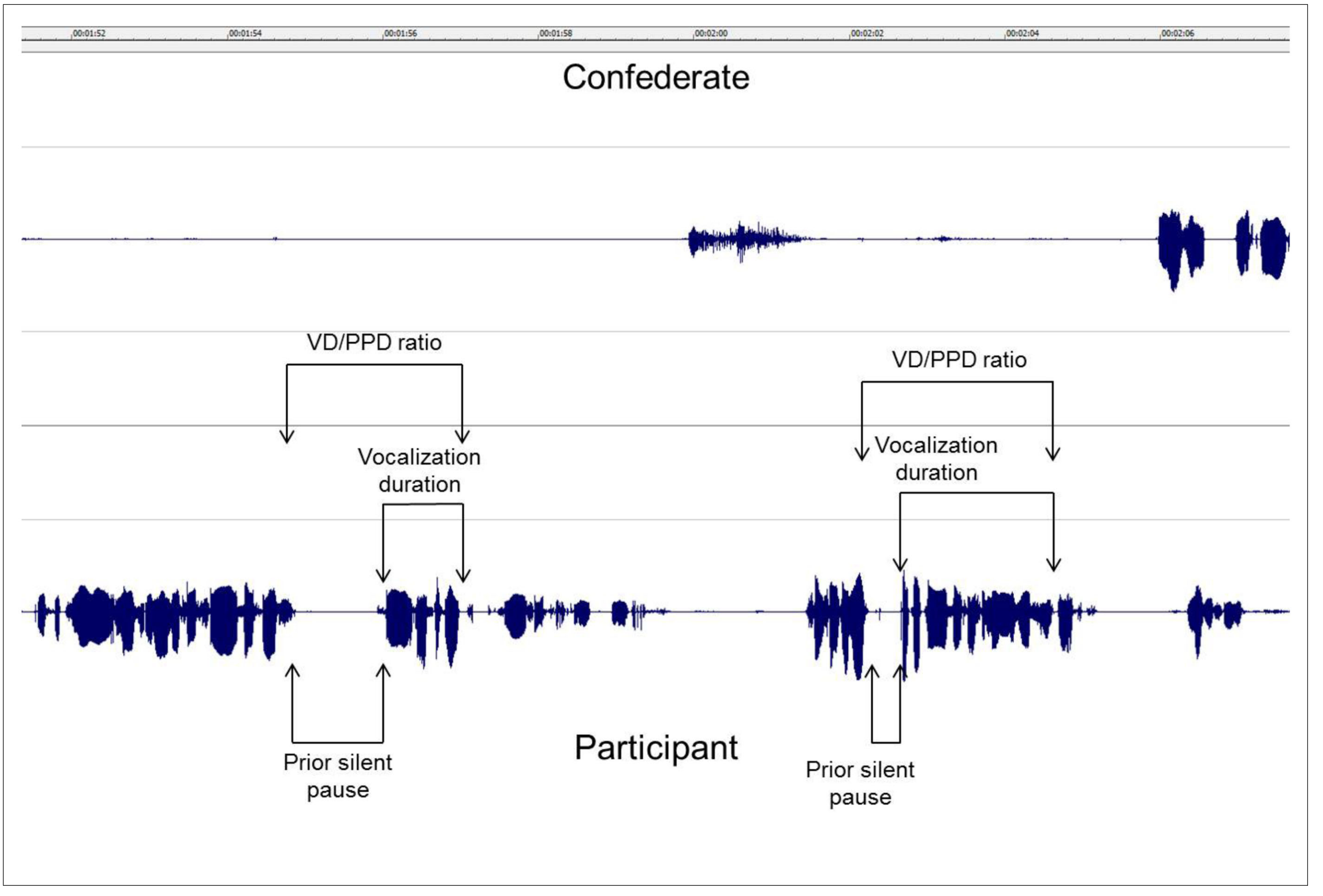
EF participant – L2 dialogue

Participant: Après ça, on va se diriger comme dans une ligne diagonale allant vers le ruisseau des roches. Comme, a ce point-là, ça va être à ta droite.
Confederate: OK...Attends... Mon ruisseau des roches est comme vraiment en bas de la page.
Participant: Oui. Oh, OK! Non, non, non! Euh... le mien... le, le ruisseau de roche sur ma page, c'est comme...c'est à la même hauteur des roches, sauf c'est comme... Ils sont séparés de 3cm ou quoi.
Confederate: Ah OK OK! Moi j'ai des rapides qui sont vraiment un peu en bas. En bas des roches. C'est comme... de ruisseau de roche ou de ce que tu m'as dit...c'est comme entre les deux. C'est ça? Es que tu as des rapides?
Participant: J'ai des eaux vives. Ça a l'air des rapides.
(Translation)
Participant: After that, we are going to go in like, a diagonal line, going towards the stone creek. Like, at that point, it's going to be on your right.
Confederate: Ok...wait....My stone creek is like, really at the bottom of the page.
Participant: Yes. Oh, OK! No nono! Um... mine... The stone creek on my page. It's like...It's at the same height as the rocks. Except that it's like... They are separated by 3cm or something.
Confederate: Oh OK OK! I have rapids that are really a bit down. Below the rocks. It's like... From stone creek or from what you told me...It's like between the two. Right? Do you have rapids?
Participant: I have white water. It looks like rapids...

FE participant – L2 dialogue

Participant: You go towards the left of the sheet.
Confederate: Aha OK. So, I just go in a straight line?
Participant: In a straight line between the stone creek and the rocks.
Confederate: Stone creek? I only have a stone creek, like, at the bottom of the page. But not...
Participant: OK, well you go in a diagonal line, at the left of the rocks.
Confederate: At the left of the rocks... Like, how many centimeters am I away from the rocks?
Participant: um... 1.
Confederate: 1? OK, So I go diagonal. Like 45 degrees?
Participant: Um...Yeah.
Confederate: And where do I stop? At the rapids?
Participant: Um... not yet!
Confederate: Not yet, OK.
Participant: When you go down in a diagonal line for maybe about 5 cm.
Confederate: 1, 2, 3, 4, 5. OK.
Participant: And then you have to go at the top of the picture of the rapids.

APPENDIX C


Confederate
VD/PPD ratio
Vocalization duration
VD/PPD ratio
Vocalization duration
Prior silent pause
Participant
Prior silent pause

frontiers in
PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 22 November 2011
doi: 10.3389/fpsyg.2011.00334

On a possible relationship between linguistic expertise and EEG gamma band phase synchrony

***Susanne Reiterer*[1,2] *, *Ernesto Pereda*[3] *and Joydeep Bhattacharya*[4,5]**

[1] Department for English Linguistics, Center for Linguistics, University of Tübingen, Germany
[2] Department of English Studies, University of Vienna, Vienna, Austria
[3] Department of Basic Physics, University of La Laguna, Tenerife, Spain
[4] Department of Psychology, Goldsmiths College, University of London, London, UK
[5] Commission for Scientific Visualization, Austrian Academy of Sciences, Vienna, Austria

Edited by:
Teresa Bajo, Universidad de Granada, Spain

Reviewed by:
Steve Majerus, Université de Liège, Belgium
Alfredo Brancucci, Università 'G.d'Annunzio' di Chieti e Pescara, Italy

***Correspondence:**
Susanne Reiterer, Department for English Linguistics, Center for Linguistics, University of Tübingen, Hoppe Seyler Strasse 3, 72076 Tübingen, Germany.
e-mail: susanne.reiterer@med.uni-tuebingen.de

Recent research has shown that extensive training in and exposure to a second language can modify the language organization in the brain by causing both structural and functional changes. However it is not yet known how these changes are manifested by the dynamic brain oscillations and synchronization patterns subserving the language networks. In search for synchronization correlates of proficiency and expertise in second language acquisition, multivariate EEG signals were recorded from 44 high and low proficiency bilinguals during processing of natural language in their first and second languages. Gamma band (30–45 Hz) phase synchronization (PS) was calculated mainly by two recently developed methods: coarse-graining of Markov chains (estimating global phase synchrony, measuring the degree of PS between one electrode and all other electrodes), and phase lag index (PLI; estimating bivariate phase synchrony, measuring the degree of PS between a pair of electrodes). On comparing second versus first language processing, global PS by coarse-graining Markov chains indicated that processing of the second language needs significantly higher synchronization strength than first language. On comparing the proficiency groups, bivariate PS measure (i.e., PLI) revealed that during second language processing the low proficiency group showed stronger and broader network patterns than the high proficiency group, with interconnectivities between a left fronto-parietal network. Mean phase coherence analysis also indicated that the network activity was globally stronger in the low proficiency group during second language processing.

Keywords: EEG, gamma band, phase synchronization, bilinguals, second language acquisition, cortical efficiency, linguistic expertise, individual differences in proficiency

INTRODUCTION

Most brain imaging studies on bilinguals/multilinguals have been conducted with either positron emission tomography (PET) or functional magnetic resonance imaging (fMRI) with a pure emphasis on localizing brain activities (e.g., see De Bot, 2008 for review). They have not specifically investigated the functional connectivity between different and distributed brain areas, yet one of the most discussed hypotheses – the influence of second language proficiency level on the extent and distribution of brain activation – would call for a method analyzing functional cooperation and interactions of brain areas. This is frequently done in the field of EEG research by using coherence or synchronization analyses (Ward, 2003; Allefeld et al., 2005; Fries, 2005; Stam, 2005). Bilingual brain organization in terms of networks and functional cooperation has been scarcely investigated hitherto (for EEG coherence see Reiterer et al., 2005a,b; and for fMRI connectivity see Dodel et al., 2005 and Majerus et al., 2008 as examples). In fact, the study by Dodel et al. (2005) and the recent study by Majerus et al. (2008) are the only examples, to the best of our knowledge, that have investigated cortical synchronization patterns by employing fMRI connectivity analyses in bilingual language (word and sentence processing, Dodel et al., 2005) as well as native language short-term memory (STM) processing (Majerus et al., 2008). Interestingly, the first connectivity study (Dodel et al., 2005) found larger and more extended networks for the bilinguals with higher proficiency levels, contrary to many studies on bilingual fluency levels, which find fewer activated areas as a function of higher fluency levels in second languages (e.g., Perani et al., 1996, 1998; Yetkin et al., 1996; Chee et al., 2001; Hasegawa et al., 2002; Briellmann et al., 2004; Xue et al., 2004; Klein et al., 2006). The second connectivity study (Majerus et al., 2008), albeit not investigating language or bilingual language processing *per se* but STM processing instead, could nevertheless differentiate high from low proficiency bilinguals by

Abbreviations: EEG, electroencephalography; Cz, C4 electrode positions on scalp over central and right hemisphere areas; fMRI, functional magnetic resonance imaging; F8, electrode position on scalp (frontal right); Fp = frontal; 8 = RH; Fp1, electrode position on scalp (prefrontal left); Fp = fronto-polar; 1 = LH; Fp2, electrode position on scalp (prefrontal right); Fp = fronto-polar; 2 = RH; HP, high proficiency group; HT, Hilbert transform; kΩ, kilo Ohm; L1, first language (mother tongue); L2, second language (foreign language); LH, left hemisphere; LP, low proficiency group; μV, micro Volt; PET, positron emission tomography; PS (index), phase synchronization (index); RH, right hemisphere; SCA, synchronization cluster analysis; SD, standard deviation; T4, electrode position on scalp over temporal right hemisphere area; T5, electrode position on scalp over temporal left hemisphere area.

means of fMRI connectivity patterns. They found the connectivity patterns to be characteristically diverse (rather than e.g., larger or smaller) for the two behaviorally different bilingual proficiency groups, with the low proficiency group showing a less specialized and less differentiated neural network underlying (serial order) STM processing, which, according to the authors, leads to a less efficient processing of serial order information in STM in the low proficiency group (a fact which is assumed to be causally connected to their generally poorer second language performance).

However, in the field of EEG synchronization, we did not find any comparable studies that investigated bilingual proficiency levels.

In an earlier study (Reiterer et al., 2005a), we analyzed EEG coherence in the lower and middle frequency ranges [from delta (1–4 Hz) to beta range (13–30 Hz)], and found a significant correlation between proficiency level and EEG coherence within the alpha band (8–12 Hz) (Reiterer et al., 2005b). The high proficiency (HP) group displayed lower coherence for both, native and foreign, language stimuli. Since the alpha band primarily reflects attentional processes, this result could possibly indicate a general language processing strategy based on general attentional processes, but not necessarily a differential language processing strategy [differentiating first (L1) from second language (L2)]. Further, the alpha band might have been too narrow to capture the differences in proficiency related to the different languages. Broad high frequency bands, such as gamma band, could be a more promising candidate to capture linguistic processes at a higher level of sophistication.

Based on these studies that revealed differences in activation patterns as a function of fluency level differences (e.g., efficient processing as in Just et al., 1996), we hypothesized that low proficiency bilinguals, as compared to high proficiency bilinguals, would be associated with a higher degree of gamma band synchronization during second language processing. Some studies (Simos et al., 2002; Micheloyannis et al., 2003; Hagoort et al., 2004; Ford et al., 2005; Weiss et al., 2005; Bastiaansen and Hagoort, 2006; Hald et al., 2006; Ihara and Kakigi, 2006; Bastiaansen et al., 2010) have already pointed to relations between gamma band synchronization and native language processing, but second language or bilingual language processing has almost not been investigated in this high frequency range. A notable exception here is the study by Ihara and Kakigi (2006), which already adverted to a putative role of the alpha and the gamma band for detecting possible differences between first and second language systems. Furthermore, we want to make a distinction between short-range or local synchronization, i.e., synchronization within a node of a functional network, and long-range synchronization, i.e., synchronization between different nodes of a network (Bastiaansen and Hagoort, 2006; Le Van Quyen and Bragin, 2007). Local gamma synchronization occurs when a large number of neurons transiently oscillate with a common phase and is primarily represented by the spectral content of the gamma band oscillation of any individual EEG electrode, whereas the long-range gamma synchronization occurs when two preferably large neuronal populations recorded by two distant EEG electrodes oscillate with a phase relationship over (at least) a few gamma oscillation cycles and is primarily represented by the degree of phase synchrony between these two EEG electrodes. The majority of the studies on native language comprehension addressed only the spectral power changes within an EEG electrode (i.e., local synchrony), while ignoring the relationship between multiple electrodes (i.e., long-range synchrony).

So, in the present work, we exclusively investigated and analyzed the long-range synchronization properties of gamma band during language comprehension in late bilinguals. To our knowledge, this is the first attempt to investigate the influence of the amount of linguistic training and expertise on long-range gamma band phase synchronization (PS).

MATERIALS AND METHODS

PARTICIPANTS

We contrasted two groups of differentially proficient second language (L2 = English) speakers, who had overall comparable educational level (University students), but differed in their proficiency levels in L2 due to different amounts of training in English. The participants in the "high proficiency group" (HP) were advanced university language students studying English language and linguistics for a master's degree (last year, 5–6 years completed). Their level of English proficiency was "very good" (so-called native speaker-like performance) or "good" according to their performances at university. Additionally, this level was verified by a certified English language teacher according to oral fluency test interviews. This rating system reflected the Austrian school marking system from one to five (max-to-min) and according to this rating system the participants were divided into the following five categories "very good," "good," "medium," "lower-level," "lowest–level." Most of the participants in the HP also studied a second foreign language (i.e., an L3) like French, Italian, or Spanish, or general linguistics. They all had high levels of linguistic training and knowledge at the time of the experiment. In other words, they have been "pre-screened" for HP at University already. As for their exposure to real life surroundings with the second language, the average amount of time they had spent abroad in an English speaking country was 10 months. The participants in the "low proficiency group" (LP) were university students of various disciplines studying for a master's degree in a subject other than language and linguistics (e.g., biology, psychology). They displayed medium to low level second language skills (corresponding to the groups "medium," "lower-level," and "lowest-level"), which were sufficient to let them pass their school leaving exams ("Matura," an equivalent to "A levels"), but since then were not developed any further. They were able to lead basic level conversations in English, but their speech was non-fluent, characterized by grammatical errors, poor pronunciation (foreign accent), slowed-down lexical access, and long pauses. The average amount of time LP participants spent abroad in an English speaking country was 5 weeks. With regard to the country where they had spent some time, the groups were homogeneous.

We strictly controlled for the variable "age of onset" of L2 learning. The average (±SD) age of onset was 9 years (1 year) and was matched between the two groups. Further controlled variables were: age, handedness, gender, mother tongue, socio-educational, and cultural background, region of residence, and non-verbal and verbal IQ. Each group consisted of 22 right-handed (measured by the Edinburgh handedness inventory; Oldfield, 1971) female students with German as their native language. We rigidly controlled for the variable gender in order to avoid possible influences of gender onto the processing of language and its neural representations. After manual and automatized artifact control we had to exclude six subjects (mostly because of muscle artifacts and/or paroxysmal oscillations in the EEG signals) from the further statistical analyses, so that finally each group was composed of 19 participants.

Mean (SD) age was 24 years (2.3 years and 2.7 years respectively for two groups) for both groups. They were also matched for socio-cultural background and education: all participants had similar social (middle class), educational (university students), and cultural (living in Vienna) background.

The two groups differed from each other mainly in the amount of second language training they were exposed to, and the difference was approximately 6 years. Summarizing, the differences between the two groups are in their linguistic experience and knowledge, hence, in their proficiency levels in English as their L2.

The study was in compliant with the Code of Ethics of the World Medical association (Declaration of Helsinki) and the experimental protocol was approved by the local ethics committee. All subjects gave their written informed consent for the study.

STIMULUS MATERIAL

The cerebral organization of language at the word and sentence level has been investigated extensively with PET, fMRI, MEG, and event-related potential studies, but much less research has to date been carried out on the processing of coherent language at the discourse level where language occurs in its natural context (i.e., where phonetic, syntactic, semantic, and pragmatic aspects of language are integrated). Therefore, in this study we adopted coherent spoken speech (radio news) as stimuli and used them in a listening comprehension and discourse processing paradigm. In cooperation with the English department at the University of Vienna, the speech samples were matched for syntactic complexity, semantic content, and genre (only reports of medium complexity level on daily politics and business were chosen), discourse structure (reports had the form of a monolog), and gender of the speaker (all male speakers). Within the framework of a block design, six blocks of coherent speech (2.0–3.2 min each) with randomly inserted baseline blocks (acoustic noise, 2.0 min each) were presented in randomized order: three blocks in condition L2 English and three blocks in condition L1 German were auditorily presented in randomized order. All stimuli were presented via earphones and a white fixation cross was presented throughout the auditory task. For visualization of stimulus presentation procedure see **Figure 1**.

The whole recording session, which began at the same time of day for each participant (9 O'clock a.m.), took approximately 3 h for each participant, including manual electrode placement, explanation of the procedure, personal questionnaires, a familiarization task, and interruptions for answering detailed open ended comprehension questions orally performed with an experimenter (a tutor of the English Department) who was blind regarding group membership. In these interview sessions, six psychological reaction parameters were assessed with the help of a behavioral questionnaire, comprising (1) The actual text comprehension (seven factual comprehension questions about the contents of the radio report were posed, 14 points = max score, 0 = min score), (2) Subjective text comprehension (participant scored himself on a rating scale from 1 to 5), (3) Self-reported attention (same procedure), (4) Cognitive work-load (same procedure), (5) Sympathy for the speaker (same procedure), and (6) Interest in the subject matter of the radio fragment (same procedure).

DATA RECORDING

We recorded multivariate EEG signals during L1 and L2 processing in a quiet, dimly lit sound-proof experimental room. Participants were monitored through a video control system during the recording session in order to control for possible movements. Nineteen gold-disk electrodes were carefully attached to the scalp with adhesive electrode gel, positioned according to the international 10/20 System (Jasper, 1958; **Figure 2**); one additional frontal electrode was used as a ground, and two separate electrodes, at the right and left ear-lobe, were used as reference electrodes. The recordings were referenced against the algebraic mean of the two ear-lobe electrodes (Essl and Rappelsberger, 1998). Eye movements were additionally controlled for by a piezo-electric device attached to the eyelid. Using a conventional Nihon-Kohden 21 channel recorder, the EEG was amplified, filtered (time constant 0.3 s),

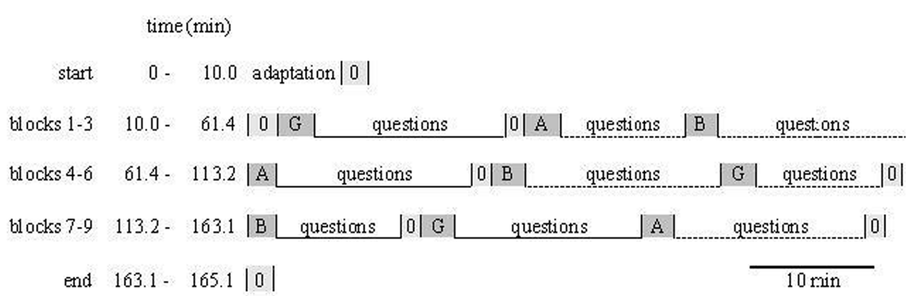


FIGURE 1 | Timing of recording session (example). Three tasks of the different language varieties (B = British English, A = American English, or G = Austrian German) were linked together to one big block comprising three small blocks with their respective baseline conditions, and the three big blocks (1–3, 4–6, 7–9) were presented in randomized order. After each language task (acoustic presentation), recording was interrupted for questions to explore factual listening comprehension, the attitude toward the speaker, interest for the contents, attention, work-load, and subjective comprehension (0 = baseline task).

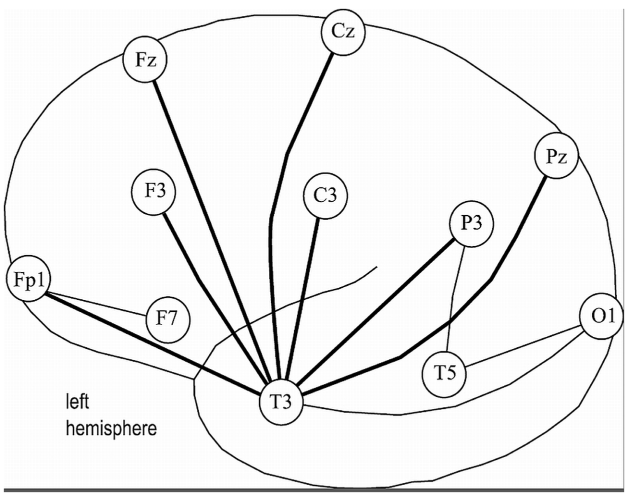


FIGURE 2 | Schematic map of the left hemisphere of EEG electrode positions (positioning according to 10/20system). Odd numbers represent loci in the left hemisphere, even numbers loci in the RH respectively. Fp = fronto-polar region, F = frontal lobe, C = central region, P = parietal lobe, T = temporal lobe, O = occipital lobe. Index letter "z" means "zero" for midline (central line, vertex). Indexing numbers on the RH would be: 4, 8, 2, and 6 instead of 3, 7, 1, and 5 respectively.

displayed and recorded at a sampling rate of 128 Hz for further processing. The electrode impedance was kept below 5 kΩ. A notch at 50 Hz was used for the elimination of power line contamination. We applied two criteria for possible artifact rejection: we rejected those epochs with amplitudes higher than 70 μV (absolute value), plus additional epochs where 2% or more samples deviated more than 3 SD from the mean value.

DATA ANALYSIS

Phase synchronization (PS) between all possible electrode pairs [$(19 \times 18)/2 = 171$ different electrode pairs] was calculated in the lower gamma frequency range (30–45 Hz, the choice was made after earlier studies, see also Bhattacharya et al., 2001; Bhattacharya et al., 2003; Bhattacharya and Petsche, 2005a,b) by three methods: (i) global PS by coarse-graining of Markov chains (CGMC) which measures the degree of PS of one electrode with all other electrodes, (ii) bivariate PS [by the recently developed phase lag index (PLI)] which measures the degree of PS between pairs of electrodes, and additionally (iii) the more conventional mean phase coherence, which measures the degree of PS between a pair of electrodes, but is more prone to volume conduction effects. However, all three methods initially require a proper estimation of the phases from the EEG signals.

Estimation of the phases

Since we were mainly interested in the gamma frequency band, each EEG signal was band-pass filtered using a zero-phase filter with 30 and 45 Hz cut-off frequencies to get the desired gamma band signal. We calculated the phases of these filtered signals, $\{x(k)\}$, by using the analytic signal approach based on Hilbert transform (HT), where the analytical signal $\zeta(t)$ is obtained:

$$\zeta(t) = x(t) + ix_{\mathrm{H}}(t) \tag{1}$$

where $x_{\mathrm{H}}(t)$ is the HT of $x(t)$, defined as:

$$x_H(t) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{x(t)}{t - t'} dt' \tag{2}$$

with $P.V.$ denoting the Cauchy principal value.

The analytic signal, which is also a complex function, can be decomposed as:

$$\zeta(t) = a_x(t) \exp^{i\phi_x(t)} \tag{3}$$

where $a_x(t)$ is the instantaneous amplitude and $\phi_x(t)$ is the instantaneous phase of $x(t)$.

In this way, the phases of 19 EEG channels, $\phi_i(t)$ $(i = 1, \ldots, 19)$, were estimated and subsequently used to assess the degree of PS in each situation, as explained below.

Estimating global phase synchronization: Coarse-Graining of Markov Chains (CGMC)

The collective synchronization of the ensemble of 19 electrodes was studied by means of a recently derived method (CGMC; Allefeld, 2006; Allefeld and Bialonski, 2007). CGMC is a multivariate method that allows the detection of synchronization clusters from the 19×19 matrix of bivariate PS indexes (in our case, PLI). Briefly, this matrix is translated into a stochastic matrix P describing a finite-state Markov process, and subsequently, it is possible to estimate the number of clusters present in the data via the eigenvalue decomposition of P. Additionally, it allows the estimation of the strength of each cluster as well as the degree of participation of each electrode in the cluster it belongs to.

The relevant fact about CGMC is that it is truly multivariate in the sense that, given a set of n electrodes ($n > 2$), it estimates the degree of overall synchronization among all the electrodes and their distribution in q synchronization clusters $(q \geq 1)$ to which each electrode of contributes differently. The validity of this approach in EEG applications has been demonstrated (Allefeld and Kurths, 2004; Allefeld and Bialonski, 2007).

Although the CGMC allows an automatic determination of the value of q from the data, after a preliminary exploration we fixed $q = 2$ so that the 19 electrodes are assigned to either the strongly synchronized or the weakly synchronized cluster.

Estimating bivariate phase synchronization: phase lag index (PLI)

There are many different ways of assessing the PS between a pair of EEG signals (see, e.g., (Pereda et al., 2005)). Here, we used the PLI (Stam et al., 2007), because it is less sensitive to volume conduction effects than other popular indexes of PS such as, for instance, the mean phase coherence (Mormann et al., 2000). The PLI is defined as:

$$\mathrm{PLI} = \left|\left\langle \mathrm{sgn}\left(\varphi(t_k)\right)\right\rangle\right| \tag{4}$$

where |•| indicates modulus, ⟨•⟩ indicates time average and

$$\varphi(t) = \left|\phi_i(t) - \phi_j(t)\right| \bmod (2\pi) \tag{5}$$

is the cyclic relative phase, i.e., the phase difference between $x_i(t)$ and $x_j(t)$ wrapped to the interval $[0, 2\pi]$. The PLI ranges from 0 (two signals with no phase relationship or a phase relationship symmetrical about 0 or $\pm\pi$ -which is a signature of volume conduction effects (Nolte et al., 2004; Stam et al., 2007) to 1 (two signals with complete phase synchrony); PLI is parameter free.

Estimating bivariate phase synchronization (additional): mean phase coherence

Additionally, we also used the now more conventional mean phase coherence (Hoke et al., 1989; Mormann et al., 2000) defined as:

$$\gamma_{i,j} = \sqrt{\langle \cos \varphi (t) \rangle^2 + \langle \sin \varphi (t) \rangle^2} \quad (6)$$

where $\langle \bullet \rangle$ indicates time average and

$$\varphi (t) = \left| \phi_i (t) - \phi_j (t) \right| \bmod (2\pi) \quad (7)$$

is the cyclic relative phase, i.e., the phase difference between $x_i(t)$ and $x_j(t)$ wrapped to the interval $[0, 2\pi]$. The mean phase coherence index ranges from 0 (two signals with no phase relationship) to 1 (two signals with complete phase synchrony), and has the advantage of being parameter-free.

STATISTICAL ANALYSIS

The statistical differences in the synchronization strength of the strongly and the weakly synchronized cluster was tested by means of repeated measures ANOVA test with proficiency (HP and LP) as independent (between groups) factor and language processing (L1 and L2) as dependent (within group) factors. Differences were considered significant when the p-value was lower than 0.05.

For the sake of using balanced stimulus trials in the group comparisons of L1 versus L2, we only analyzed the three blocks of the condition "British English" versus the three blocks of "Austrian German," based on our earlier experience (Reiterer et al., 2005a) that the variant of English (British or American English) neither affected the coherence patterns in the EEG nor the respective behavioral outcomes. Thus it had emerged previously (as far as it can be discriminated by EEG synchronization analyses) that L2 English was processed as L2 English and not differentiated further into its subvariants or accents.

RESULTS

BEHAVIORAL RESULTS

A comprehension questionnaire applied after each task condition revealed (**Figure 3**) that the low proficiency (LP) group understood approximately 50% of the English (L2) texts, whereas the HP group understood nearly perfectly (95%). For the comprehension questions of the control condition German (L1) no statistical difference between the groups was obtained. The HP group had again a performance accuracy of 95%, whereas the LP group scored slightly, but not significantly, worse (performance accuracy of 80%).

No differences between the two groups were found for other psychometric variables (self-reported attention, work-load, sympathy for the speaker, and interest in the subject matter).

EEG SYNCHRONIZATION RESULTS

First we could differentiate L1 from L2 processing by gamma band global synchronization clustering patterns (CGMC) but on a more subtle scale, HP and LP bilinguals could further be significantly differentiated by bivariate gamma band synchronization measures (gamma band mean phase coherence and PLI) predominantly, or, almost exclusively when processing L2.

On a global scale, i.e., when estimating the "global PS" by the method of CGMC where the collective synchronization of the ensemble of 19 electrodes is studied with respect to cluster strength, we obtained very similar general results of clustering strength for both groups investigated (high and low proficiency group alike), however significantly different for the two languages, mother tongue German, and second language English. Cluster strength is significantly higher in both groups for the L2 than for the L1 (**Figure 3**).

When taking a more fine-grained view for the distribution of the electrodes belonging to the stronger cluster within the gamma band (measured with bivariate PS by means of the "PLI" which measures the PS between the single pairs of electrodes), we obtained more subtle differences in synchronization characteristics between the high and the low proficiency group.

The topography is markedly different between the HP and the LP group (**Figure 4**) with the LP group showing a strongly and significantly synchronized cluster only for processing L2 within left temporo-parietal areas preponderantly. The other topographical clusterings (areas) shaded in light blue over central and right hemisphere areas did not reach significance. To work out these between group differences in detail, we provided an additional figure (**Figure 5**) where these group differences in PLI are depicted in percentages (percentages of pertinence to the cluster) – the positive values (yellow/red color) corresponding to those electrodes

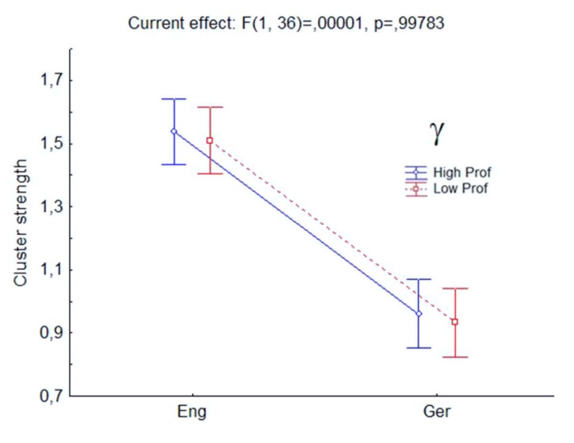


FIGURE 3 | Synchronization cluster (strength) analysis by "Coarse-Graining Markov Chains" (CGMC) within the (γ) gamma-band (30–45 Hz) showed significant differences between native language German (Ger, *x*-axis), cluster strength (*y*-axis) around 0.9, and second language English (Eng) with cluster strength significantly higher (around 1.5). Red line denotes the low proficiency and blue line the high proficiency group.

which belong to the strong cluster in a greater percentage to subjects in the LP group. In the LP group one can see a marked increase in the left temporo-parietal/central region while listening to L2 English whereas the increase is much lower and right temporal for the listening of German (**Figure 5**).

According to this result for the L1 we can assume that an increase up to 20% is within the statistical fluctuation (no difference is expected usually between the groups while listening to L1), so that an increase of 25% or above might be considered significant. These results indicate that the low proficiency group recruits more often the left temporo-parieto-central part of the cortex than the HP group when listening to L2 English, which is clearly also their less proficient language.

To obtain a topographical scalp distribution of the group differences in average participation of electrodes for the strongly synchronized cluster (PLI, within the gamma band) we calculated an index for the participation of each electrode and subtracted the values of the HP group from those of the LP group (**Figure 6**). Here we found that the greatest increment of involvement of the electrodes belongs to the LP group during the L2 language condition and is topographically most pronounced over left frontal, or fronto-central areas.

The positive values (red color) here indicate a higher participation of electrodes for the low proficiency group, again with a preponderance over the left hemisphere and stronger for L2 than for L1. The topographies are significantly different for the groups ($p < 0.01$, sign ranked paired test) when listening to the second language English, whereas for listening to the native language topographies are not significantly different between the groups (average difference is 5×10^{-4}, i.e., equal to zero).

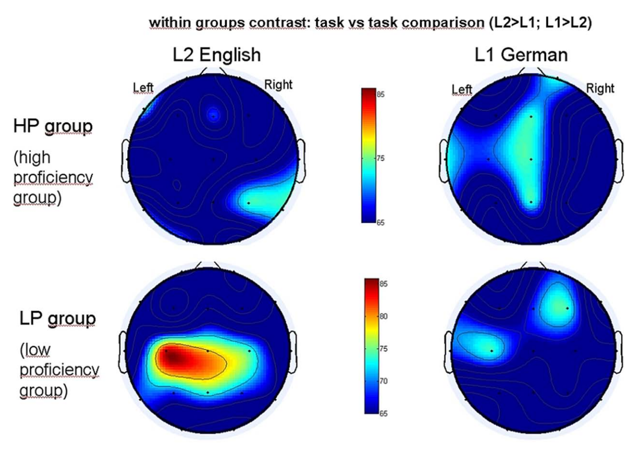


FIGURE 4 | Brain maps depicting gamma synchronization measured by phase lag index (PLI) during first (L1, right panel) and second (L2, left panel) language processing in the high proficiency (HP) group (upper row) and low proficiency (LP) group (lower row). Task versus task comparisons measuring within group differences at $p < 0.05$, corrected for multiple comparisons. Brain maps are to be seen from bird's eye perspective, with the frontal parts showing toward top of the page. Red color in color bar indicates percentage of electrodes' pertinence to a given cluster.

To get an impression about the most importantly involved electrodes, we depicted the results described in **Figure 6** additionally in a topographical map showing synchronization lines in the LP group (connections are given for interhemispheric long-range connections and within left/right hemisphere separately). Within the strongest participating left hemispheric frontal cluster, the greatest involvement of long-range interhemispheric gamma synchronization concerns the electrode positions: F3, F7, Fp1 to C4, P4, and T3 to F4. Within the right hemisphere only (corresponding to the right panel in **Figure 6**) the long-range connections involve electrode positions connecting frontal with parietal regions.

The results for the gamma band synchronization differences between the high and low proficiency group were corroborated by an additional analysis (**Figure 8**, calculation of mean phase coherence, see Estimating Bivariate Phase Synchronization (Additional): Mean Phase Coherence in Materials and Methods).

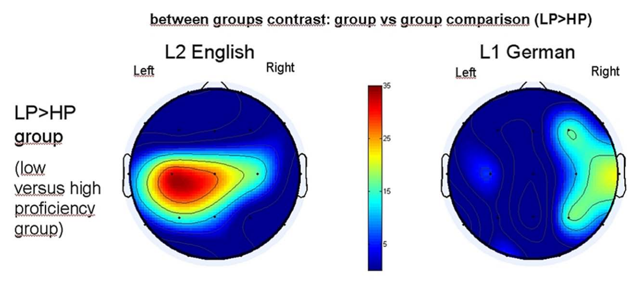


FIGURE 5 | Between groups contrast of the groups depicted in Figure 4. Shown here: LP (low proficiency group) > HP (high proficiency) group. Red color in color bar indicating which electrodes increase more their percentage of pertinence to the strongly synchronized cluster. Left panel: group contrast (indicting LP group) during L2 processing and right panel indicating L1.

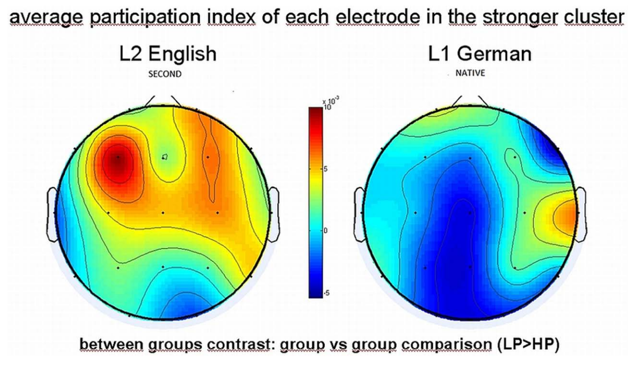


FIGURE 6 | Topography distribution brain maps of the average participation of each electrode (synchronization cluster index) in the strongly synchronized cluster within the gamma band detected by PLI. Comparison shown here: LP > HP group (group contrast, HP subtracted from LP group). Color bar (red color) indicating the greatest increment of involvement of the electrodes in left frontal/fronto-central areas during L2 (left panel) processing in the LP group.

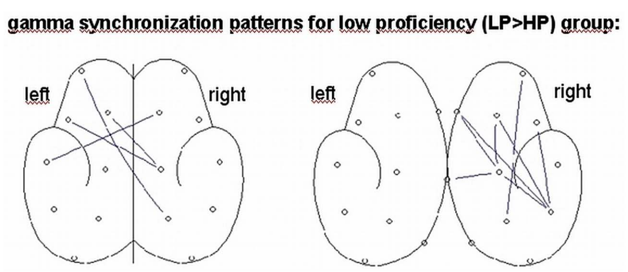


FIGURE 7 | Connectivity brain maps visualizing the most synchronized and exact electrode positions (synchronization patterns) within the strongly synchronized cluster in the gamma band for the interhemispheric connections only (left panel) and the within hemisphere (left and right) connections, for the condition L2 English in the LP > HP group.

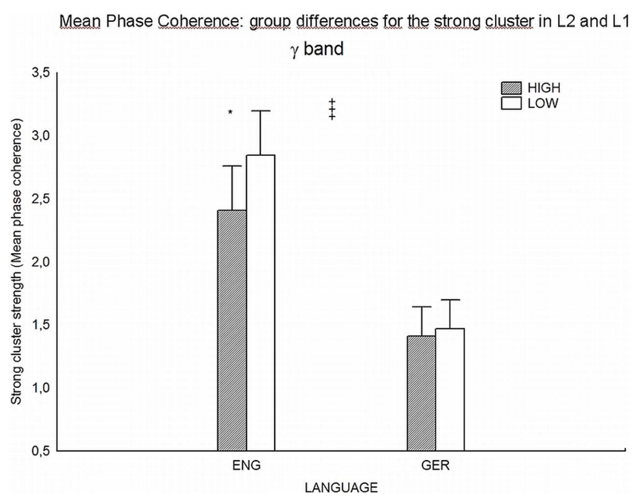


FIGURE 8 | Mean Phase Coherence: differences between the groups during L2 ("ENG," left double bar) and L1 ("GER," right double bar) processing. Dark bars indicate the high proficiency group and white bars the low proficiency group. Difference in mean phase coherence is significant ($p = 0.42$) between the groups during processing L2 English (increased synchronization strength for LP group), but not during processing L1 German.

Where the mean phase coherence yields a significant difference between the HP and LP group ($p = 0.42$) for the strongest synchronized gamma-band cluster during the processing of English L2, the same significant group difference cannot be found for native language German. This analysis corroborates the main finding for stronger synchronization in gamma band during processing of L2 for the lower proficiency group.

DISCUSSION

In this study, we primarily showed that different levels of "cortical control" or cortical processing mechanisms accompany the processing of second and first language, and furthermore, that differently proficient bilinguals can be differentiated by their cortical connectivity patterns, especially while processing their less fluent language. We suppose that ease of language processing might in part be instantiated by the brain trough different levels of synchronization between language network areas, with stronger synchronization between larger and more extended networks reflecting the recruitment of more resources (cortical effort), either because the task is generally more difficult as in the case of processing a later learned second language by late bilinguals or – on a more subtle scale – because proficiency differences due to differences in long-term language training make the task at hand more effortful and hence call for the integration of more global workspace in the brain.

Generally speaking, our results offer two insights (one more of a factual, one more of a methodological nature): first, language learners who are highly proficient in their L2 and have undergone extensive linguistic training seem to use different language processing strategies reflected in different cortical patterns on the level of synchronized electrophysiological activity in the brain, and secondly, these behavioral differences in cognitive processing can be made visible by measuring synchronized activity within the EEG gamma frequency range. Our results indicate a speculative role of gamma band as a further method to investigate the neural substrate of bilingual proficiency level.

EEG COHERENCE/SYNCHRONIZATION PATTERNS IN LATE BILINGUALS

By employing a recently derived technique of EEG global synchronization analysis (CGMC, see Data Analysis) we found pronounced differences in synchronization strength in the gamma frequency range between auditory text processing/comprehension in mother tongue and a later learned (around 9 years of age) second language. This result points to subtle differences in cortical control mechanisms at the level of interconnectedness between language areas and surrounding tissue and their connectivity patterns – possibly as a function of differences in "language entrenchment" (Mac Whinney, 2010).

Cooperative activity, interactions and communication between neuronal assemblies through coherent oscillations subserving cognitive processes have been traced and investigated with electrophysiological methods by means of coherence and synchronization analyses within various frequency ranges of the ongoing brain responses (Engel and Singer, 2001; Ward, 2003; Fries, 2005). Increased synchronization between and within frequency ranges (e.g., most prominently alpha, theta) was found to reflect increased working memory demands, short-term memory work-load, and cognitive effort (Sarnthein et al., 1998; Sauseng et al., 2005; Schack et al., 2005). Recently gamma band analyses of the human EEG have become very promising sources for gaining new insights into higher-order cognitive information processing. What formerly has been discarded, or cut-off as "noise" or contaminated EEG, is now looked upon as a valuable tool for investigating the most sophisticated mental processes including music perception (Bhattacharya and Petsche, 2005b) and artistic imagination (Bhattacharya and Petsche, 2005a). Gamma band oscillations (an indicator of local or short-range synchronization) are said to reflect gestalt perception (Kaiser and Lutzenberger, 2003) or a kind of matching process between bottom-up and top-down information (e.g., comparing memory contents with incoming stimulus related information

(Herrmann et al., 2004). Further, gamma band PS (an indicator of long-range synchronization) is thought to reflect cognitive "binding" phenomena, feature integration, STM, higher-order integrated thinking associated with quick high-density information processing, and transient associations of neural assemblies (Fell et al., 2003).

Within the domain of language, some authors have already investigated the role of gamma band oscillations in (native) language processing (Pulvermüller et al., 1997), for example, for the syntactic and semantic domain (Braeutigam et al., 2001; Micheloyannis et al., 2003; Hagoort et al., 2004; Ihara and Kakigi, 2006; Bastiaansen et al., 2010) for verbal performance and intelligence (Jausovec and Jausovec, 2005) as well as for correlations with semantic complexity (Simos et al., 2002). In addition, gamma band oscillations and synchronization phenomena have also been reported for L1 processing (Ford et al., 2005; Weiss et al., 2005; Ihara and Kakigi, 2006). For example, Ford et al. (2005) found that binding mechanisms in sentence processing were reflected in fronto-temporal gamma synchrony. Effects of sentence complexity on gamma coherence have been reported by the Weiss et al. (2005) study. Although all these EEG coherence studies were performed purely on native language processing, it seems not too far-fetched to compare L1 with L2 phenomenologically and look for bilingual language processing as well in the gamma frequency range. We would like to argue that it is only a logical consequence that bilingual language processing and cortical control mechanisms that are related to individual differences in the mastery of languages, can be revealed by adopting the above described method. We suggest that the basic mechanisms which underlie second language processing are similar to the ones in first language processing from a theoretical and empirical point of view, e.g., (Newman-Norlund et al., 2006), since it is compatible with recent brain imaging studies which find (at least partially) overlapping areas of activation for L1 and L2 (Hasegawa et al., 2002; Chee et al., 2003; Marian et al., 2003; Vingerhoets et al., 2003; Lucas et al., 2004; Ojima et al., 2005; Reiterer et al., 2005a,b, 2009; Klein et al., 2006; Gandour et al., 2007). Bilingual brain organization in terms of networks and functional connectivity has rarely been investigated so far (for EEG coherence see Reiterer et al., 2005a,b). With the method of fMRI connectivity the only studies to date are a study by Dodel et al. (2005) and Majerus et al. (2008). Both of them investigated cortical synchronization patterns by employing fMRI connectivity analyses in bilingual language processing (Dodel et al., 2005) and native language STM processing (Majerus et al., 2008). Interestingly, the Dodel et al. found bigger and more extended networks for the bilinguals with higher proficiency levels with fMRI connectivity, contrary to many studies on bilingual fluency levels, which find fewer activated areas as a function of higher fluency levels in second languages (e.g., Perani et al., 1996, 1998; Yetkin et al., 1996; Chee et al., 2001; Hasegawa et al., 2002; Briellmann et al., 2004; Xue et al., 2004; Klein et al., 2006). The very recent study by Majerus et al. (2008) could however differentiate high from low proficiency bilinguals by fMRI connectivity patterns.

DIFFERENCES IN PROFICIENCY LEVEL

However, we did not only find significant differences in connectivity strength between first and second language, but also between the different proficiency groups. This finding *per se* appears to be rather intuitive, since the participants in our study were no early bilinguals, but mixed proficiency late bilinguals who were exposed to the second language for the first time around 9 years of age. More importantly than age of onset even (Birdsong, 2006), they received most of their foreign language input through formal classroom training and very little through natural exposure in an L2 setting. Hence, we believe that the reason for this striking difference in gamma band synchronization strength (**Figure 3**) is mostly due to differences in exposure, entrenchment, and language learning methods, less to age of onset of learning the language. The recent brain imaging literature on bilingual or multilingual language learning increasingly supports the viewpoint that proficiency differences have more impact on brain organization in bilinguals than "pure" age of onset (Kotz, 2009; Reiterer, 2010). Proficiency on the other hand is a "fuzzy" term insofar as it needs to be clarified in the first place which factors led to a certain level of proficiency (be it a special long-term exposure, an early onset, an intensive training paradigm, a special aptitude or predisposition for language learning etc.). Differences in proficiency level can be reached by various different factors, or, more realistically, a combination of those.

By employing further analyses in the gamma range on the two groups we investigated (proficiency levels due to different amounts and quality of language training) we found differences in connectivity patterns reflecting the differences in level of fluency/proficiency in L2.

Differently proficient bilinguals, who had either higher or lower amounts of linguistic training and expertise in their second language, could be differentiated by their EEG network activity or synchronization patterns in the gamma frequency band by mean phase coherence analysis and topographical differences of the underlying employed networks by means of PLI. More specifically, we found that that during processing the second language, the low proficiency bilinguals, as compared to HP bilinguals, produced more strongly synchronized patterns of functional connectivity especially in left fronto-parietal areas. The low proficiency speakers seemed to recruit those areas in a concerted manner more often than their HP counterparts.

Our findings related to proficiency differences in L1 and L2 processing, are based on two results: (1) group × language analysis (**Figure 3**) mainly reflecting differences of the two language systems (L1 and L2) on a cortical processing level (increased synchronization strength in L2 for both groups), and (2) within and between groups analysis (**Figures 4–7**) indicating that linguistic training can alter L2 processing demands on a cortical as well as on a behavioral level and less proficient second language users have to recruit broader language networks (in left fronto-parietal areas) more strongly (with higher connection strength).

BILINGUAL LANGUAGE PROCESSING AND THEORIES OF CORTICAL EFFICIENCY

What has been observed many times in various domains outside of language processing that increased cognitive demands are accompanied by increased activity levels or extended area recruitment (e.g., for intelligence see Haier et al., 1988; Haier et al., 1992; Grabner et al., 2006, for music processing: Lotze et al., 2003, for working memory: Sarnthein et al., 1998; Sauseng et al., 2005), has also been reported for L1 as well as L2 processing (Raichle et al., 1994; Just et al., 1996; Yetkin et al., 1996; Perani et al., 1998; Rypma and D'Esposito, 1999; Dräger et al., 2004; Xue et al., 2004; Reiterer et al., 2005a,b; Abutalebi, 2008; Kotz, 2009; Leonard et al., 2011). What has been explicitly called "cortical efficiency" could be termed "proficiency level differences" in the field of L2 processing. With native language processing several authors found that comparable to a "compensation mechanism," brain activation increases with the complexity in linguistic processing (Just et al., 1996; Rypma and D'Esposito, 1999; Dräger et al., 2004) or reduces with increased repetition and practice (Raichle et al., 1994; Thompson-Schill et al., 1999). In the field of second language processing as well, various research groups have detected the "cortical efficiency" phenomena by revealing that proficiency level (either attained by practice, higher exposure or by formal training or as occurring more naturally by both) has an influence on the extent and intensity of cortical activation in a bilingual's brain (Yetkin et al., 1996; Perani et al., 1998; Chee et al., 2001; Hasegawa et al., 2002; Wartenburger et al., 2003; Briellmann et al., 2004; Xue et al., 2004; Zhang et al., 2005; Klein et al., 2006). The usual observation can be summarized in the following terms: lower proficiency, more distributed activity (i.e., a Larger network) and higher proficiency, more focal activity (smaller network). Our results of higher gamma band long-range synchronization in L2 going hand in hand with lower proficiency level in bilinguals and lower gamma synchronization with higher proficiency level, are pointing into the same direction, possibly revealing a compensation mechanism in the domain of second language processing.

LIMITATIONS OF THE STUDY

The current study also has a few limitations. No verbal or non-verbal IQ test was performed as control, because of limitations of time and laboratory use and the theoretical consideration that language abilities do not correlate with non-verbal intelligence. We want to point to this shortcoming and are aware that this might limit the interpretation of our results. It the same vein it needs to be mentioned that we carefully chose two groups with a closely matching educational level, preselected by university exams (participants were all students with completed Bachelor's degree, studying for a Master's) and this might enhance group similarity with respect to higher cognitive and intellectual abilities.

Finally, we would like to mention one additional point in here, which could be regarded as limitation or as interesting outcome likewise. This is the behavioral result that our two groups behaved slightly differently already when tested (comprehension questions) in their mother tongue. The low proficiency group with regard to L2 scored slightly worse when tested on L1 comprehension and text recall. This we like to call "the L1 paradox." Usually the L1 is implemented as a control condition where the groups should behave in exactly the same way, because native speakers are perceived as a "homogeneous mass." This is the classical intuition, but our data as well as other research (Pakulak and Neville, 2010; Reiterer et al., 2011) show that also mother tongue speakers can differ in their L1 proficiency and competence levels (a fact which is also affirmed by the existence of congenital language disorders). Variation within L1 competence might be smaller than within L2 interlanguages, but nevertheless existing. Recent research shows that also smaller differences in L1 competence and ability levels can be traced by brain imaging techniques and detected in such a way which would not have been possible with pure behavioral measurements. Even in the case of our present study, the result of the L1 differences behaviorally was just a marginal one, statistically speaking only a "trend." Such small differences in mother tongue processing can, however, point to important underlying principles. What they show is individual differences in the linguistic abilities of L1 speakers. One of the theoretical concepts that tries to capture this phenomenon is language aptitude. Theoretical assumptions and new brain research data likewise (Wells, 1985; Abrahamsson and Hyltenstam, 2008; Golestani et al., 2011; Reiterer et al., 2011) show that the variable of general language aptitude is a possible hidden driving force behind individual differences in L1 as well as L2 proficiency and ability levels. Very often, this variable is neglected in the whole field of bilingualism research, be it behavioral or neurocognitive. What we could have hit upon in our present study by detecting those small L1 trend differences, is pre-existing differences between the groups in language aptitude. General language aptitude might also drive and determine career and study choice. In our case here aptitude might have partly driven the language students to study foreign languages and linguistics (our HP group). We acknowledge that it is very difficult to "control" for all these pre-existing variables, but at the same time it needs to be said that they might be very important in explaining a lot of variance in bilingual data. Thus, what is missing in many studies of bilinguals or/and second language learners (including the present study) is a sound testing for individual differences in language aptitude. This is a methodological as well as a theoretical issue and an important point to be considered in future studies.

CONCLUSION

We have shown here that by looking at EEG gamma band phase synchronization patterns, one can differentiate second (classroom-learned) from first language, and within the later learned second language (L2) learners with lower amounts of linguistic training and expertise from those with higher amounts of expertise by the different ways by which they employ synchronized activation. The observed patterns could be explained by the theory of cortical efficiency because we found different network patterns for high and low proficiency learners, with more widely distributed synchronized networks in left fronto-parietal areas more often recruited by lower proficiency learners. Our findings further indicate that EEG gamma band phase synchronization measures are sensitive to differences in second language processing and control strategies due to experience/proficiency-driven differences.

ACKNOWLEDGMENTS

Joydeep Bhattacharya is supported by JST.ERATO project. The author Susanne Reiterer is supported by the German Research Foundation (DFG, project AC-55/7-1).

REFERENCES

Abrahamsson, N., and Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Stud. Sec. Lang. Acquis.* 30, 481–509.

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychol. (Amst)* 128, 466–478.

Allefeld, C. (2006). About the derivation of the SCA algorithm. *Int. J. Bifurcat. Chaos* 16, 3705–3706.

Allefeld, C., and Bialonski, S. (2007). Detecting synchronization clusters in multivariate time series via coarse-graining of Markov chains. *Phys. Rev. E* 76, 066207.

Allefeld, C., Frisch, S., and Schlesewsky, M. (2005). Detection of early cognitive processing by event-related phase synchronization analysis. *Neuroreport* 16, 13–16.

Allefeld, C., and Kurths, J. (2004). An approach to multivariate phase synchronization analysis and its application to event-related potentials: synchronization cluster analysis. *Int. J. Bifurcat. Chaos* 14, 417–426.

Bastiaansen, M., and Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Prog. Brain Res.* 159, 179–196.

Bastiaansen, M., Magyari, L., and Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during online sentence comprehension. *J. Cogn. Neurosci.* 22, 1333–1247.

Bhattacharya, J., Pereda, E., and Petsche, H. (2003). Effective detection of coupling in short and noisy bivariate data. *IEEE Trans. Syst. Man Cybern. B* 33, 85–95.

Bhattacharya, J., and Petsche, H. (2005a). Drawing on mind's canvas: differences in cortical integration patterns between artists and non-artists. *Hum. Brain Mapp.* 26, 1–14.

Bhattacharya, J., and Petsche, H. (2005b). Phase synchrony analysis of EEG during music perception reveals changes in functional connectivity due to musical expertise. *Signal Process.* 85, 2161–2177.

Bhattacharya, J., Petsche, H., and Pereda, E. (2001). Long-range synchrony in the gamma band: role in music perception. *J. Neurosci.* 21, 6329–6337.

Birdsong, D. (2006). Age and second language acquisition and processing: a selective overview. *Lang. Learn.* 56, 9–49.

Braeutigam, S., Bailey, A. J., and Swithenby, S. J. (2001). Phase-locked gamma band responses to semantic violation stimuli. *Brain Res. Cogn. Brain Res.* 10, 365–377.

Briellmann, R. S., Saling, M. M., Connell, A. B., Waites, A. B., Abbott, D. F., and Jackson, G. D. (2004). A high-field functional MRI study of quadri-lingual participants. *Brain Lang.* 89, 531–542.

Chee, M. W., Hon, N., Lee, H. L., and Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. Blood oxygen level dependent. *Neuroimage* 13, 1155–1163.

Chee, M. W., Soon, C. S., and Lee, H. L. (2003). Common and segregated neuronal networks for different languages revealed using functional magnetic resonance adaptation. *J. Cogn. Neurosci.* 15, 85–97.

De Bot, K. (2008). The imaging of what in the multilingual mind? Review article. *Sec. Lang. Res.* 24, 111–133.

Dodel, S., Golestani, N., Pallier, C., Elkouby, V., Le Bihan, D., and Poline, J. B. (2005). Condition-dependent functional connectivity: syntax networks in bilinguals. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 360, 921–935.

Dräger, B., Jansen, A., Bruchmann, S., Forster, A. F., Pleger, B., Zwitserlood, P., and Knecht, S. (2004). How does the brain accommodate to increased task difficulty in word finding? A functional MRI study. *Neuroimage* 23, 1152–1160.

Engel, A., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci. (Regul. Ed.)* 5, 16–25.

Essl, M., and Rappelsberger, P. (1998). EEG Coherence and reference signals: experimental results and mathematical explanations. *Med. Biol. Eng. Comput.* 36, 399–406.

Fell, J., Fernandez, G., Klaver, P., Elger, C. E., and Fries, P. (2003). Is synchronized neuronal gamma activity relevant for selective attention? Brain Res. *Brain Res. Rev.* 42, 265–272.

Ford, J. M., Gray, M., Faustman, W. O., Heinks, T. H., and Mathalon, D. H. (2005). Reduced gamma-band coherence to distorted feedback during speech when what you say is not what you hear. *Int. J. Psychophysiol.* 57, 143–150.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci. (Regul. Ed.)* 9, 474–480.

Gandour, J., Yunxia, T., Talavage, T., Wong, D., Dzemidzic, M., Yisheng, X., Xiaojian, L., and Lowe, M. (2007). Neural basis of first and second language processing of sentence-level linguistic prosody. *Hum. Brain Mapp.* 28, 94–108.

Golestani, N., Price, C. J., and Scott, S. K. (2011). Born with and ear for dialects? Structural plasticity in the expert phonetician brain. *J. Neurosci.* 31, 4213–4220.

Grabner, R. H., Neubauer, A. C., and Stern, E. (2006). Superior performance and neural efficiency: the impact of intelligence and expertise. *Brain Res. Bull.* 69, 422–439.

Hagoort, P., Hald, L., Bastiaansen, M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441.

Haier, R., Siegel, B., Nuechterlein, K., Hazlett, E., Wu, J., Peak, J., Browning, H. L., and Buchsbaum, M. S. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with PET. *Intelligence* 11, 199–218.

Haier, R., Siegel, B., Tang, C. H., Abel, L., and Buchsbaum, M. (1992). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence* 16, 415–426.

Hald, L. A., Bastiaansen, M. C., and Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain Lang.* 96, 90–105.

Hasegawa, M., Carpenter, P. A., and Just, M. A. (2002). An fMRI study of bilingual sentence comprehension and workload. *Neuroimage* 15, 647–660.

Herrmann, C., Munk, M., and Engel, A. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn. Sci. (Regul. Ed.)* 8, 347–355.

Hoke, M., Lehnertz, K., Pantev, C., and Lütkenhöner, B. (1989). "Spatiotemporal aspects of synergetic processes in the auditory cortex as revealed by magnetoencephalogram," in *Series in Brain Dynamics*, eds E. Basar and T. H. Bullock, (Berlin: Springer), 84–105.

Ihara, A., and Kakigi, R. (2006). Oscillatory activity in the occipitotemporal area related to the visual perception of letters of a first/second language and pseudoletters. *Neuroimage* 29, 789–796.

Jasper, H. H. (1958). The ten/twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.

Jausovec, N., and Jausovec, K. (2005). Differences in induced gamma and upper alpha oscillations in the human brain related to verbal/performance and emotional intelligence. *Int. J. Psychophysiol.* 56, 223–235.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., and Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science* 274, 114–116.

Kaiser, J., and Lutzenberger, W. (2003). Induced gamma-band activity and human brain function. *Neuroscientist* 9, 475–484.

Klein, D., Watkins, K. E., Zatorre, R. J., and Milner, B. (2006). Word and nonword repetition in bilingual participants: a PET study. *Hum. Brain Mapp.* 27, 153–161.

Kotz, S. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain Lang.* 109, 68–74.

Le Van Quyen, M., and Bragin, A. (2007). Analysis of dynamic brain oscillations: methodological advances. *Trends Neurosci.* 30, 365–373.

Leonard, M., Torres, C., Travis, K., Brown, T., Hagler, D., Dale, A., Elman, J., and Halgren, E. (2011). Language proficiency modulates the recruitment of non-classical language areas in bilinguals. *PLoS ONE* 6, e18240. doi:10.1371/journal.pone.0018240

Lotze, M., Scheler, G., Tan, H. R., Braun, C., and Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage* 20, 1817–1829.

Lucas, T. H., McKhann, G., and Ojemann, G. (2004). Functional separation of languages in the bilingual brain: a comparison of electrical stimulation language mapping in 25 bilingual patients and 117 monolingual control patients. *J. Neurosurg.* 101, 449–457.

Mac Whinney, B. (2010). "A tale of two paradigms," in *Language Acquisition across Linguistic and Cognitive Systems*, eds M. Kail and M. Hickmann (Amsterdam: John Benjamins), 17–33.

Majerus, S., Belayachi, S., De Smedt, B., Leclercq, A. L., Martinez, T., Schmidt, C., Weekes, B., and Maquet, P. (2008). Neural Networsk for short-term memory for order differentiate high and low proficiency bilinguals. *Neuroimage* 42, 1698–1713.

Marian, V., Spivey, M., and Hirsch, J. (2003).Shared and separate systems in bilingual language processing: converging evidence from eye-tracking and brain imaging. *Brain Lang.* 86, 70–82.

Micheloyannis, S., Vourkas, M., Bizas, M., Simos, P., and Stam, C. J. (2003). Changes in linear and nonlinear EEG measures as a function of task complexity: evidence for local and distant signal synchronization. *Brain Topogr.* 15, 239–347.

Mormann, F., Lehnertz, K., David, P., and Elger, C. E. (2000). Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. *Physica D* 144, 358–369.

Newman-Norlund, R. D., Frey, S. H., Petitto, L. A., and Grafton, S. T. (2006). Anatomical substrates of visual and auditory miniature second-language learning. *J. Cogn. Neurosci.* 18, 1984–1997.

Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., and Hallet, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clin. Neurophysiol.* 115, 2292–2307.

Ojima, S., Nakata, H., and Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *J. Cogn. Neurosci.* 17, 1212–1228.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Pakulak, E., and Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *J. Cogn. Neurosci.* 22, 2728–2744.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F., and Mehler, J. (1996). Brain processing of native and foreign languages. *Neuroreport* 7, 2439–2444.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S. F., Fazio, F., and Mehler, J. (1998). The bilingual brain. Proficiency and age of acquisition of the secon language. *Brain* 121, 1841–1852.

Pereda, E., Quian Quiroga, R., and Bhattacharya, J. (2005). Nonlinear multivariate analysis of neurophysiological signals. *Prog. Neurobiol.* 77, 1–37.

Pulvermüller, F., Birbaumer, N., Lutzenberger, W., and Mohr, B. (1997). High-frequency brain activity: its possible role in attention, perception and language processing. *Prog. Neurobiol.* 52, 427–445.

Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M., Pardo, J. V., Fox, P. T., and Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cereb. Cortex* 4, 8–26.

Reiterer, S. (2010). "The cognitive neuroscience of second language acquisition and bilingualism: factors that matter in L2 acquisition - a neurocognitive perspective," in *Language Acquisition across Linguistic and Cognitive Systems*, eds M. Kail and M. Hickmann (Amsterdam: John Benjamins), 307–321.

Reiterer, S., Hemmelmann, C., Rappelsberger, P., and Berger, M. L. (2005a). Characteristic functional networks in high- versus low-proficiency second language speakers detected also during native language processing: an explorative EEG coherence study in 6 frequency bands. *Brain Res. Cogn. Brain Res.* 25, 566–578.

Reiterer, S., Berger, M. L., Hemmelmann, C., and Rappelsberger, P. (2005b). Decreased coherence between prefrontal electrodes: a correlate of high language proficiency? *Exp. Brain Res.* 163, 109–113.

Reiterer, S., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., Winkler, S., and Ackermann, H. (2011). Individual differences in speech imitation/pronunciation aptitude in late bilinguals: functional neuroimaging and brain morphology. *Front. Psychol.* 2:271, 1–12.

Reiterer, S., Pereda, E., and Bhattacharya, J. (2009). Measuring second language proficiency with EEG synchronization: how functional cortical networks and hemispheric involvement differ as a function of proficiency level in second language speakers. *J. Sec. Lang. Res.* 25, 77–106.

Rypma, B., and D'Esposito, M. (1999). The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. *Proc. Natl. Acad. Sci. U.S.A.* 96, 6558–6563.

Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., and von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7092–7096.

Sauseng, P., Klimesch, W., Schabus, M., and Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *Int. J. Psychophysiol.* 57, 97–103.

Schack, B., Klimesch, W., and Sauseng, P. (2005). Phase synchronization between theta and upper alpha oscillations in a working memory task. *Int. J. Psychophysiol.* 57, 105–114.

Simos, P. G., Papanikolaou, E., Sakkalis, E., and Micheloyannis, S. (2002). Modulation of gamma-band spectral power by cognitive task complexity. *Brain Topogr.* 14, 191–196.

Stam, C. J. (2005). Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin. Neurophysiol.* 116, 2266–2301.

Stam, C. J., Nolte, G., and Daffertshofer, A. (2007). Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Hum. Brain Mapp.* 28, 1178–1193.

Thompson-Schill, S. L., D'Esposito, M., and Kan, I. P. (1999). Effects of repetition and competition on activity in left prefrontal cortex during word generation. *Neuron* 23, 513–522.

Vingerhoets, G., Van Borsel, J., Tesink, C., Deblaere, K., Seurinck, R., Vandemaele, P., and Achten, E. (2003). Multilingualism: an fMRI study. *Neuroimage* 20, 2181–2196.

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends Cogn. Sci. (Regul. Ed.)* 7, 553–559.

Wartenburger, I., Heereken, H., Abutalebi, J., Cappa, S., Villringer, A., and Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–170.

Weiss, S., Mueller, H. M., Schack, B., King, J. W., Kutas, M., and Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *Int. J. Psychophysiol.* 57, 129–141.

Wells, G. (1985). *Language Development in the Pre-School Years.* Cambridge: Cambridge University Press.

Xue, G., Dong, Q., Jin, Z., and Chen, C. (2004). Mapping of verbal working memory in nonfluent Chinese-English bilinguals with functional MRI. *Neuroimage* 22, 1–10.

Yetkin, O., Zerrin, T., Yetkin, F., Haughton, V., and Cox, R. W. (1996). Use of functional MR to map language in multilingual volunteers. *AJNR Am. J. Neuroradiol.* 17, 473–477.

Zhang, Y., Kuhl, P. K., Imada, T., Kotani, M., and Tohkura, Y. (2005). Effects of language experience: neural commitment to language-specific auditory patterns. *Neuroimage* 26, 703–720.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 April 2011; paper pending published: 11 June 2011; accepted: 29 October 2011; published online: 22 November 2011.

*Citation: Reiterer S, Pereda E and Bhattacharya J (2011) On a possible relationship between linguistic expertise and EEG gamma band phase synchrony. Front. Psychology **2**:334. doi: 10.3389/fpsyg.2011.00334*

This article was submitted to Frontiers in Cognition, a specialty of Frontiers in Psychology.