The role of executive control in bilingual language production and reading

Irina Pivneva

Department of Psychology, Faculty of Sciences
Centre for Research on Brain, Language and Music

McGill University
Montreal, Quebec, Canada
September 2014

A thesis submitted to McGill University
In partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Experimental Psychology

©Copyright 2014 All rights reserved
Table of Contents

Acknowledgement .................................................................................................................................................. viii
Abstract .............................................................................................................................................................. xi
Abrégé .................................................................................................................................................................. xiii
Funding ............................................................................................................................................................... xvi
Statement of Originality ...................................................................................................................................... xvi
Contribution of Authors .................................................................................................................................. xviii
Chapter 1: General Introduction .......................................................................................................................... 21
  Definition of executive control used in this thesis ............................................................................................. 28
  Definition of bilingualism used in this thesis ....................................................................................................... 31
  Bilingual language production processes ......................................................................................................... 34
  Models of bilingual language production ......................................................................................................... 38
    Mental firewall accounts of bilingual language production ......................................................................... 40
    Inhibitory accounts of bilingual language production ................................................................................. 45
  Bilingual language production studies ........................................................................................................... 48
    Picture-word interference paradigm ............................................................................................................. 49
    Language-switching paradigm ....................................................................................................................... 51
  Executive control as a candidate for language control ..................................................................................... 55
  The present work ................................................................................................................................................. 56

Chapter 2:
Inhibitory control and L2 proficiency modulate bilingual language production: Evidence from spontaneous monologue and dialogue speech (Pivneva, Palmer, & Titone, 2012, Frontiers in Psychology) ................................................................................................................................. 61
Effects of name agreement, L1 vs. L2 production, and pure vs. mixed language blocks .......................................................... 124

Effects of individual differences in executive control .................................................. 127

Discussion ..................................................................................................................... 133

Experiment 2: Single picture naming ........................................................................ 136

Methods ......................................................................................................................... 136

Participants ................................................................................................................... 136

Materials ......................................................................................................................... 137

Procedure ....................................................................................................................... 138

Results ............................................................................................................................. 139

Effects of name agreement, L1 vs. L2 production, and pure vs. mixed language blocks .................................................................................................................. 142

Effects of individual differences in executive control .................................................. 146

Discussion ..................................................................................................................... 147

General discussion ........................................................................................................ 148

Preface to Chapter 4 ...................................................................................................... 159

Chapter 4:

Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements (Pivneva, Mercier, & Titone, 2014, Journal of Experimental Psychology: Learning, Memory, and Cognition) .................................................................................................... 165

Abstract ......................................................................................................................... 166

Methods ......................................................................................................................... 170
Appendix B. A transcribed example of produced conversational speech from Chapter 2 ......277

Appendix C. A graphic depiction of a typical acoustic output produced in Chapter 2...........278

Appendix D. A graphic representation of the anti-saccade task from Chapters 2 and 4...........279

Appendix E. A graphic representation of the non-linguistic Simon and Stroop tasks from Chapters 2-4 ........................................................................................................................................280

Appendix F. A graphic representation of the Number Stroop task from Chapters 2 and 4 .....281

Appendix G. Example stimuli for Chapter 3........................................................................282
Dedication

To my mother who instilled in me the love of learning.

To Michael who instilled in me the love of argument.

To my children who instilled in me the love of patience.
Acknowledgement

I would like to thank a small army of people without whom my tenure at McGill University and, in part, this thesis would not have come to fruition. First and foremost, I would like to acknowledge Dr. Debra Titone. Your dedication to research and to your students is exceptional. Thank you for investing hard work, time, and energy to help and guide me through academic and personal development labyrinth. A special thanks for helping me to identify and to work on my weaknesses from early on. Thank you for countless times when you uplifted my “fallen” research spirits: be it interpreting complex result patterns or resurrecting a manuscript draft. Thank you for the multitude of networking opportunities at all conferences, workshops, and gatherings. Especially, thank you for teaching me how to strategize with respect to projects and ever-fleeting time. You have truly been my working model in how to compliment and balance research, and academic and service appointments, all while raising a young family.

Among graduate committee members, I would like to thank Fred Genesee who has been the cornerstone of my committee from the start. Thank you, Dr. Genesee, for your feedback and the endless letters of support. I would like to thank Dr. Kris Onishi for her feedback on my presentations and my manuscripts, and letters of support. I also would like to thank Drs. Meghan Clayards, Yuriko Oshima-Takane, and Shari Baum for their thoughtful feedback on my manuscripts. In addition, I would like to thank Drs. Caroline Palmer and Jelena Ristic for their valuable feedback on my presentations. Thank you, Drs. Yoshio Takane and Heungsun Hwang, for helping me navigate through novel and powerful statistical methods. Lastly, a special thanks to my former undergraduate supervisor, Dr. David Forman; without whose early guidance I would not have learned how to take charge of research projects in a structured way.
Within the McGill Language and Multilingualism Laboratory, I would like to thank all past and present members for their help, feedback, and support. Special thanks to Naveed Sheikh and Kyle Lovseth for your perky feedback and willingness to open your mind to digressions in research and popular culture. Thank you, Veronica Whitford, for listening to my rants and for the help and support with students and frequent emergencies. Thank you, Dr. Julie Mercier, for opening your project to me in my first year and for always keeping a balanced outlook on life. Thank you to Dr. Maya Libben, to whom I looked up as a newly-minted graduate student. A special thanks to Drs. Georgie Columbus and Annie Gilbert, who have supported and guided me during the most challenging times at McGill. Most importantly, thank you to all undergraduate research assistants without whom the work reported in this thesis would not have been possible. I would like to thank Giselle Lehman, Sonia Kroll, Stacey Wilson, Emily Quinan, Erin Luxenberg, Alissa Yip, Deborah Baremberg, Vivienne Walz and Nabeel Rkieh for their help with testing, transcribing, conversation cleaning and coding. Thank you, Emily Delpero, for being my first honor research student, for your brilliant questions, and for proudly claiming me as “your graduate student.” Thank you, Yasmine Khawajkie, for the help not only with research projects but also thinking of me and my family in your travels. A special thank you to Abigail Free for seeking my advice in academic and personal matters and for your unparalleled help with projects, transcription and student management. Most importantly, the largest thanks is to 18 students volunteers who have spent countless hours transcribing and coding production data, especially Verona Soliman-Saad, Emilie Bolduc, and Catherine Degre for going above and beyond their duties.

I would like to thank the Auditory Cognitive Neuroscience CREATE program and Dr. Caroline Palmer for providing me with ample networking and training opportunities, and
especially an internship with Dr. Sylvain Baillet at the Montreal Neurological Institute (MNI) to extend my knowledge to neuroimaging methods. Thank you, Dr. Baillet, for welcoming me to your laboratory to learn magnetoencephalography to study bilingual language processing. A special thank you to MEG lab members, namely Elizabeth Bock, Dr. Benjamin Morillon, Maryse Thomas, and Sebastien Dery, for their feedback, support, and help with my projects.

To my friends, whom I do not see often, thank you for your support and encouragement. Especially, thank you, Dana Hayward, for your help, support, and advice within the last two years. I admire your tenacity and brilliance. Thank you to my family near and far for supporting me and challenging me to defend my choices. Thank you to my younger brother, who has trailed with me across the globe. Thank you to my late father who taught me from an early age how to strategize and negotiate. Special thanks to my older sister, Elena, who has always served as my guiding principle with her unsurpassing wisdom and wit. I have always looked up to you. Especially, thank you to my eldest niece, Ekaterina, whose advice has been wise beyond her years. A very special thanks to my daughters, Kayden-Alexandra and Jayden-Alice, who have supplied and helped to pilot endless research ideas, and supported and loved me unconditionally. Another special thanks to Michael for doing more than his “fair share” and supporting me with actions if not with words. Without your presence and support none of this would have happened. The biggest thank you is reserved for my dear mother, who from an early age has encouraged and supported the search for knowledge, despite the limited resources and unfavorable third-world conditions. Thank you for your sacrifices, your tenacity, and all of your kept promises.
Abstract

This thesis investigates the link between domain-general cognitive system and bilingual language processing system, specifically linking executive control to bilingual language production and reading across varied communicative contexts. A fundamental question in bilingual language processing is whether language processing system is solely responsible for linguistic operations involved in language production and reading or whether it recruits, on demand, domain-general cognitive resources. Some bilingual language processing accounts (e.g., “mental firewall” accounts) posit that language processing system is architecturally blocked from the domain-general cognitive system, especially during earliest stages of processing. Contrastingly, inhibitory accounts posit that language processing system actively recruits domain-general cognitive resources, especially when task demands increase and often during earliest stages of processing. Effectively, we investigated the link between domain-general, non-linguistic executive control among bilinguals and bilingual language production in spontaneously produced monologues and dialogues (Chapter 2); in sentence (Chapter 3, Experiment 1) and single word production (Chapter 3, Experiment 2) across single- and dual-language contexts; and sentence reading (Chapter 4).

In Chapter 2 (Pivneva, Palmer, Titone, 2012), we found that individual differences in executive control related to acoustic indices but not to content indices of bilingual language production. Specifically, greater executive control related to more efficient ongoing speech planning and production processes, especially in cognitively more demanding L2 dialogues.

In Chapter 3 (Pivneva & Titone, resubmitted), the findings were two-fold. First, bilinguals recruited executive control to a greater extent when producing grammatically well-formed sentences in Experiment 1 versus single words in Experiment 2 and also when speaking
in single- versus dual-language contexts. Second, greater executive control related to greater
global versus local language control in sentence production in Experiment 1. Global language
control often indexes early cross-language attenuation, while local language control indexes later
attenuation of specific words.

In Chapter 4 (Pivneva, Mercier, & Titone, 2014), we found that greater executive control
related to reduced early-stage cross-language activation, as indexed by semantically conflicting
interlingual homographs (e.g., *chat* a casual conversation in English; French for *cat*) versus
semantically congruent cognates (e.g., *piano* across English and French). Additionally, greater
executive control guided enhanced later-stage use of sentence constraint when reading sentences
with language-unique control words.

Thus, bilinguals recruit executive control when processing demands increase, such as
conversing in L2, or producing and reading L2 sentences. Bilinguals also recruit executive
control, on demand, across early and late stages of language processing. Taken together, the
results reported here support active ongoing interactions between domain-general cognitive and
language processing systems throughout language processing stages and especially when task
demands increase. Lastly, the results of this thesis are discussed with respect to current
directions in bilingual language processing, general associations between bilingualism and
cognition, and directions for future research.
La présente thèse porte sur le lien entre les contraintes cognitives générales et le traitement du langage chez le locuteur bilingue, notamment par rapport à la relation entre le contrôle exécutif et la production du langage dans différents contextes communicatifs, ainsi que la lecture. Dans ce champ de recherche, une question fondamentale demeure à savoir si les processus de production du langage et de lecture peuvent relever uniquement du système de traitement langagier ou s’ils dépendent aussi des ressources cognitives générales. Par exemple, certaines approches (ex. : pare-feu mental ou « mental firewall ») suggèrent que le système de traitement du langage est structurellement indépendant du système cognitif général, surtout durant les premières étapes de traitement du langage. À l’opposer, d’autres approches dites « inhibitrices » avancent plutôt que le système de traitement langagier fait activement appel aux ressources cognitives générales du locuteur, notamment lors de tâches difficiles ou durant les premières phases de traitement. Le présent projet doctoral porte donc sur le lien entre les contraintes cognitives générales, par l’entremise du contrôle exécutif extralinguistique, et la production orale spontanée chez le locuteur bilingue dans divers contextes communicatifs, soit : la production de monologues et de dialogues (Chapitre 2), d’énoncés grammaticaux (Chapitre 3, Expérience 1), de mots isolés (Chapitre 3, Expérience 2) et la lecture d’énoncés en contexte monolingue et bilingue (Chapitre 4).

L’étude présentée au Chapitre 2 (Pivneva, Palmer, Titone, 2012) a permis d’établir que le contrôle exécutif ne semble pas avoir d’impact sur le contenu linguistique produit par le locuteur bilingue, mais influence plutôt sa réalisation des indices acoustiques impliqués. Ainsi, un meilleur contrôle exécutif favorise une meilleure
planification en continu de la production, de même qu’une exécution plus précise, particulièrement en contexte de dialogue en langue seconde, une tâche considérée exigeante du point de vue cognitif.

L’étude présentée au Chapitre 3 (Pivneva & Titone, sous révision) a permis de démontrer deux points différents. Tout d’abord, que les locuteurs bilingues modulent leur recours au contrôle exécutif selon la tâche à exécuter. Ainsi, la production d’énoncés grammaticaux complets (Expérience 1) requiert une plus grande implication du contrôle exécutif que la production de mots isolés (Expérience 2), de même que la production en contexte monolingue requiert un plus grand recours au contrôle exécutif que la production en contexte bilingue. De plus, cette étude a aussi permis de démontrer que le fait de bénéficier d’un contrôle exécutif supérieur favorise le contrôle linguistique global au détriment du contrôle linguistique local. Le contrôle linguistique dit « global » est souvent associé à la désactivation de la langue concurrente dans son ensemble durant les premières étapes de traitement du stimulus langagier alors que le contrôle dit « local » est associé au blocage de mots spécifiques.

L’étude présentée au Chapitre 4 (Pivneva, Mercier, & Titone, 2014), démontre quant à elle qu’un meilleur contrôle exécutif chez le locuteur bilingue entraîne une réduction de l’activation interlinguistique durant les premières étapes de la lecture. Ces résultats ont été obtenus à l’aide d’une tâche de lecture de mots partageant une même orthographe en français et en anglais (homographes interlinguistiques) et dont le sens était soit le même dans les deux langues (p. ex. : « piano » réfère au même instrument de musique dans les deux langues) ou différent (p. ex. : « chat » en français réfère à un animal alors qu’en anglais il réfère à une conversation informelle). Qui plus est, les
locuteurs bénéficiant d’un meilleur contrôle exécutif étaient aussi mieux en mesure d’utiliser les contraintes phrastiques durant les étapes ultérieures de la lecture d’énoncés ne contenant pas d’homographes interlinguistiques. Donc, les locuteurs bilingues font appel au contrôle exécutif lorsque le niveau de difficulté augmente comme lors de conversations en langue seconde, ou lors de la production spontanée et la lecture de phrases en langue seconde. Les locuteurs bilingues sont aussi en mesure de mobiliser le contrôle exécutif au besoin, à diverses étapes du traitement langagier.

Dans leur ensemble, les résultats rapportés dans la présente thèse démontrent l’interaction constante des contraintes cognitives générales avec les systèmes de traitement du langage, et ce particulièrement durant les tâches plus exigeantes.

Enfin, l’ensemble des résultats obtenus est mis en relation avec les approches théoriques actuelles sur le traitement du langage chez le bilingue situées dans le cadre plus général de l’étude des liens entre le bilinguisme et la cognition. Finalement, nous présentons diverses pistes de recherche complémentaires.
Funding

This research was made possible by the generous support from various funding sources. The studies presented in this thesis were generally supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada Discovery Award (awarded to Titone). The authors also gratefully acknowledge additional support from the Canada Research Chairs (CRC) Program (awarded to Palmer and Titone). Additional support was provided by the Frederick Banting and Charles Best Canada Graduate Scholarship (CGS) Master's Award from the Canadian Institute for Health Research (CIHR) (awarded to Pivneva), by the Postgraduate Scholarships (PGS) Award from the Fonds de Recherche du Quebec sur Nature et Technologies (FQRNT) (Pivneva), the NSERC Postgraduate Scholarships – Doctoral (PGS D) Award (Pivneva), and by the CIHR CGS Doctoral Award (Mercier).

Statement of originality

Research presented in this thesis investigated interactions between cognition and bilingual language processing.

Chapter 2 describes research on how individual differences in executive control and L2 proficiency relate to the efficiency with which bilinguals produce speech in spontaneous bilingual monologues and dialogues. These results extend findings from prior single-word production studies showing that greater executive control relates to more efficient bilingual production. These data were published in 2012 in a manuscript as part of a special issue on Bilingualism and Cognitive control in the journal Frontiers in Psychology.
Chapter 3 describes research on how individual differences in executive control relate to production of grammatically well-formed sentences in Experiment 1 and single words in Experiment 2 in single- and dual-language contexts. The results of the study reported in Chapter 3 demonstrate that executive control is involved to a greater extent in sentence but not single-word production. It is also important for global language control rather than local language control. These results replicate findings reported in Chapter 2 and extend results from single word production studies using eye movement measures of bilingual language production, a paradigm previously used only in monolingual language production studies. The manuscript based on these data has been revised and resubmitted to Cognition in January 2014, following an editorial invitation to resubmit as a full manuscript instead of a short report.

Chapter 4 describes research on how individual differences in executive control relate to bilingual sentence reading. The results of the study reported in Chapter 4 demonstrate that executive control significantly reduces cross-language activation during early stages of processing and especially in the presence of semantic conflict. Executive control also relates to enhanced use of contextual information during second language reading. These results extend prior work on the role of executive control in monolingual sentence reading and context processing and prior studies with respect to general context processing in second language reading. These data were published in a manuscript that appear online in January 2014 ahead of print May 1, 2014 issue, in the journal of Journal of Experimental Psychology: Learning, Memory, and Cognition.
Contribution of authors

Chapter 2:

Irina Pivneva (first author) – With respect to responsibilities, as a first author I conceptualized and designed the study. I also conducted background research; selected and adapted stimuli to bilingual production; programmed the experiment; trained research assistants on experimental protocol. With respect to participant recruitment and testing, I actively recruited participants and participated as the confederate in all experimental sessions. With respect to data analysis, I was responsible for pre-processing data and for statistical analyses. With respect to manuscript preparation, I was responsible for writing the majority of the manuscript. I integrated revisions suggested by Drs. Debra Titone and Caroline Palmer and external reviewers, and prepared the manuscript for publication.

Caroline Palmer (second author) – Dr. Palmer was involved in the development of software that was used to pre-process acoustic data. She also significantly assisted with manuscript preparation and revision.

Debra Titone (third author) – Dr. Titone supervised and provided input throughout the research process, including research design and stimuli development; data analyses, interpretation of the results, and assistance in manuscript preparation and revision.

Chapter 3:

Irina Pivneva (first author) – With respect to responsibilities, as a first author I conceptualized and designed the study. I also conducted background research; selected stimuli; programmed the experiment; trained research assistants on experimental protocol. With respect to participant recruitment, I actively recruited participants, supervised research assistants conducting experimental sessions and transcribing acoustic
data. With respect to data pre-processing, I trained research assistants on transcribing procedures, supervised the transcription process, and was solely responsible for alignment of acoustic and eye movement data. I was responsible for performing all statistical analyses. With respect to manuscript preparation, I was responsible for writing the majority of the manuscript. I integrated revisions suggested by Dr. Debra Titone and external reviewers, and prepared the manuscript for publication.

Debra Titone (second author) – Dr. Titone supervised and provided input throughout the research process, including research design development; data analyses, interpretation of the results, and assistance in manuscript preparation and revision.

Chapter 4:

Irina Pivneva (first author) – With respect to responsibilities, as a first author I assisted Drs Debra Titone and Julie Mercier in study conceptualization and design. I also conducted background research and assisted in training research assistants on experimental protocol. With respect to data analysis, I was responsible for pre-processing data and for statistical analyses. With respect to manuscript preparation, I was responsible for writing the majority of the manuscript. I integrated revisions suggested by Drs. Debra Titone and Julie Mercier and external reviewers, and prepared the manuscript for publication.

Julie Mercier (second author) – Dr. Mercier programmed the task battery assessing executive control tasks, some of which were also used in studies reported in Chapter 3 and 4. Dr. Mercier assisted in training research assistants on experimental protocol and supervised research assistants conducting experimental sessions. She also significantly assisted with manuscript preparation and revision.
Debra Titone (third author) – Dr. Titone supervised and provided input throughout the research process, including research design and stimuli development; data analyses, interpretation of the results, and assistance in manuscript preparation and revision. In addition, Dr. Titone, together with Dr. Maya Libben, significantly contributed to stimuli development used in the bilinguals sentence reading task, originally published in Libben and Titone (2009).
CHAPTER 1: General introduction
Humans have the remarkable ability to acquire one or more languages (i.e., to become bilingual or multilingual), oftentimes obtaining native-like fluency in all known languages. Research that investigates the relation between bilingualism and human cognition is important in its own right, and also because of the constant increase in globalization and the need to communicate in trade, travel or immigration. In fact, a large proportion of the global population today is bilingual or multilingual (Grosjean, 2010). In Canada alone, nearly one in six Canadians can communicate in both official languages, English and French (Lepage & Corbeil, 2013). In the province of Quebec, nearly 43 percent of the population is bilingual and the rate of bilingualism increases faster in Quebec than in other Canadian provinces (Lepage & Corbeil, 2013). In Montreal, the largest metropolitan centre of Quebec, nearly 54 percent of the population is bilingual with 18.2 percent reporting to speak 3 or more languages (Statistics Canada, 2012).

Unsurprisingly, research investigating the link between language and cognition in bilinguals has gained much needed attention during the last four decades (e.g., Bialystok, 2009; Bialystok, Craik, Green, & Gollan, 2009; Cook, 1997; Hamers & Blanc, 1989; Kroll & Bialystok, 2013; Kroll & Gollan, 2014; Lambert, 1981; Macnamara, Krauthammer, & Bolgar, 1968; Peal & Lambert, 1962). Historically, bilingualism was believed to cause linguistic delays in development (reviewed in Hakuta & Diaz, 1985), which has been countered by recent work showing that bilingualism promotes neurocognitive change in brain and behavior, such as advantages in cognitive development in childhood or symptom-onset delays for neurodegenerative disease in aging (reviewed in Baum & Titone, in press; Kroll & Bialystok, 2013). Of special note, this line of research, which promotes the existence of bilingual advantages (rather than disadvantages), was pioneered by researchers from this department in the
1960’s: Wallace Lambert and John Macnamara. More recently, this line of work has been taken up by Ellen Bialystok, Fergus Craik, and Judy Kroll, among others (Bialystok, 2010; Bialystok, Craik, Green, et al., 2009; Kroll & Bialystok, 2013; Kroll, Bobb, & Hoshino, 2014; Macnamara et al., 1968; Peal & Lambert, 1962), who argue that the main source of bilingual advantages, where observed, arise from the fact that bilinguals engage in rampant cross-linguistic activation, even in single-language contexts (reviewed in Kroll & Gollan, 2014). Many studies show that bilinguals activate both languages when 1) reading for comprehension in their L2 and even in their L1 (reviewed in Van Assche, Duyck, & Hartsuiker, 2012); 2) hearing L2 words that share sounds with the L1 (reviewed in Blumenfeld & Marian, 2013); and 3) speaking, especially in the L2 (reviewed in Kroll et al., 2014; reiewed in Kroll & Gollan, 2014).

Indeed, graduate work that my colleagues and I conducted, in addition to the studies presented in this thesis, addressed such issues in the domains of reading and listening. In the domain of reading, my colleagues and I found that bilinguals activate word meanings in both languages (i.e., an index of cross-language activation) when reading for comprehension sentences with cognates (e.g., words that share spelling and meaning; piano across English and French) and interlingual homographs (e.g., words that share spelling but differ in meaning; chat French for cat and English for casual conversation) in their L2, as shown by my colleagues (Libben & Titone, 2009) and in their L1, as shown by my colleagues and me (Titone, Libben, Mercier, Whitford, & Pivneva, 2011). Within that work we had found that cross-language activation appears relatively early in processing (e.g., around 250 ms.) but it can be attenuated by sentence constraint during later processing stage (e.g., around 600-800 ms.) (Libben & Titone, 2009) and by individual differences in L2 proficiency among bilinguals (Titone et al., 2011).
Currently, we are extending these research questions to healthy bilingual older adults for whom cognitive processing is sub-optimal (Pivneva, Mercier, Sudarshan, Baum, & Titone, 2014).

In the domain of listening, my colleagues and I found that bilinguals also activate both languages when hearing English words that begin with similar sounds as other English words (e.g., target *field* versus *feet*) and also French words (e.g., target *field* versus *fille* French for girl), which compete for selection on activation (Mercier, Pivneva, & Titone, 2013, accepted), consistent with work by Marian and colleagues (Blumenfeld & Marian, 2011, 2013; Marian & Spivey, 1999, 2003; Shook & Marian, 2013; Spivey & Marian, 1999). Interestingly, in that work we have demonstrated that greater executive control among bilinguals helps to attenuate within-and cross-linguistic activation when the task is performed in the L2 and within-language activation when the task is performed in the L1. Similar to the reading domain, we are extending these research questions to healthy bilingual older adults who often demonstrate decline in auditory comprehension (Mercier, Sudarshan, Pivneva, Baum, & Titone, under review).

Thus, it is clear from prior work, including my own, that bilinguals automatically activate both languages (i.e., cross-language activation) even in single-language contexts, although bilinguals can flexibly modulate the degree of cross-language activation across different contexts. We expect the same to be true for language production, an essential part of this dissertation. In the domain of language production, such adaptive modulation of cross-language activation enables bilinguals to restrict their output to one language when speaking to monolinguals and/or to fluently mix their languages when speaking to other bilinguals. This linguistic flexibility arises from complex interactions between cortical and subcortical brain areas associated with language processing and, more interestingly, with general cognitive processing, including executive control networks (e.g., Abutalebi, 2008; Abutalebi et al., 2012;
Abutalebi & Green, 2007, 2008). Indeed, bilingual language production process is thought to be so demanding that some researchers have hypothesized that domain-general cognitive resources, such as those involved in executive control, are required for bilinguals to attain language control necessary to achieve linguistic flexibility across communicative contexts (Abutalebi & Green, 2007; Green, 1998; Green & Abutalebi, 2013; Kroll & Bialystok, 2013; Kroll & Gollan, 2014).

One way to link executive control to language processing is to compare groups of bilingual and monolingual speakers, consistent with studies alluded to previously (Kroll & Bialystok, 2013). Studies from this line of research reveal that bilinguals recruit executive control to a greater extent than monolinguals (Rodríguez-Pujadas et al., 2013) and that they do so differently (e.g., more efficient recruitment of the anterior cingulate cortex as reported by Abutalebi et al., 2012). Likely, the manner and the extent to which bilinguals recruit executive control might be unique to bilingualism because bilinguals do linguistically 1) “more” of (e.g., knowing more words for the same objects) and/or 2) different from (e.g., resolving linguistic cross-talk; switching between languages) what monolinguals already do (Abutalebi, 2008; Abutalebi & Green, 2007, 2008; Baum & Titone, in press; Chang, 2014; Li, Legault, & Litcofsky, 2014; Pivneva, Sheikh, Whitford, Mercier, & Titone, 2014). Therefore, different linguistic demands, in effect, naturally “train” executive control over time, often relating to cognitive advantages in bilinguals versus monolinguals (for reviews see Bialystok, Craik, & Luk, 2012; Cook, 1997; Kroll & Bialystok, 2013; Prior & Gollan, 2013).

While important, studies comparing bilinguals versus monolinguals oftentimes reveal inconsistent patterns of behavioral results (Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009; e.g., Hilchey & Klein, 2011; Kousaie & Phillips, 2012; reviewed in Li et al., 2014; Paap, 2014; Paap & Greenberg, 2013; Paap & Liu, 2014). More importantly, they fail to capture
the intrinsic differences among bilinguals (e.g., age and manner of second language (L2) acquisition; the current amount and manner of native language (L1) and/or L2 daily usage; individual differences on executive control), and the extent to which bilinguals engage in language behaviors that demand general versus specific kinds of executive control (Baum & Titone, in press; Pivneva, Sheikh, et al., 2014).

Thus, another way to link executive control to language processing is to investigate whether and how individual differences within a particular group relate to language processing. Based on this approach, we know from work in monolinguals that executive control is implicated in reading for comprehension, understanding and producing speech, among other linguistic operations (e.g., Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005; Faust, Balota, Duchek, Gernsbacher, & Smith, 1997; Friederici, Steinhauer, Mecklinger, & Meyer, 1998; Gernsbacher, 1997; Gernsbacher & Faust, 1991; Gernsbacher, Varner, & Faust, 1990; Hussey & Novick, 2012; Miyake, Just, & Carpenter, 1994; Novick, Kan, Trueswell, & Thompson-Schill, 2009; Novick, Trueswell, & Thompson-Schill, 2005; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997; Ye & Zhou, 2009). With respect to language production, a main focus of this thesis, executive control networks have been linked to the success with which monolinguals retrieve, select and name a verb associated with a target noun among multiple competing candidates (e.g., noun rope is associated with verbs hand, tie, enclose, etc.) (reviewed in Ye & Zhou, 2009). Additionally, greater executive control among monolingual aphasic patients relates to better communicative abilities and more successful recovery after the stroke (Brownsett et al., 2014).

While similar work on bilinguals is still in its infancy, we now know that individual differences in executive control among bilinguals relate to success with which they understand spoken speech (Blumenfeld & Marian, 2011, 2013; Mercier et al., 2013, accepted). Therefore, in
this thesis we focus on whether and how individual differences in executive control among bilinguals relate to bilingual language production, an area of research that is less well understood, despite its integral role in day-to-day communication. We then extend this focus to another skill integral to daily life, reading, to investigate whether and how individual differences in executive control among bilinguals relate to bilingual reading.

Thus, the three manuscripts presented in this thesis collectively investigate the link between individual differences in executive control among bilinguals and bilingual language production across different communicative contexts, with a subsequent extension to bilingual reading. The first manuscript (Pivneva, Palmer, & Titone, 2012; Published, Frontiers in Psychology, Bilingualism and Cognitive Control Special Issue) investigates the role of executive control in a natural communicative context, such as producing speech in a conversation. The second manuscript (Pivneva & Titone, under review; Cognition) investigates whether executive control demands vary in contexts with greater experimental control, such as when bilinguals produce single words and grammatically well-formed sentences. The third manuscript (Pivneva, Mercier, & Titone, 2014; Published, Journal of Experimental Psychology: Learning, Memory, & Cognition) investigates a secondary but important issue pertaining to the role of executive control during bilingual sentence reading. In the sections that follow, we first provide a brief definition for the concepts of executive control and bilingualism to specify how they were operationalized in this thesis. We then present bilingual language production processes and describe non-selective lexical activation during bilingual language production. Finally, we present bilingual language production models and supporting experimental evidence examining bilingual language production.
Definition of executive control used in this thesis

Given that the construct of executive control is a key component of this thesis, it is important to state what we mean by the term. Executive control is generally used as an umbrella term to describe domain-general cognitive processes that are often found paired with activations in the prefrontal cortex and cortico-striatal networks (for reviews see Alvarez & Emory, 2006; Chan, Shum, Toulopoulou, & Chen, 2008; Diamond, 2013; Jurado & Rosselli, 2007). From a behavioral perspective, executive control, is thought to be involved in decision-making, selectively attending to relevant information in the environment, inhibiting irrelevant information, shifting attention, monitoring and planning behavior (e.g., Alvarez & Emory, 2006; Aron, 2007; Baddeley & Hitch, 1974; Braver & Cohen, 2001; Braver, Paxton, Locke, & Barch, 2009; Chan et al., 2008; Clark et al., 2008; Diamond, 2013; Elliott, 2003; Engle, Kane, & Tuholski, 1999; Hasher, Lustig, & Zacks, 2007; Meiran, 2000, 2010; Miller & Cohen, 2001; Miyake & Friedman, 2012; Miyake et al., 2000; Monsell, 2003; Stuss & Alexander, 2000).

While executive control is critical for day-to-day human functioning and decision-making, it is a matter of active theoretical debate whether executive control consists of a unified set of distinct processes tapping into a core common component, reflecting inhibitory processes among others (De Frias, Dixon, & Strauss, 2006; Dempster, 1992; Duncan, Emslie, Williams, Johnson, & Freer, 1996; Kimberg, D'Esposito, & Farah, 1997; Munakata et al., 2011; Parkin & Java, 1999), or whether it is composed of several independent processes tapping into distinct components (e.g., Braver & Cohen, 2001; Braver et al., 2009; Godefroy, Cabaret, Petit-Chenal, Pruvo, & Rousseaux, 1999; Miyake et al., 2000; Stuss & Alexander, 2000).

Unified or distinct executive control processes views are supported by models derived from work investigating executive control processes in healthy individuals and individuals who
have suffered neurological damage (for reviews see Alvarez & Emory, 2006; Chan et al., 2008; Jurado & Rosselli, 2007). For example, Miyake and colleagues (2000) proposed a model of distinct processes of executive control based on work by Teuber (1972). Their unity and diversity of executive control framework includes three distinct executive control processes, namely updating (information monitoring), shifting (switching between tasks), and inhibition (suppressing prepotent non-target responses). To extract the distinct executive control processes, the authors conducted confirmatory factor analysis on several typical but different executive control tasks. Factor analysis and other similar statistical techniques (e.g., structural equation modeling) allow researchers to derive unobserved latent variables, which contain shared variance, from observed variables across several tasks. Based on their results, Miyake and colleagues posited that the unity in their framework reflected a large proportion of variance shared between the three processes. The diversity in the framework reflected a relatively large proportion of variance specific and unique to each process.

After a decade of research across different populations, including preadolescent children (Rose, Feldman, & Jankowski, 2011), older adults (Vaughan & Giovanello, 2010), and twins (Friedman, Miyake, Robinson, & Hewitt, 2011), most recent accounts now tend to favor a unified view of executive control, which argue that executive control is a common “entity” that imposes control in information processing (Miyake & Friedman, 2012; Munakata et al., 2011). For example, Miyake and Friedman (2012) modified their original model to propose a new unity/diversity framework. The new framework consists of a one common component (e.g., inhibition) with two sub-components (e.g., updating- and shifting-specific abilities). The common component captures nearly 100 percent of variance associated with inhibition and explains a large proportion of variance of both sub-components (Friedman et al., 2008; 2011).
The common inhibitory component represents the ability to actively maintain goal-directed information and behavior to bias lower-level processing, including processes involved in response inhibition (Munakata et al., 2011).

While the debate about executive control composition has shifted towards the unified view, often accounts propose that executive control imposes at least two types of control (reviewed in Jurado & Rosselli, 2007; but see Stuss, 1992; for the third process of “awareness of self and the environment”). Some views propose that control only varies in its degree. A now classical framework by Norman and Shallice (1986) represents executive control as a Supervisory Attentional System (SAS) to control information processing with varying degrees of involvement. The SAS is thought to be involved in planning future actions, decision-making, working with novel information or manipulating information in a novel manner. Within this framework, information processing is thought to occur in an automatic (routine) or controlled (non-routine) manner. Routine type of information processing (e.g., reading, walking) requires minimal to no involvement of the SAS. Conversely, non-routine type of information processing (e.g., reading word red printed in blue ink, walking on ice) requires heavy involvement of the SAS, as the information processing must occur in a novel or less-practiced manner.

A more recent account, the Dual Mechanisms framework of executive control, also argues that executive control operates through two independent processes (Braver, 2012). They are early proactive control and late reactive control. Proactive control engages when anticipating increases in cognitive demands. To do so, proactive control maintains goal-relevant information all while down-regulating non-target information early on. Reactive control engages as a correction mechanism when cognitive demands increase during later processing stages. To do so, reactive control inhibits specific interfering non-target information as it occurs in real time.
In contrast to dual process accounts, a recent Unified framework of executive control, proposes that executive control represents a single “construct” inhibition (Munakata et al., 2011). However, control is still expressed via two types of inhibitory processes, which are not mutually exclusive and often work in unison. Two processes consist of directed global inhibition and indirect competitive inhibition. Directed global inhibition involves subcortical and prefrontal regions and reflects global neural control processes, which provide contextual information to inhibit non-target processing in a top-down manner. Indirect competitive inhibition occurs within prefrontal regions and involves inhibition of non-target information indirectly (i.e., lateral inhibition) by increasing activation of the target information.

Given the ongoing debate about the composition of the executive control, we took an empirical, task-based approach to executive control by using a variety of tasks that been successfully used in prior studies linking executive control to bilingual language processing (Blumenfeld & Marian, 2011, 2013; Linck, Hoshino, & Kroll, 2008; Linck, Schwieter, & Sunderman, 2012; Mercier et al., 2013, accepted; Pivneva, Mercier, & Titone, 2014). More importantly, we also were especially cognizant to select executive control tasks that were non-verbal to minimize any influence of L2 ability on domain-general cognitive processing (e.g., Luo, Luk, & Bialystok, 2010), which is not always true in the prior literature. The rationale of this decision is that we wanted to obtain measures of executive control that were as independent as possible of language function specifically.

Definition of bilingualism used in this thesis

Given that the construct of bilingualism is also a key component of this thesis, it is important to state what we mean by the term. What does it mean to be “bilingual?” While
seemingly intuitive, defining whether one is bilingual, and characterizing how bilinguals may differ has long been a challenge to both lay people and researchers. Broadly speaking, to be bilingual means to know and use two languages. However, “to know and use two languages” can vary drastically from one individual to the next. For example, a stricter criterion for bilingualism may include bilinguals who know and use their two languages with “native-like” fluency (Bloomfield, 1933, p.56). However, this definition presents a rather limited view, such that it excludes a large proportion of individuals who know and use two languages daily but without achieving native-like fluency. Moreover, it is unclear what “native-like” fluency means, even among monolingual speakers of the language, as language abilities vary greatly from one individual to the next. To that effect, some popular “myths” about bilingualism include beliefs that bilinguals possess “carbon-copies” of their two languages, such as the same vocabulary knowledge across two languages, no accent in either one, advanced understanding of figurative language (e.g., idioms, metaphors, and proverbs), among others (Grosjean, 2010). More importantly, as researchers we need to remember that a bilingual is not two monolinguals combined in one (e.g., Grosjean, 1989, 2010; Malt & Sloman, 2003).

Thus, most often, researchers reject strict all-or-none definitions of bilingualism and adopt more graded definitions. This more inclusive and empirically derived view of bilingualism recognizes that bilinguals are a heterogeneous group that vary in their degree of L1 and L2 knowledge and usage (see Butler & Hakuta, 2006 for bilingualism definition and classification issues). For example, bilinguals vary in the age they learned their L2. Some bilinguals learn their L2 together with their L1 from birth, known as simultaneous bilinguals. Others learn their L2 after having acquired their L1 either relatively early or late in life, known as early and late sequential bilinguals respectively (e.g., Genesee et al., 1978; Grosjean, 1982). Bilinguals can
vary as a function of age of L2 acquisition with respect to structure and function of brain areas recruited in language processing (Abutalebi & Green, 2007). For example, there are structural changes in the brain associated with whether a bilingual is an early or a late sequential or a simultaneous bilingual or a monolingual (Klein, Mok, Chen, & Watkins, 2013). Bilinguals also vary in their relative L1 and L2 proficiency, independent of age of acquisition. For some bilinguals the relative proficiency across their two languages is equal (i.e., balanced bilinguals), while for others the relative proficiency is stronger (i.e., unbalanced bilinguals; including those for whom L2 proficiency is greater than L1, known as reverse-dominance bilinguals) for one language versus another (Peal & Lambert, 1962). Bilinguals can also vary in their current degree and manner of daily use of two languages. For example, some bilinguals compartmentalize their languages in daily use, such that they use one language at home and another at work, while others use both languages at home and work (Green, 2011; Grosjean, 2010; Pivneva, Sheikh, et al., 2014). Given the multitude of ways in which bilinguals can differ, this thesis quantifies individual differences among bilinguals and explicitly assesses their role in bilingual language production or statistically accounts for variability brought on by such differences.

In this thesis, bilinguals will be defined as individuals who demonstrate communicative ability in English and French, obtained through varied means (e.g., from family, friends, formal instruction, etc.), with varied levels of relative proficiency in their two languages. To partake in the experiments reported in this thesis, bilinguals will have had to identify themselves as bilingual with either English or French as the native language and French or English as the L2. Prior to participation, participants were also forewarned that they needed a working knowledge of English and French to complete tasks in English and in French as task materials assessing bilingual language production were presented in both languages. The vast majority of the
participants had such working knowledge and reported their native language as being currently their most-dominant. Simultaneous bilinguals who acquired both languages from birth were also included in the analyses. We classified them as native-English or native French based on the native language of their mother to account for prenatal linguistic input.

**Bilingual language production processes**

The first two manuscripts of this thesis involve bilingual language production, thus, the first part of this section will describe processes involved in bilingual language production and the second part of this section will describe the consequences of automatic cross-language activation observed in bilingual language production. When people intend to speak, language production proceeds through a series of stages in a top-down manner with the speaker maintaining control of production processes. Language production stages include conceptualization of the intended message, formulation (e.g., retrieving and combining words in the message) and articulation (e.g., voicing the message) (Kempen & Hoenkamp, 1987; Levelt, 1989). For example, when a native French-English bilingual, a student named Marie, sees a picture of a couch, she first activates conceptual cues for the picture during conceptualization (Figure 1; De Bot, 1992; Hermans, 2000; Poulisse & Bongaerts, 1994). Then, Marie selects a language appropriate to a given context and activates lexical information that best represents the picture at the lemma level during formulation (e.g., word label, gender, word class, among others), a process known as lexical access. In parallel, she also activates lexical information for sofa and other concepts that are semantically related to the picture within a language (e.g., loveseat, armchair, etc.) and cross-linguistically (e.g., divan; French for couch) even if she plans to say couch in English, her L2 (Kroll, Bobb, Misra, & Guo, 2008). Subsequently, she must select the most appropriate lexical
information for the picture to realize the output all while regulating all irrelevant lexical information that activates within- and across-languages (Green, 1998; Kroll et al., 2008), a process known as lexical selection. Lastly, Marie activates phonological information to formulate and articulate the selected lexical information at a rate of about 150-300 words per minute (Goldman-Eisler, 1968).

\[\text{Figure 1. Graphical representation of stages of language production based on Poulisse and Bongaerts, 1994 and Hermans, 2000 (adapted from Kroll, Bobb and Wodniecka (2006))}\]

Of note, these language production processes occur incrementally, such that Marie transfers partially prepared fragments of the message from one stage to the next before completely preparing the message in its entirety. Consequently, she begins articulating earlier
parts of the message before fully retrieving all sound patterns for the message. Across these stages, Marie likely recruits non-linguistic executive control to help her with speech planning processes and to attenuate linguistic cross-talk for word forms and sound patterns to maintain fluency during speech production and to monitor her output (Green, 1998).

Given the top-down nature of production processes, evident through the intention to convey the message, it is often assumed that bilinguals selectively activate and regulate early on which language they will use during speech planning. In other words, bilinguals can selectively activate and lexically access only the target language that is relevant in a given context, while disregarding the non-target language during earliest stages of speech planning. However, while this view carries a certain intuitive appeal, research has consistently shown that bilinguals activate both languages automatically and non-selectively, especially during L2 production (e.g., Colomé, 2001; Colomé & Miozzo, 2010; Costa, 2005; Costa, Caramazza, & Sebastian-Galles, 2000; Costa, Colomé, Gomez, & Sebastian-Galles, 2003; De Bot, 1992; De Groot, 2011; Gollan & Kroll, 2001; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Hoshino & Thierry, 2011; Kroll et al., 2008; Kroll, Bobb, & Wodniecka, 2006; Kroll, Dussias, Bogulski, & Valdes Kroff, 2012; Kroll & Gollan, 2014; Poulisse & Bongaerts, 1994). In other words, there is linguistic cross-talk that occurs early on, which oftentimes can create competition throughout the planning stages. The extent to which linguistic cross-talk creates competition varies as a function of the relative activation of the target and non-target lexical information within- and cross-linguistically (e.g., Finkbeiner, Almeida, Janssen, & Caramazza, 2006; Finkbeiner, Gollan, & Caramazza, 2006; Starreveld & La Heij, 1996). The degree of competition from linguistic cross-talk can be measured by the Luce ratio, which reflects the inverse relationship between the relative activation levels of competing lexical information and the time required to select the target
lexical information (Finkbeiner, Gollan, et al., 2006). Simply put, the closer two lexical items are in meaning to each other, the more competition there is between them. In the case of bilingual language production, translation equivalents share a common semantic representation and likely can be classified as more closely related than semantically related words within a language, thus, creating a “hard problem” in bilingual language production (Finkbeiner, Gollan, et al., 2006).

Linguistic cross-talk also promotes dynamic interactions between languages with L2 influencing L1 and vice-versa (Dussias, 2003; Gollan, Montoya, Cera, & Sandoval, 2008; Jared & Kroll, 2001; Kroll, Dussias, et al., 2012; Kroll & Gollan, 2014; Whitford & Titone, 2012). Likely as a result of the dynamic cross-linguistic interactions, bilinguals on average are slower and less accurate in accessing information in any one of their two languages when compared to monolinguals (e.g., Gollan & Ferreira, 2009; Hanulová, Davidson, & Indefrey, 2010; Ivanova & Costa, 2008; reviewed in Kroll & Gollan, 2014; Linck et al., 2008; reviewed in Runnqvist, Strijkers, & Costa, 2014; Sandoval, Gollan, Ferreira, & Salmon, 2010) and these differences increase with increases in L2 usage (Gollan et al., 2008; Whitford & Titone, 2012). However, bilinguals also rarely make language errors (e.g., speaking the wrong language to a monolingual), especially highly L2 proficient bilinguals for whom competition between two strongly activated languages would be the greatest. To that extent, it is thought that bilinguals develop superior language control abilities to select and monitor their language output appropriately in a given context, overcoming the “hard problem” and/or reducing the competition effects (reviewed in Kroll & Gollan, 2014).
Models of bilingual language production

Four decades of research linking language and cognition in bilinguals have led to many theoretical models of bilingual language processing. Some models were developed as extensions of monolingual language processing models. For example, the Bilingual Interactive Activation model (Dijkstra & Van Heuven, 1998; Van Heuven, Dijkstra, & Grainger, 1998) and its extension, the Bilingual Interactive Activation plus model (BIA+; Dijkstra & van Heuven, 2002) were developed as extensions of the Interactive Activation (IA) model of monolingual word recognition proposed by McClelland and Rumelhart (1981). The bilingual versions of the model capture how non-selective lexical activation occurs when bilinguals read for comprehension. In the domain of spoken word recognition, the Bilingual Model of Lexical Access (BIMOLA; Grosjean, 1988, 1997) was developed as an extension of the TRACE model of spoken word recognition proposed by McClelland and Elman (1986). In the domain of language acquisition, the Self-Organizing Model of Bilingual Processing (SOMBIP; Li & Farkas, 2002; for a review see Li & Zhao, 2013) was developed as an extension of several models of normal and pathological language development, such as the Dyslexic and Category-Specific Aphasic Impairments in a Self-Organizing Feature Map Model of the Lexicon (DISLEX; Miikkulainen, 1997), the Developmental Lexicon (DevLex; Li, Farkas, & MacWhinney, 2004) and its extension (DevLex-II; Li, Zhao, & Mac Whinney, 2007), and Hyperspace Analogue to Language (HAL; Burgess & Lund, 1997).

In contrast, other models were developed to account for specific types of bilingual language processing, rather than extensions of monolingual language processing models. For example, the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994) was developed to model asymmetries observed in translation during L2 learning. This model is particularly strong
in explaining the earliest stages of L2 learning and processes involved when transitioning to higher L2 proficiency levels. More recently, the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS; Shook & Marian, 2013) was developed to model bilingual spoken language comprehension.

While an exhaustive review of all models of bilingual language processing is beyond the scope of this thesis, we will discuss models that have been specifically developed for bilingual language production. Of note, all models of bilingual language production generally agree that both languages activate even briefly during speech planning. However, they differ with respect to how bilinguals select the target language amidst cross-linguistic activation and also with respect to the hypothesized role of executive control in bilingual language production. One class of bilingual language production models de-emphasizes executive control in arguing that bilinguals boost activation of the target language to surpass that of the non-target language (e.g., Costa, 2005; Costa & Caramazza, 1999; La Heij, 2005). For the purposes of this thesis, we will call this class of models the “mental firewall” accounts, a term originally proposed by Kroll and colleagues (2006). Another class of bilingual language production models emphasizes executive control in arguing that bilinguals actively inhibit the non-target language to select the target language (Abutalebi & Green, 2007, 2008; Green, 1998; Green & Abutalebi, 2013; Guo, Liu, Misra, & Kroll, 2011; Kroll et al., 2008; Kroll & Gollan, 2014; Linck, Kroll, & Sunderman, 2009; Philipp, Gade, & Koch, 2007; Verhoef, Roelofs, & Chwilla, 2009). Within this class of models we will devote particular attention to the Inhibitory Control model of bilingual language production, because this model serves as the main theoretical framework of this thesis (Green, 1998).
Mental firewall accounts of bilingual language production

The mental firewall accounts of bilingual language production generally posit that even though bilinguals non-selectively activate their two languages during production, this linguistic cross-talk often does not result in competition between languages, because bilinguals can use environmental cues to minimize cross-linguistic competition before it starts. For example, according to the Concept Selection (CS) model, bilinguals select the target language and resolve cross-language activation during the conceptualization (i.e., preverbal) stage of speech planning before accessing word meanings. To do so, bilinguals capitalize on language cues during the conceptual stage of production to boost activation of the target language such that only the target lexical information becomes available during subsequent stages (La Heij, 2005). Of note, the language-cue parallels “register” (e.g., speaking to a friend versus a professor) in Levelt’s (1989) serial model of monolingual language production.

In a model proposed by Costa and colleagues, bilinguals select the target language and resolve cross-language activation during the formulation stage, when accessing word meanings (Caramazza, 1997; Costa, 2005; Costa & Caramazza, 1999; Costa et al., 2000; Costa, Miozzo, & Caramazza, 1999; Finkbeiner, Almeida, et al., 2006; Finkbeiner & Caramazza, 2006). Of note, the model proposed by Costa and colleagues represents different levels of language production in nodes instead of stages (Figure 2). Accordingly, activation of semantic nodes reflects activation of meaning; lexical nodes reflect complete phonological representation for a given word (i.e., how it sounds); and sublexical nodes represent phonological components (known as “phonological lexemes” in Caramazza’s model (1997)), such as a single phoneme (i.e., a specific sound within a word). Within this model, semantic, lexical and sublexical nodes activate automatically and non-selectively when bilinguals prepare to speak, such that semantic nodes
activate and propagate activated information to the lexical nodes, which in turn further propagate it to the sublexical nodes. As the activation occurs, a language-specific mechanism, inherent to the language processing system, can selectively 1) ignore the activation of the non-target lexical information and 2) boost activation for the target lexical information, such that the activated non-target lexical information does not pass an activation threshold, which Kroll and colleagues (2006) have referred to as a mental firewall. Because non-target lexical information does not pass through the mental firewall, little to no cross-linguistic competition occurs. Should competition occur (albeit minimally), it can be resolved via lateral inhibition (i.e., locally inhibiting specific lexical candidates) within the bilingual language system without relying on external domain-general cognitive resources (Howard, Nickels, Coltheart, & Cole-Virtue, 2006). Critically, mental firewall accounts de-emphasize the role of executive control in bilinguals because language processing system is architecturally blocked from the domain-general cognitive processing system. Of note, the language-specific selection mechanism within the mental firewall accounts parallels a grammatical class selection mechanism (e.g., planning to produce a noun, so activation is boosted only for nouns vs. other grammatical classes), in Dell’s (1986) interactive model of monolingual language production.
Figure 2. Graphic representation of a bilingual language production model proposed by Costa and colleagues (Costa, 2005).

To illustrate, when our native French-English bilingual student, Marie, plans to produce *couch* in English, the strongest activated competitor theoretically should be the translation-equivalent *divan* (French for *couch*) from her non-target language, French, and *sofa* from the
target language, English. The goal for Marie is to overcome the activation of the irrelevant *divan*, so that she may correctly select and produce *couch*. According to the mental firewall accounts, Marie uses contextual cues from her environment (e.g., she is at an English university speaking to her monolingual English friend) and engages her language-specific selection mechanism within the language-processing system to boost activation to lexical information in English, while the lexical information from French remains less activated and for the most part ignored, as it cannot pass through the mental firewall. As such, when Marie begins speaking *couch*, *divan* presents little to no competition since it is not included in the pool of possible candidates for production and, thus, does not interfere during production (reviewed in Runnqvist et al., 2014).

Some researchers have argued that the language-specific selection mechanism may not function as efficiently as hypothesized across different levels of processing, as several studies have overwhelmingly demonstrated the influence of the non-target language across the semantic (Jescheniak & Schriefers, 1998; Peterson & Savoy, 1998; Rahman & Aristei, 2010), lexical (Finkbeiner, Almeida, et al., 2006; Schwieter & Sunderman, 2009), and phonological (Hermans et al., 1998) levels of bilingual speech production. To address this problem some mental firewall accounts have proposed that cross-language activation and competition, where observed, occur only during later stages of speech planning (e.g., when bilinguals are about to articulate the word). According to one such account, the Response Exclusion account, bilinguals prepare articulatory plan for the target and the non-target lexical information to relay to the output buffer for articulation (Finkbeiner & Caramazza, 2006; Janssen, Schirm, Mahon, & Caramazza, 2008; Mahon, Costa, Peterson, Vargas, & Caramazza, 2007). The output buffer can only hold one response at a time. If the response in the non-target language activates stronger than in the target
language, the non-target lexical information will occupy the output buffer until the selection mechanism engages to clear it, so that the target lexical information can enter the output buffer (Costa, Alario, & Caramazza, 2005; Mahon et al., 2007; Miozzo & Caramazza, 2003).

Other researchers have argued that the function of the language-specific selection mechanism may be task-dependent (Bloem & La Heij, 2003; Schriefers, Meyer, & Levelt, 1990) or may not even come online until bilinguals attain certain levels of L2 proficiency (Costa & Santesteban, 2004; Schwieter & Sunderman, 2008). Accordingly, it is thought that bilinguals with low L2 proficiency likely rely on domain-general cognitive resources (e.g., inhibitory processes) to obtain language control, because they lack experience managing their two languages and because L2 processing is more cognitively demanding for them. With increasing L2 usage and, hence, higher L2 proficiency, bilinguals begin to develop a language-specific selection mechanism, which they then can use to maintain a better form of language control (reviewed in Runnqvist et al., 2014; Schwieter & Sunderman, 2008).

To summarize, the mental firewall accounts adopt a non-competitive approach to bilingual language production. Bilinguals resolve cross-language activation by selectively boosting activation of the target language, such that the non-target language, even if activated, does not pass the mental firewall. Importantly, mental firewall accounts de-emphasize the role of domain-general executive control during bilingual language production; however, some recent proposals have started to discuss the role of executive control in bilingual language production.
Inhibitory accounts of bilingual language production

The inhibitory accounts of bilingual language production argue that when bilinguals non-selectively activate their two languages during production, the linguistic cross-talk almost always results in competition across different stages of production (reviewed in Kroll et al., 2014; Kroll & Gollan, 2014). Typically, competition for selection is strongest when bilinguals plan to speak in the usually less-dominant L2, such that the usually more-dominant L1 activates automatically and creates cross-linguistic interference that must be resolved (e.g., Abutalebi & Green, 2007; Green, 1998). Bilinguals overcome cross-language activation and resolve cross-linguistic competition by inhibiting the more-dominant L1, so that the successful L2 realization occurs. Interestingly, strength of inhibition is reactive, such that the more dominant-language will require greater inhibition and the less-dominant language will require less inhibition. However, greater inhibition can sometimes result in significant costs (e.g., slower naming and/or worse accuracy) to the more-dominant language when it needs to be reactivated for subsequent production, a curious asymmetry in language switching (e.g., Meuter & Allport, 1999). An interesting point to consider is whether repeated L1 inhibition can carry significant consequences to L1 production in general, not only transiently but perhaps more permanently (Guo et al., 2011; Kroll et al., 2014; Kroll et al., 2008; Kroll & Gollan, 2014; Linck et al., 2009; Philipp et al., 2007).

One theoretical framework that captures relatively well the effects described above is Green’s (1998) Inhibitory Control (IC) model. The IC model is also central to this thesis, because it assigns a fundamental role to executive control during bilingual language production (Figure 3). According to the IC model, language production is a goal-directed, communicative action. Like other non-linguistic physical actions, language production consists of mental task
schemas. Task schemas represent action sequences that are implemented by the conceptualizer (C) to achieve particular goals (G). Task schemas may be routine (L1 production) or non-routine (L2 production). During production multiple task schemas activate in parallel and compete to control output (O). To resolve competition, executive control, namely the supervisory attentional system (SAS; Shallice & Burgess, 1996), inhibits routine goals, and monitors the successful implementation of non-routine goals, based on the input (I) from the bilingual lexico-semantic system. As well, within the bilingual lexico-semantic system, executive control fine-tunes the relative activation and inhibition of words within each language to select and output appropriate words for production.

**Figure 3.** Graphic representation of the Inhibitory Control model, the central theoretical framework of this thesis (Green, 1998).
To illustrate, when our native French-English bilingual student, Marie, engages in a dialogue with John, her monolingual English friend, she will recruit executive control to suppress globally the irrelevant, more-dominant French language task schema. Likely, Marie will have to recruit executive control to a greater extent when she speaks English relative to when she speaks French. In addition, Marie will further monitor her output and fine-tune the relative activation and inhibition of semantically related within- (e.g., *loveseat*, *armchair*) and between-language (e.g., *divan*) lexical items.

Abutalebi and Green (2007) extended the IC model to incorporate neurocognitive evidence for bilingual language production. They identified a network of cortical regions (e.g., prefrontal, inferior parietal and anterior cingulate cortices) and subcortical structures (e.g., basal ganglia, the head of the caudate nucleus in particular) that modulate competition between the L1 and L2 activation during bilingual language production. Within this framework subcortical structures (i.e., basal ganglia) modulate the global activation of L1 or L2 task schemas, whereas frontal cortical structures modulate local activation of L1 and L2 lexical items. Within the extension of the IC model the L2 proficiency can also significantly modulate bilingual language production. L2 language production is more controlled and less automatic (see also Favreau & Segalowitz, 1983; Segalowitz, 2010; Segalowitz & Hulstijn, 2005), especially when L2 proficiency is low, thus requiring executive control (prefrontal function, in particular; Petrides, 1998). In contrast, L2 production becomes more automatic and less dependent on executive control when L2 proficiency is high. However, L1 production effort might instead increase with high L2 proficiency due to weaker links between L1 word forms and concepts (Bialystok, Luk, Peets, & Yang, 2010; Gollan et al., 2008; Gollan et al., 2011; Ivanova & Costa, 2008; Michael & Gollan, 2005; Whitford & Titone, 2012).
To summarize, the inhibitory accounts adopt a competitive approach to bilingual language production. Bilinguals resolve the “hard problem” by actively inhibiting the non-target language, so that successful realization of the target language can occur. Importantly, inhibitory accounts make specific predictions for the role of domain-general executive control during bilingual language production: (1) L2 production should require greater executive control 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 versus less demanding communicative task should limit the resources available for executive control to engage), and (3) bilinguals should differ in the extent to which they successfully produce language insofar that they differ in their executive control capacity.

**Bilingual language production studies**

This section will review typical empirical evidence supporting different classes of bilingual production models described above. Over decades researchers have used many methods to study bilingual language production. While a comprehensive review of all possible methods is beyond the scope of this thesis, this section will discuss two most commonly used paradigms: a bilingual version of the picture-word interference paradigm and a language-switching paradigm. As will be seen, the bottom-line of this review is that studies using picture-word interference paradigm often have provide evidence for the mental firewall accounts, while studies using language-switching paradigm often provided evidence for the inhibitory accounts. More importantly, researchers now tend to favor proposals to develop new methods to study
bilingual language production, as both paradigms, picture-word interference and language-switching, sometimes produce mixed results (Runnqvist et al., 2014).

**Picture-word interference paradigm**

In the picture-word interference paradigm participants see pictures one at a time with printed words superimposed on them (e.g., Glaser, 1992; Glaser & Düngelhoff, 1984; Glaser & Glaser, 1989; Hentschel, 1973; Lupker, 1979; Rosinski, 1977; Schriefers et al., 1990). Participants are instructed to name the picture all while ignoring the printed word. Printed words vary in how closely they relate semantically to the picture. In the unrelated condition the printed word (e.g., *rocks*) does not relate semantically to the picture (e.g., *couch*), whereas in the semantically related condition the printed word (e.g., *loveseat*) relates semantically to the picture (e.g., *couch*). It is expected for semantically related versus unrelated words to interfere with picture-naming and delay naming times, an effect known as picture-word semantic interference often observed in studies of monolingual language production (e.g., Alario, Segui, & Ferrand, 2000; Bloem & La Heij, 2003; Bloem, Van Den Boogaard, & La Heij, 2004; Damian & Bowers, 2003; Damian, Vigliocco, & Levelt, 2001; Glaser & Düngelhoff, 1984; La Heij, Dirks, & Kramer, 1990; La Heij, Starreveld, & Steehouwer, 1993; Rahman & Melinger, 2009; Roelofs, 1992; Rosinski, Golinkoff, & Kukish, 1975; Schriefers et al., 1990; Starreveld & La Heij, 1995, 1996; Wheeldon & Monsell, 1994). Semantic interference is thought to arise because lexical candidates activate and compete for selection. To resolve competition, speakers must inhibit the activated non-target candidates, a process which slows naming.

In a bilingual version of the task, participants name pictures in one language with the superimposed words printed in the other language. This version of the picture-word interference paradigm allows researchers to address whether bilinguals experience between-language
competition when candidates from both languages activate. Bilingual picture-word interference effect is expected to be greatest when the superimposed word is the translation equivalent for the presented picture, because of shared conceptual representation. For example, when Marie names a picture of a *couch* in English, the French translation equivalent *divan* would receive activation and compete for selection, significantly slowing the naming speed for the picture of the *couch*.

Several studies have found that translation equivalent printed words interfere with naming, revealing semantic interference effects similar to monolingual studies (e.g., Finkbeiner, Almeida, et al., 2006; Hermans et al., 1998; Jescheniak & Schriefers, 1998; Peterson & Savoy, 1998; Rahman & Aristei, 2010; Schwieter & Sunderman, 2009). Some studies have even shown that semantic interference effects can occur relatively early in the time course of speech planning (e.g., Hoshino & Thierry, 2011). These findings have been used to support the inhibitory accounts of bilingual language production, which argue that the parallel activation of two languages results in competition. To resolve competition, bilinguals must inhibit the translation equivalent from the non-target language; a process that slows naming.

Interestingly, other studies show the exact opposite pattern of effects. Specifically, bilinguals on average name pictures faster in the translation equivalent versus unrelated condition, revealing semantic facilitation effects (e.g., Costa & Caramazza, 1999; Costa et al., 1999; Runnqvist et al., 2014). More interestingly, some studies have observed facilitatory effects even when translation equivalents were not presented explicitly (Gollan & Acenas, 2004; Gollan, Montoya, Fennema-Notestine, & Morris, 2005). These findings have been used to provide support for the mental firewall accounts of bilingual language production, which argue that the initial parallel activation still occurs but it does not result in competition. To that effect, the language-specific mechanism boosts the activation of the target language, such that it passes
through the mental firewall leaving the non-target lexical information on the “wrong” side of the mental firewall. Because the non-target lexical information is no longer competitive, it compliments speech planning and, in fact, facilitates processing through priming of the commonly activated semantic nodes.

However, one potential drawback with this interpretation is that the activation itself might be already language-specific, instead of the language-specific selection mechanism. While it may, the mental firewall accounts disagree. Importantly, if the activation itself is language-selective (i.e., activate one language only), then naming the translation equivalent should not differ from naming the semantically unrelated word because both would be “unrelated” to the target word. However, given that there is a significant difference in naming, supports the non-selective activation. Given that the direction of the effect is facilitatory, supports the idea of a language-specific selection mechanism (De Groot, 2011).

**Language-switching paradigm**

Language-switching paradigms have been a staple of bilingual language production research (reviewed in Kroll & Gollan, 2014; Runnqvist et al., 2014). Typically, researchers have used language-switching paradigms to test whether cross-linguistic activation results in competition and, thus, requires inhibition. In one type of a language-switching paradigm, bilinguals name stimuli (e.g., pictures, numbers, etc.) in one language and switch or not switch to a different language on the next trial based on a cue (e.g., background color, country’s flag).

For example, Meuter and Allport (1999) conducted, what is now, a canonical study, in which bilinguals named numerals (e.g., 2, 9) in the L1 or L2 cued by the background color.
There were two types of trials, switch and no switch. Bilinguals were on average slower to name numerals in the switch trials than in the no switch trials. Interestingly, Meuter and Allport observed a curious asymmetry when bilinguals switched into their more-dominant L1 versus their less-dominant L2, such that switching into the L1 revealed greater costs than switching into the L2. In fact, L1 naming speed following a switch was even slower than the L2 naming speed. The results of this study and those that were subsequently inspired by Meuter and Allport’s work have frequently been used to support the inhibitory control accounts of bilingual language production. Specifically, bilinguals are thought to inhibit L1 during L2 naming and when they are cued to switch into the L1, it takes longer to re-activate previously inhibited L1, resulting in slower naming.

The interpretation that asymmetrical switch costs result from inhibition of the L1 has been systematically challenged in the literature. There are numerous studies that replicate switch costs (e.g., slower and/or less accurate on switch versus no-switch trials) but fail to replicate the asymmetry in switching (reviewed in Runnqvist et al., 2014). Some studies have found asymmetry in switching but only in specific settings. For example, Costa and Santesteban (2004) found that low L2 proficiency bilinguals reveal asymmetry in language switching but high L2 proficiency bilinguals switch symmetrically between their L1 and L2 (even switching between L1 and L3). As well, there are studies that found symmetrical switch costs even for low L2 proficiency bilinguals (Christoffels, Firk, & Schiller, 2007; Gollan & Ferreira, 2009; Verhoef, Roelofs, & Chwilla, 2010). Furthermore, there are studies that demonstrate that asymmetrical switch costs can vary as a function of task demands, such as voluntary switching; response predictability, response preparation time, and response valence among others (Finkbeiner, Almeida, et al., 2006; Gollan & Ferreira, 2009; Runnqvist et al., 2014; Verhoef et al., 2009).
Taken together, these findings are generally interpreted to support the mental firewall accounts (reviewed in Runnqvist et al., 2014).

Another way researchers use language-switching paradigm to study the effects of cross-linguistic competition is by comparing whether and how naming differs when bilinguals name single pictures in pure- (L1-only, L2-only) and mixed-language (L1/L2) blocks (e.g., Christoffels et al., 2007; Guo et al., 2011; Kroll & Gollan, 2014). In the pure-language blocks, bilinguals name all pictures only in the L1 or the L2. In the mixed-language blocks, bilinguals name pictures in the L1 or the L2 on any given trial based on a cue. Of significant importance to this thesis, bilinguals are often pre-trained on the stimuli or name the same stimuli across languages and/or across pure- and mixed-language blocks. It is thought that naming the same image twice should result in faster and more accurate naming times during the second time, because of repetition priming (reviewed in Kroll & Gollan, 2014).

By comparing naming across languages or across blocks, researchers can estimate different types of inhibition that may be required for language control (e.g., De Groot, 2011). For example, when researchers compare naming across pure-language blocks (L1-only vs. L2-only), they can estimate global language control (i.e., inhibition of the complete language system) by manipulating the language order of the blocks. Researchers can also compare naming across languages within a mixed-language (L1/L2) block to estimate local language control (i.e., inhibition of specific lexical candidates) by manipulating language order of the trials. Of note, some researchers estimate local language control by comparing naming across pure- and mixed-language blocks for a particular language (e.g., L1 separately, L2 separately Guo et al., 2011), an effect also known as mixing cost (e.g., Macnamara et al., 1968).
Overwhelmingly, the results of these studies demonstrate that the L1 is differentially affected when the L2 is active across neuroimaging and behavioral methods (reviewed in Kroll & Gollan, 2014). While the L1 naming on average is faster and more accurate than the L2 naming (Gollan & Ferreira, 2009; Hanulová et al., 2010; Ivanova & Costa, 2008; Linck et al., 2008; Sandoval et al., 2010), L1 naming is severely slowed when it follows rather precedes the L2 in block-to-block switching (L1-follow vs. L1-precede blocks) and also in trial-to-trial switching of the mixed-language block (L1-switch vs. L1-no switch trials), similar to asymmetrical switch costs reported above. L1 naming is also slowed by language-mixing (e.g., L1 in the mixed block slower than L1 in the pure block), whereas L2 naming actually is either unaffected or is faster in the mixed-language blocks (reviewed in Kroll & Gollan, 2014). While the L1 inhibition effects occur transiently, some researchers have recently shown long-term effects spanning several blocks (e.g., Misra, Guo, Bobb, & Kroll, 2012).

L1 naming deficits in contexts when the L2 is active support the inhibitory accounts of bilingual language production. Mental firewall accounts counterproposed that while bilinguals may, in fact, use inhibition as a strategy in language-switching paradigms, the inhibition itself per se does not represent language control (e.g., Bobb & Wodniecka, 2013; Verhoeof et al., 2009). More interestingly, there are researchers who have recently launched a call to develop new methods to assess how bilinguals achieve language control during bilingual language production, as the currently developed paradigms, not excluding those discussed here, may no longer adequately assess what happens during bilingual language production (reviewed in Runnqvist et al., 2014). Furthermore, language-switching between single words may not be the universal behavior for all bilinguals, as communicative contexts vary from one bilingual to the next and can also vary as bilinguals adapt to a given task (Green, 2011; Green & Abutalebi, 2013).
Executive control as a candidate for language control

Clearly, bilinguals achieve language control to regulate successfully production. However, it is unclear whether language control is recruited from within the bilingual lexico-semantic system as per mental firewall accounts or whether it is recruited from the domain-general cognitive network outside the bilingual lexico-semantic system, as per inhibitory accounts. If language control is recruited from the domain-general cognitive network, is executive control a viable candidate to be “the language control” that bilinguals use to regulate their production?

Neuroimaging studies of language switching in bilingual language production demonstrate that bilinguals use neural regions networks similar to those responsible for domain-general executive control, especially the left inferior frontal gyrus, the left dorsolateral prefrontal cortex (DLPFC), the left parietal lobule, caudate nuclei, anterior cingulate cortices, and fronto-parietal and cortico-subcortical networks (Abutalebi et al., 2007; Abutalebi, Della Rosa, Ding, et al., 2013; e.g., Abutalebi, Della Rosa, Gonzaga, et al., 2013; Abutalebi et al., 2012; Abutalebi & Green, 2008; reviewed in Collette, Hogge, Salmon, & Van Der Linden, 2006; Hernandez, 2009; Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; reviewed in Hervais-Adelman, Moser-Mercer, & Golestani, 2011; Hirshorn & Thompson-Schill, 2006; Price, Green, & Von Studnitz, 1999; Quaresima, Ferrari, Van Der Sluijs, Menssen, & Colier, 2002; Rodriguez-Fornells, Rotte, Heinze, Nosselt, & Munte, 2002). Additionally, some studies correlated performance on executive control and language-switching tasks. These studies generally find that greater executive control capacity relates to smaller switching costs during bilingual language production (e.g., Linck et al., 2008; Linck et al., 2012). Most recent work on monolingual
language processing further delineates specific role of brain areas associated with executive control in sentence production but not in silent articulator/mouth movements or counting (Geranmayeh, Wise, Mehta, & Leech, 2014).

Given the overwhelming evidence for shared networks for language-switching and executive control, executive control can, in fact, be “the language control” that bilinguals use to regulate their production. However, there is evidence contrary to the claim, arguing that there are inherently qualitative differences between executive control and language control. More specifically, studies measuring linguistic (e.g., switching between languages while naming pictures) and non-linguistic-switching (e.g., switching between naming a color or shape) in the same participants fail to find a one-to-one relationship, often revealing costs for non-linguistic but not linguistic switching (e.g., Calabria, Hernandez, Branzi, & Costa, 2012; Weissberger, Wierenga, Bondi, & Gollan, 2012). These studies, mostly supporting mental firewall accounts, argue that language control in bilingual language production is “not completely subsidiary to the domain-general executive control system” (Calabria et al., 2012). They also argue that the neuroimaging studies reported above often compare different groups of bilinguals performing linguistic-switching and non-linguistic-switching tasks (but see Abutalebi et al., 2012).

**The present work**

Based on the literature reviewed above, it is clear that both known languages activate when bilinguals plan to speak. However, the extent to which cross-language activation results in competition and whether executive control plays a role in bilingual language production varies across different theoretical accounts and across studies. Mental firewall accounts argue that cross-linguistic activation does not result in competition. These accounts, in general, are
agnostic about the role of executive control in production. Conversely, inhibitory accounts argue that cross-linguistic activation almost always results in competition with consequences to processing, especially the L1. With respect to executive control, these accounts make specific predictions about the role of executive control in bilingual language production. The most recent developments now further delineate its role in different communicative contexts.

While efforts to develop these accounts have been gargantuan, little is known about whether and how individual differences among bilinguals on executive control relate to the success with which they produce L1 and L2. While a small set of studies has started to investigate the role of executive control in bilingual language production, they have almost exclusively used single word production methods (e.g., picture naming). Thus, it is unclear whether the results observed in these studies scale up to methods using full sentence and conversational interaction levels. In addition, the vast majority of prior studies often did not assess individual differences among bilinguals on executive control to investigate its role in bilingual language production. As such, the main goal of this thesis is to present novel data to investigate whether and how individual differences in executive control relate to bilingual language production and reading across different communicative contexts.

The thesis contains five chapters, of which three present manuscripts with original data. Across studies in the first two Chapters, native French bilinguals proficient in English (i.e., French-English), and native English bilingual proficient in French (i.e., English-French) completed language production tasks in their respective L1 and L2. Tasks were designed to reflect language production across different communicative contexts, such as semi-spontaneous dialogue/monologue speech, grammatically well-formed sentences, and isolated single words. In the third Chapters native French-English bilinguals completed an L2 sentence reading task. In
all studies, participants also completed a task battery assessing individual differences in L1 and L2 proficiency/knowledge and executive control.

Across studies we applied a similar analytic strategy, in which we used linear mixed effects (LME) regression to model the effects of executive control on bilingual language production (Baayen, 2008). Unlike traditional analysis of variance (ANOVA) type methods, LME regression allows to capture variability brought on by individual differences across participants as it allows to integrate both categorical and continuous variables in the same model. More importantly, LME regression allows researchers to cross participants and items as random effects to capture systematic variability brought on by extraneous factors such as sensitivity to linguistic factors and/or task demands (Baayen, Davidson, & Bates, 2008).

The strength of the present thesis lies in the fact that the effect of executive control on bilingual language production is assessed in three different communicative contexts, such as producing isolated single words, semi-spontaneous grammatically well-formed sentences, and naturalistic conversations. While studies share a common theme, each study also has a particular goal.

**Chapter 2 (Pivneva, Palmer, & Titone, 2012, Frontiers in Psychology):** This chapter presents a study that assessed a link between executive control and spontaneous L1 and L2 speech produced in conversational settings. Additionally, this study assessed whether the link between executive control and bilingual language production were to be modulated by L2 proficiency. Lastly, this study was particularly strong to assess whether executive control related to content (i.e., what was produced) and/or acoustic (i.e., how efficiently speech was produced) indices of bilingual language production. Of note, in this study we referred to executive control as inhibitory control to parallel Green’s Inhibitory Control model (1998). Research
developments after publication of this study have argued that inhibition may in fact be the primary component of executive control but executive control may also impose different types of control (Miyake & Friedman, 2012).

Chapter 3 (Pivneva & Titone, under review, Cognition): This chapter describes a study that was the first to assess whether executive control demands vary as a function of communicative demands, such as producing grammatically well-formed sentences versus single words across single- and dual-language contexts in more experimentally controlled settings. This study was particularly strong to assess bilingual language production across two Experiments: using a novel eye movement monitoring method to study grammatically well-formed sentences in Experiment 1 and traditional behavioral methods to study single word production in Experiment 2.

Chapter 4 (Pivneva, Mercier, & Titone, 2014): This chapter presents a study that assessed a link between executive control and bilingual sentence reading. The particular strength of this study was to link domain-general executive control to attenuation of cross-linguistic activation using eye movement monitoring method and to do so during earliest stages of processing.

Chapter 5 summarizes the results of the studies presented in Chapters 2-4 and discusses them in terms of theoretical models presented in the introduction of the thesis. Further, this chapter discusses implications for the reported findings and presents novel theoretical developments with respect to the role of executive control in bilingual language production and reading. Lastly, this chapter discusses future directions of this line of research.

The final section of this thesis contains References for the research discussed throughout the thesis.
CHAPTER 2:

Inhibitory control and L2 proficiency modulate bilingual language production:

Evidence from spontaneous monologue and dialogue speech

(Pivneva, Palmer, & Titone, 2012, Frontiers in Psychology: Bilingualism &
Cognitive Control Special Issue)
Abstract

Bilingual language production requires that speakers recruit inhibitory control to optimally balance the activation of more than one linguistic system when they produce speech. Moreover, the amount of inhibitory control 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, forty-two 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.
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 & 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 (Dell, Chang, & Griffin, 1999; Griffin & Ferreira, 2006; Levelt, 1989).

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 (Colomé & Miozzo, 2010; De Groot, 2011; Kroll et al., 2008). These added demands include a greater need to manage cross-language competition arising from parallel activation of two languages (Kroll et al., 2008; Kroll et al., 2006), less practice using inhibitory control during L2 speech production (Abutalebi & Green, 2007), and weaker links between conceptual and linguistic representations in the L2 and possibly L1 (Gollan et al., 2008; Poulisse & Bongaerts, 1994). Indeed, recent work suggests that the added
demands of bilingual language processing might lead to enhanced non-linguistic cognitive function for processes necessary for reducing cross-language competition, such as inhibitory capacity and selective attention (Bialystok, 2009; Bialystok, Craik, Klein, & Viswanathan, 2004).

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 Inhibitory Control (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 (Abutalebi & Green, 2007; Green, 1998). 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 (Shallice & Burgess, 1996) suppresses routine goals via inhibitory control operations, 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 inhibitory control to globally suppress the irrelevant but more routine L1 dialogue language schema. As well, within the bilingual lexico-semantic system, inhibitory control 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. Within 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 & Segalowitz, 1983; Segalowitz, 2010; Segalowitz & Hulstijn, 2005), thus requiring inhibitory control (prefrontal function, in particular; see also Petrides,
In contrast, when L2 proficiency is high, L2 production is automatic and less dependent on inhibitory control, although L1 production effort might instead increase due to a weakening of the links between word forms and concepts in the L1 (Bialystok, 2001; Bialystok et al., 2010; Gollan et al., 2008; Gollan et al., 2011; Ivanova & Costa, 2008; Michael & Gollan, 2005).

Thus, the IC model (Green, 1998) and its extension (Abutalebi & Green, 2007) make several logical predictions about the role of inhibitory control during bilingual language production: (1) L2 language production should require greater inhibitory control 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 inhibitory control to occur), and (3) bilinguals should successfully produce language insofar that they intrinsically possess inhibitory control 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 (Gollan & Ferreira, 2009; Hanulová et al., 2010; Linck et al., 2008; Sandoval et al., 2010), and also when they produce extended speech (Towell, Hawkins, & Bazergui, 1996). Moreover, as L2 proficiency increases, language production in a less-dominant L2 improves (De Jong & Wempe, 2009; Kormos, 2006; Poulisse & Bongaerts, 1994). For example, at high L2 proficiency levels picture-naming speed and accuracy become more similar across L2 and L1 (Costa & Caramazza, 1999; Costa &
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 & Dénes, 2004).

Increased L2 proficiency also relates to increased L1 processing effort when bilinguals produce single words in response to a picture (Gollan, Bonanni, & Montoya, 2005; Gollan et al., 2008; Gollan et al., 2011; Ivanova & 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 & 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 & 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 dysfluencies (e.g., uhs, ums) when answering speculation questions (e.g., What makes an ideal friend?) during a dyadic interview than when producing speech without a
conversational partner (e.g., telling a story from a picture) (Fehringer & 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 (Costa, Pickering, & Sorace, 2008; Garrod & Pickering, 2004; Hartsuiker & Pickering, 2008; Hartsuiker, Pickering, & Veltkamp, 2004; Kootstra, Van Hell, & Dijkstra, 2010; Pickering & Garrod, 2004). 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 & 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, Hoshino, and Kroll (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 (Fehringer & Fry, 2007; Gollan, Bonanni, et al., 2005; Green, 1998; Ivanova & Costa, 2008; Kroll et al., 2008; Linck et al., 2008; Poulisse & Bongaerts, 1994). We also predicted that dialogue speech would be more effortful than monologue speech, particularly in the L2 vs. L1 context (Fehringer & 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 inhibitory control 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 & 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 & 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 (the 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 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 (Goldman-Eisler, 1968; Henderson, Goldman-Eisler, & Skarbek, 1966; Kormos, 2006; Kormos & Dénes, 2004; Segalowitz, 2010), and that prior silent pause durations reflect speech planning effort (Chaffe, 1980; Ferreira, 1991; Levelt, 1983; Lindsley, 1975; 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.

**Method**

**Participants**

A total of 22 English-French and 20 French-English bilingual adults (N = 42, M = 21.21, SD = 2.52; 7 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, Blumenfeld, & Kaushanskaya, 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, Samelson, Lee, and Tomblin (2010). Table 1 summarizes self-assessed L2 proficiency measures.
Table 1

_Self-assessed L2 proficiency ratings, language history and standardized L2 proficiency scores (n=42)_

<table>
<thead>
<tr>
<th>Rating scales (0-10)</th>
<th>English-French (n=22)</th>
<th>French-English (n=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Speaking ability</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Reading ability</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Writing ability</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Translating ability</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Listening comprehension</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Fluency</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Grammatical ability</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Overall competence</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

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 (yrs. 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  

<table>
<thead>
<tr>
<th></th>
<th>L1***</th>
<th></th>
<th>L2***</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>82</td>
<td>9</td>
<td>48</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>7</td>
<td>50</td>
<td>17</td>
</tr>
</tbody>
</table>

* Two-tailed independent samples t-test significant at 0.05  
** Two-tailed independent samples t-test significant at 0.01  
*** Two-tailed independent samples t-test significant at 0.001
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 (Appendix A). 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 and a language background questionnaire. Testing session lasted approximately two hours.
Figure 2. Illustration of the procedure in the current study.
Across all monologue and dialogue versions, participants always served as the instruction giver. In the English version of the Map task, we instructed participants in English to verbally guide the instruction follower through a printed route from start to finish, in English (Appendix B). 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, Richardson, & Dale, 2009). During monologues and dialogues, participants viewed maps on a 20-inch 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 & 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 Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999).

**Anti-saccade task.** This task assessed ability to inhibit the pre-potent tendency to look towards a peripherally presented target (Hallett, 1978). We used an Eye-Link 1000 tower mounted system (SR-Research, Ontario, Canada) with a sampling rate of 1kHz to monitor and record fixation durations of the right eye. Participants were presented randomly intermixed prosaccade 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 towards (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, Craik, & Ryan, 2006), where we subtracted the average reaction time of all anti-saccade trials from the average reaction time of all pro-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 & 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 > .05). Using these measures of inhibitory capacity, we computed a standardized composite inhibition cost score (McMurray et al., 2010).

Table 2

Minima, maxima, means, and standard deviations for individual difference measures

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Max</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L1 verbal ability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASI score</td>
<td>8</td>
<td>18</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td><strong>Simon task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent</td>
<td>376</td>
<td>660</td>
<td>489</td>
<td>65</td>
</tr>
<tr>
<td>Incongruent</td>
<td>424</td>
<td>728</td>
<td>527</td>
<td>73</td>
</tr>
<tr>
<td>Cost</td>
<td>-48</td>
<td>-68</td>
<td>-38***</td>
<td></td>
</tr>
<tr>
<td><strong>Stroop task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent</td>
<td>275</td>
<td>844</td>
<td>480</td>
<td>108</td>
</tr>
<tr>
<td>Incongruent</td>
<td>211</td>
<td>783</td>
<td>510</td>
<td>106</td>
</tr>
<tr>
<td>Cost</td>
<td>64</td>
<td>61</td>
<td>-30***</td>
<td></td>
</tr>
<tr>
<td><strong>Number Stroop task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent</td>
<td>446</td>
<td>717</td>
<td>582</td>
<td>72</td>
</tr>
<tr>
<td>Incongruent</td>
<td>475</td>
<td>833</td>
<td>652</td>
<td>87</td>
</tr>
<tr>
<td>Cost</td>
<td>-29</td>
<td>-116</td>
<td>-70***</td>
<td></td>
</tr>
<tr>
<td><strong>Anti-saccade task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-saccade</td>
<td>230</td>
<td>415</td>
<td>304</td>
<td>47</td>
</tr>
<tr>
<td>Anti-saccade</td>
<td>341</td>
<td>528</td>
<td>411</td>
<td>45</td>
</tr>
<tr>
<td>Cost</td>
<td>-111</td>
<td>-113</td>
<td>-107***</td>
<td></td>
</tr>
<tr>
<td><strong>Inhibition cost score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English-French</td>
<td>-0.0099</td>
<td>0.0087</td>
<td>-0.0002</td>
<td>0.0048</td>
</tr>
<tr>
<td>French-English</td>
<td>-0.0070</td>
<td>0.0088</td>
<td>0.0002</td>
<td>0.0037</td>
</tr>
</tbody>
</table>

*** Two-tailed paired samples t-test significant at 0.001

**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.

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 & 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.

Table 3

<table>
<thead>
<tr>
<th>Language</th>
<th>Clarity of instructions</th>
<th>Speaker fluency</th>
<th>Speaker nativeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>6.62</td>
<td>7.49</td>
<td>8.86</td>
</tr>
<tr>
<td>L2</td>
<td>6.05</td>
<td>7.28</td>
<td>6.98</td>
</tr>
</tbody>
</table>

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 < .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 ($t = 2.96$, $p < .01$). 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 ($t = 2.09$, $p < .05$). 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 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.*

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Clarity of instructions</th>
<th>Speaker fluency</th>
<th>Speaker nativeness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>SE</td>
<td>t-value</td>
</tr>
<tr>
<td>Intercept</td>
<td>6.34</td>
<td>0.25</td>
<td>25.08***</td>
</tr>
<tr>
<td>Speech type (monologue, dialogue)¹</td>
<td>0.95</td>
<td>0.32</td>
<td>2.96**</td>
</tr>
<tr>
<td>Language (L1, L2)²</td>
<td>-0.65</td>
<td>32</td>
<td>-2.02*</td>
</tr>
<tr>
<td>L2 proficiency</td>
<td>0.59</td>
<td>0.28</td>
<td>1.75</td>
</tr>
<tr>
<td>Language group³ (English-French vs. French-English)</td>
<td>0.35</td>
<td>0.23</td>
<td>1.53</td>
</tr>
<tr>
<td>Speech type * Language</td>
<td>0.41</td>
<td>0.45</td>
<td>0.90</td>
</tr>
<tr>
<td>Speech type * L2 proficiency</td>
<td>-0.30</td>
<td>0.39</td>
<td>-0.77</td>
</tr>
<tr>
<td>Language * L2 proficiency</td>
<td>0.82</td>
<td>0.39</td>
<td>2.09*</td>
</tr>
<tr>
<td>Speech type * Language * L2 proficiency</td>
<td>-0.81</td>
<td>0.55</td>
<td>-1.46</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random effects</th>
<th>Variance</th>
<th>Variance</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>0.00</td>
<td>0.23</td>
<td>0.16</td>
</tr>
<tr>
<td>Item</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Residual</td>
<td>2.10</td>
<td>1.03</td>
<td>1.01</td>
</tr>
</tbody>
</table>

* *p<0.05 level  ** p<0.01 level  *** p<0.001 level
¹ baseline = monologue
² baseline = L1
³ baseline = English-French
Figure 3. Graphical representation of partialled 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.

**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 ($t = -4.59, p < .001$). 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.

Figure 4. Graphical representation of partialled 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.

**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 < .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 < .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 < .01). Figure 5 shows this interaction across left and right 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.

Figure 5. Graphical representation of partialled 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.
**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 $t’s < 1.53$, $p > .05$).

**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, Homel, Merewether, Martz, & Lomask, 1986; Welkowitz, Bond, Feldman, & Tota, 1990). For the purpose of this study, we 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 (Alpert et al., 1986; Goldman-Eisler, 1968; Kormos, 2006; Segalowitz, 2010; Welkowitz et al., 1990; Wilson & Wilson, 2005). 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.
Table 5

Mean values (ms), standard errors 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

<table>
<thead>
<tr>
<th>Language</th>
<th>Mean (ms)</th>
<th>SE (ms)</th>
<th>Mean observation count</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vocalization duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>2272</td>
<td>280</td>
<td>32.30</td>
</tr>
<tr>
<td>L2</td>
<td>1959</td>
<td>274</td>
<td>40.72</td>
</tr>
<tr>
<td>Dialogue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>1624</td>
<td>237</td>
<td>52.14</td>
</tr>
<tr>
<td>L2</td>
<td>1422</td>
<td>225</td>
<td>66.33</td>
</tr>
<tr>
<td><strong>Prior silent pause duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>689</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>698</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Dialogue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>584</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>596</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td><strong>VD/PPD ratios</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>4.14</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>3.58</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Dialogue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>3.58</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>3.08</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td><strong>Total sample duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monologue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>151169</td>
<td>9292</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>154476</td>
<td>10775</td>
<td></td>
</tr>
<tr>
<td>Dialogue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>494494</td>
<td>27188</td>
<td></td>
</tr>
<tr>
<td>L2</td>
<td>468151</td>
<td>21076</td>
<td></td>
</tr>
</tbody>
</table>

Of note, our custom software also distinguishes speech signal for 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 ($t = -3.93, p < .001$). VD/PPD ratios were smaller in dialogues ($M = 3.33$) than monologues ($M = 3.86$), resulting in a main effect of speech type ($t = -3.48, p < .001$).

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 ($t = -2.60, p < .05$), 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 ratios 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.

![Figure 6](image)

*Figure 6.* Graphical representation of partialled 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.

**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 x speech type x 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 ($t = -0.85$, $p > .05$). However, VD/PPD ratios decreased as inhibitory capacity decreased (inhibition cost increased), resulting in a main effect of inhibitory capacity ($t = -2.20$, $p < .05$). 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 ($t = 2.70$, $p < .01$). 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.
Table 6

Linear mixed effects models for the temporal measure, VD/PPD ratios (ratio of vocalization durations over their prior silent pause durations) to illustrate interactions between speech type, language L2 proficiency and inhibitory capacity.

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>SE</th>
<th>tvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>3.85</td>
<td>0.29</td>
<td>13.16**</td>
</tr>
<tr>
<td>Speech type (monologue, dialogue)(^1)</td>
<td>-0.60</td>
<td>0.17</td>
<td>-3.48***</td>
</tr>
<tr>
<td>Language (L1, L2)(^2)</td>
<td>-0.72</td>
<td>0.18</td>
<td>-3.93***</td>
</tr>
<tr>
<td>L2 proficiency</td>
<td>-0.42</td>
<td>0.28</td>
<td>-1.48</td>
</tr>
<tr>
<td>Inhibitory capacity</td>
<td>-119.51</td>
<td>54.29</td>
<td>-2.20*</td>
</tr>
<tr>
<td>Language group (^3) (English-French vs. French-English)</td>
<td>0.54</td>
<td>0.39</td>
<td>1.39</td>
</tr>
<tr>
<td>Speech type * Language</td>
<td>0.20</td>
<td>0.23</td>
<td>0.87</td>
</tr>
<tr>
<td>Speech type * L2 proficiency</td>
<td>0.50</td>
<td>0.21</td>
<td>2.38*</td>
</tr>
<tr>
<td>Language * L2 proficiency</td>
<td>0.71</td>
<td>0.22</td>
<td>3.25**</td>
</tr>
<tr>
<td>Speech type * Inhibitory capacity</td>
<td>68.83</td>
<td>40.91</td>
<td>1.68</td>
</tr>
<tr>
<td>Language * Inhibitory capacity</td>
<td>13.80</td>
<td>42.82</td>
<td>-0.32</td>
</tr>
<tr>
<td>L2 proficiency * Inhibitory capacity</td>
<td>-120.38</td>
<td>68.69</td>
<td>-1.75</td>
</tr>
<tr>
<td>Speech type * Language * L2 proficiency</td>
<td>-0.71</td>
<td>0.28</td>
<td>-2.60*</td>
</tr>
<tr>
<td>Speech type * Language * Inhibitory capacity</td>
<td>9.80</td>
<td>52.29</td>
<td>0.19</td>
</tr>
<tr>
<td>Speech type * L2 proficiency * Inhibitory capacity</td>
<td>136.17</td>
<td>50.54</td>
<td>2.70**</td>
</tr>
<tr>
<td>Language * L2 proficiency * Inhibitory capacity</td>
<td>-17.54</td>
<td>53.69</td>
<td>-0.33</td>
</tr>
<tr>
<td>Speech type * Language * L2 proficiency * Inhibitory capacity</td>
<td>-55.43</td>
<td>64.88</td>
<td>-0.85</td>
</tr>
</tbody>
</table>

**Random effects**

<table>
<thead>
<tr>
<th></th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>1.30</td>
</tr>
<tr>
<td>Item</td>
<td>0.00</td>
</tr>
<tr>
<td>Residual</td>
<td>12.12</td>
</tr>
</tbody>
</table>

\(^* p < 0.05\) \(^** p < 0.01\) \(^*** p < 0.001\)
\(^1\) baseline = monologue
\(^2\) baseline = L1
\(^3\) baseline = English-French
Figure 7. Graphical representation of partialled 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.

**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.

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 & Pickering, 2004; Hartsuiker & Pickering, 2008; Hartsuiker et al., 2004; Kootstra et al., 2010; Pickering & Garrod, 2004). 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 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 (Chaffe, 1980; Ferreira, 1991; 
Levelt, 1983; Lindsley, 1975; 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.

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 (Fehringer & Fry, 2007; Gollan & Ferreira, 2009; Hanulová et al., 2010; Hernandez, Martinez, & Kohnert, 2000; Kormos & Dénes, 2004; Sandoval et al.,
2010), and that L2 proficiency is an important determinant of L1 and L2 production performance (Costa & Caramazza, 1999; Gollan, Bonanni, et al., 2005; Gollan et al., 2008; Gollan et al., 2011; Ivanova & Costa, 2008; Kormos & Dénes, 2004; Poulisse & Bongaerts, 1994). 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 & 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 & 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 & Pickering, 2004; Hartsuiker & Pickering, 2008; Hartsuiker et al., 2004; Kootstra et al., 2010; Pickering & Garrod, 2004).

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 for all measures 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 interactions. 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 x L2 proficiency x inhibitory capacity interaction, \( t = -2.20, p < .05 \)), 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 with this hypothesis, the correlation between L2 proficiency and inhibitory capacity in our sample was not significant ($r = -.16, p = .31$), 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 inhibitory control, 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 (Abutalebi & Green, 2007; Green, 1998) 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, 2009; Bialystok et al., 2004; Kroll & Bialystok, 2013). 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.
Preface to Chapter 3

The study presented in Chapter 2 investigated how individual differences in executive control relate to spontaneously produced L1 and L2 conversational speech. The study also investigated whether L2 proficiency modulated the effects brought on by executive control. A particular strength of the study was to assess the relationship between executive control and the content versus acoustic measures of production. Another strength was to develop semi-automated methods to study temporal dynamics of bilingual language production in more naturalistic contexts.

Using a modified version of the Map Task, the study revealed that individual differences in executive control related only to acoustic measures of how efficiently bilingual produce conversational speech, such that bilinguals with greater executive control planned and produced speech more efficiently than bilinguals with weaker executive control. In addition, executive control did not relate to the content of what was produced. Unlike executive control, individual differences in L2 proficiency related to both content and acoustic measures of production (i.e., what was produced and how efficiently). Taken together, the study was the first to demonstrate that individual differences in executive control relate to the success of conversational interactions.

While important in its own right, this study represented naturally occurring bilingual language production, which lacks certain degree of experimental control, as we were experimentally at the mercy of what the participants chose to say. The study also did not allow to assess directly how bilinguals resolve linguistic cross-talk. Therefore, in the study reported in Chapter 3 we assessed whether and how executive control relates to bilingual language production using experimentally more controlled methods. The particular aim of the study was
to compare whether and how executive control demands varied as a function of different communicative contexts across two Experiments. In Experiment 1, the study used eye movement measures of bilingual language production to investigate executive control demands when bilinguals produced grammatically well-formed sentences (i.e., more constrained by the experimental task). In Experiment 2, the study used traditional behavioral methods to investigate executive control demands when bilinguals produced single words (i.e., less constrained by the experimental task). We also assessed across both Experiments whether and how demands also varied across single- and dual-language contexts.

Lastly, we assessed cross-linguistic activation differently from the Picture-Word Interference paradigm discussed in the introduction. Specifically, we used name agreement, which measures the many plausible ways an object can be described within a particular language, ranging from high (e.g., sock) to low (e.g., couch vs. sofa) in name agreement. Prior work has successfully demonstrated that naming low name agreement pictures activates brain areas associated with executive control in monolingual speakers (Kan & Thompson-Schill, 2004). No study to our knowledge has assessed name agreement effects in bilingual speakers.

As will be seen in Chapter 3, the results suggest that executive control is overall important in bilingual language production. However, executive control is particularly important when bilinguals produce grammatically well-formed sentences in Experiment 1 versus single words in Experiment 2. In addition, executive control is more specifically important for resolving cross-language activation when inhibiting an entire non-target linguistic system versus specific words found within a language.
CHAPTER 3:

Bilingual language production and individual differences in executive control: Differential effects of producing words in sentences vs. isolation

(Pivneva, & Titone, under review, Cognition)
Abstract

We used eye movement recordings (Experiment 1) and single picture naming (Experiment 2) to investigate the link between bilingual language production, L2 proficiency and individual differences in executive control. In Experiment 1, forty-eight bilinguals produced short sentences to describe picture arrays containing three objects ("The hose and the couch are above the bridge") while their eye movements were monitored, where the second target picture varied in name agreement (e.g., sofa/couch versus book). In Experiment 2, forty-eight bilinguals named the same pictures individually in a single picture naming task that maintained the exact order of presentation as Experiment 1. Across both Experiments, participants encountered L1-only, L2-only, and mixed L1-L2 blocks, and engaged in a separate battery of executive control and language background tests. L2 production was more effortful than L1 production in both experiments, particularly when target pictures had low name agreement. However, greater executive control among bilinguals related to smaller L2 production costs, irrespective of name agreement or language mixing context, in the eye movement sentence production task (Experiment 1) but not during single-picture naming (Experiment 2). This suggests that automatic access of a non-target language during bilingual production recruits domain-general executive control, particularly when language production requires a full multiword grammatical specification (i.e., words are produced in full sentences) but not when words are produced alone. Also noteworthy is that domain-general executive control is recruited even under circumstances that do not involve language mixing.
Bilingual language production and individual differences in executive control: Differential effects of producing words in sentences vs. isolation

Bilinguals exhibit remarkable flexibility in being able to restrict their speech output to only one known language in single language situations (e.g., when speaking to monolinguals), and to fluently switch their languages in multilingual situations (e.g., when speaking to other bilinguals). Indeed, bilinguals typically make few noticeable cross-language errors during production (e.g., saying the French word *divan* when intending to produce the English word *couch*), despite compelling evidence that both languages are automatically activated (Costa, 2005; Hermans et al., 1998; Kroll et al., 2008; Kroll et al., 2006; Kroll, Dussias, et al., 2012; Kroll & Gollan, 2014; Kroll, Guo, & Misra, 2012; Starreveld, De Groot, Rossmark, & Van Hell, 2013). Consequently, non-selective activation of both known languages is as close to being a fact for bilingual production as it is for spoken (Blumenfeld & Marian, 2011; Marian & Spivey, 1999; Mercier et al., 2013; Shook & Marian, 2013; Spivey & Marian, 1999) and written bilingual comprehension (Dijkstra & van Heuven, 2002; Duyck, van Assche, Drieghe, & Hartsuiker, 2007; Van Assche et al., 2012; Van Assche, Duyck, Hartsuiker, & Diependaele, 2009). However, one open question is the extent to which bilingual language production draws upon domain-general cognitive resources, such as executive control (Braver, 2012; Miyake & Friedman, 2012), similar to what has been observed for native language production (Acheson, Hamidi, Binder, & Postle, 2011; Acheson & MacDonald, 2009; Hartsuiker & Barkhuysen, 2006; Pivneva et al., 2012).

Indeed, models of bilingual production differ strikingly in their commitment to whether and how domain-general executive control influences the language production process (Costa, 2005; De Bot, 1992; Green, 1998; Poulisse & Bongaerts, 1994). In this paper, we are
particularly interested in characterizing the role of domain-general executive control as a function of distinct language production demands. We pursue this interest in two ways. First, we examine the role of communicative context on language production across experiments by having bilinguals produce L1 and L2 words in the context of grammatically well-formed sentences (E1) and in isolation (E2). Second, we examine the role of specific kinds of production demands within each communicative context. These include local demands (i.e., suppressing activation of multiple object labels both within- and across-languages) and global demands (i.e., suppressing activation of an entire language system as reflected by a cost for L2 versus L1 production in general). As will be seen, our results suggest that executive control is fundamental to managing global bilingual production demands only when there is a need to produce words in fully elaborated sentences in an L2. However, in building to this conclusion, we first review the literature that leads to our experimental questions and specific manipulations.

Language production is generally thought to occur in a series of cascading stages that include conceptualization of a message, followed by linguistic formulation and articulation of the message (e.g., Kempen & Hoenkamp, 1987; Levelt, 1989). For example, when presented with a picture of an object to be named (e.g., couch), speakers first conceptualize the message and generate an intention to produce the object’s label. Subsequent to conceptualization, speakers retrieve relevant lexical information to formulate this message linguistically, such as retrieving word forms for a presented object (e.g., couch, sofa, loveseat, etc.), and other information (e.g., gender markings). Lastly, speakers retrieve phonological information for lexical candidates that are activated, and then generate an articulatory plan to output the message (e.g., saying the word couch in English). Throughout speech planning, bilinguals must inhibit linguistic information (e.g., lexical forms, such as choosing couch over divan in French during formulation) from the
automatically activated irrelevant language, so that successful production can occur (reviewed in Kroll & Gollan, 2014).

Of relevance here, the effort associated with speech planning is heightened by lexical competition that occurs during the time-course of language production. For example, when speakers retrieve lexical information about an object during the formulation stage, competition may arise locally because a given object has multiple distinct labels within a particular language (e.g., couch vs. sofa vs. loveseat), semantically related neighbors (e.g., couch vs. chair), or phonologically related neighbors (e.g., couch vs. crouch). For successful production to occur, speakers must choose an intended label to produce all while inhibiting other irrelevantly activated labels (e.g., choosing couch and inhibiting sofa, loveseat, chair, crouch).

Here we focus specifically on local competition arising when an object can be described with only one- (e.g., door) or multiple distinct labels (e.g., couch vs. sofa) within a particular language (Alario et al., 2004; Johnson, Paivio, & Clark, 1996), precisely because of its direct relevance to the bilingual situation. By virtue of knowing a second language, bilinguals have at least two labels for each object (e.g., door vs. porte in French), one corresponding to each language (for work on multiple translations for a word see Degani, Prior, & Tokowicz, 2011; Degani & Tokowicz, 2013). Consequently, bilinguals have even more labels for objects that have multiple distinct labels within a language (e.g., couch vs. sofa vs. divan in French). Thus, it is unclear if local competition and speech planning effort vary when bilinguals name objects with multiple distinct labels (e.g., couch vs. sofa vs. divan in French) versus one label (e.g., door vs. porte in French) in the first or second language.

Specifically, objects that have only one usual label within a language are said to have high name agreement, whereas objects with multiple distinct labels are said to have low name
agreement (Snodgrass & Vanderwart, 1980). Variations in within-language name agreement exert robust local effects on picture naming latencies in English (Friedman, Ritter, & Snodgrass, 1996; Szekely et al., 2003; Szekely et al., 2004) and French (Alario & Ferrand, 1999; Alario et al., 2004; Bonin, Chalard, Meot, & Fayol, 2002; Bonin, Peereman, Malardier, Meot, & Chalard, 2003), in that low name agreement pictures are generally more difficult to produce than high name agreement pictures. For example, speakers fixate objects that have low- versus high-name agreement longer before naming them (Griffin, 2001), take longer to produce them (Alario et al., 2004), and preferentially recruit brain areas associated with executive control during lexical selection (e.g., left prefrontal cortex Kan & Thompson-Schill, 2004). Importantly, name agreement effects on language production are independent of those attributable to word frequency and age of acquisition (Lachman, Shaffer, & Hennrikus, 1974; Vitkovitch & Tyrrell, 1995), or other factors such as visual complexity, imageability, image agreement, concept familiarity, number of syllables and number of phonemes (Alario et al., 2004). Thus, an open question is whether bilinguals recruit domain-general executive control to help minimize such demands in real time (e.g., Lev-Ari & Keysar, 2013; Misra et al., 2012).

In addition to local demands arising from factors such as low name agreement, a second and potentially related form of language control is based on global demands that arise as a function of particular communicative contexts. For example, a bilingual speaker may actively down-regulate English at her French-speaking work-place (and up-regulate French), and subsequently do the reverse after returning to her English family. Indeed, the ability of bilinguals to exert global control over all aspects of a language is also likely to require domain-general executive control, similar to what was described above for local demands. Consistent with this view, several studies from the behavioral and electrophysiological domains suggest that
the human brain is capable of engaging in the distinct process of whole language suppression (Christoffels et al., 2007; Guo et al., 2011; Kroll & Gollan, 2014). Thus, it is also possible that domain-general executive control modulates the up- or down-regulation of all lexical forms tagged to a particular language (global demands) in addition to particular word forms (local demands) (De Groot, 2011).

Questions concerning the role of domain-general executive control in bilingual language production are important for distinguishing between bilingual production models. Language-oriented models, which do not emphasize the role of general cognitive capacities, typically extend monolingual production models in two ways (De Bot, 1992; Poulisse & Bongaerts, 1994). First, they posit that bilinguals represent all known languages in an integrated memory store, the contents of which can be automatically accessed following the intention to speak. Second, they posit that bilinguals encode and represent language-specific cues, which become activated when the intention is specific to a particular language (e.g., English, if a French-English bilingual is speaking to someone who only knows English). These language cues provide feed-forward control that biases (partially or otherwise) production output to the intended language (Costa, 2005; Costa et al., 1999; Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; De Bot, 1992; Green, 1998; Poulisse & Bongaerts, 1994; see also BIA+ (Dijkstra & Van Heuven, 2002); Runnqvist, Strijkers, Alario, & Costa, 2012). Those two provisions account for the cognitive mechanisms by which local competition can arise, and also potentially how bilinguals globally suppress all word representations specific to one language. However, they do not provide explicit accounts of how domain-general control modulates the ongoing stages of bilingual language production, given that inhibitory processes within such models are
instantiated only through lateral inhibitory links between representations at any particular level of linguistic processing, or across levels of processing within a core language system. In contrast, other bilingual production models take an action-oriented approach, and place a heavier emphasis on domain-general executive control. The Inhibitory Control (IC) model (Green, 1998), for example, conceives of language production as a communicative action, analogous to non-linguistic physical actions. Like physical actions, bilingual language production consists of routine (usually L1) and non-routine (usually L2) mental task or language schemas that compete to control language output. Thus, for successful L2 production to occur, bilinguals must recruit executive control from the supervisory attention system (SAS; Shallice & Burgess, 1996) to suppress dominant L1 schemas, and to monitor implementation of L2 schemas. Executive control is also hypothesized to fine-tune the relative activation and inhibition of words tagged within each language, so that speakers can select and output the appropriate L2 words. Indeed, both of these control operations occur as a function of L2 proficiency, given that L2 proficiency directly tracks the relative dominance of L1 and L2 language schemas (see also Favreau & Segalowitz, 1983; Segalowitz, 2010; Segalowitz & Hulstijn, 2005).

The IC model is supported by neuroimaging studies showing that a bilingual language production (in the form of single picture naming) recruits a distributed control network that overlaps substantially with that for domain-general cognitive control tasks (Abutalebi, Della Rosa, Ding, et al., 2013; Abutalebi, Della Rosa, Gonzaga, et al., 2013; Guo et al., 2011). The IC model is also supported by behavioral studies of bilingual language production. For example, we previously found that individual differences among bilinguals in executive control and L2 proficiency related to acoustical indices of language production effort in a spontaneous speech
study that used the map task (Pivneva et al., 2012). Specifically, French-English and English-French bilinguals generated navigational instructions to an actual or hypothetical listener in both their L1 and L2, counterbalanced over participants, and also completed a battery of tasks assessing executive control and L2 proficiency. Consistent with the IC model, increased executive control and L2 proficiency among bilinguals patterned with reduced production effort, particularly during L2 speech production. Interestingly, the content of what was produced was unaffected by individual differences in executive control, as assessed by independent ratings of instruction clarity, speaker fluency, or nativeness. While these data cohere with the idea that executive control recruitment is important for the initial stages of speech planning, this study did not allow us to disentangle the independent contributions of locally inhibiting specific word forms from globally inhibiting a non-target language generally. As well, it remains unclear whether these effects would also occur in the context of single picture naming.

Thus, we now investigate this issue across two experiments that have greater methodological specificity. Specifically, bilinguals named pictures that varied in name agreement (sofa/couch vs. door) in their L1 alone, L2 alone, and in L1/L2-mixed language blocks. Participants also completed a battery that assessed both language history and executive control. We tested three predictions across both experiments: (1) L2 production would be more effortful than L1 production as reflected by naming accuracy and latency measures; (2) L1 and L2 production effort would increase as name agreement decreased; and (3) increased executive control would relate to decreased production effort particularly when task demands were high. More specifically, we were particularly interested in whether executive control capacity would relate to the need to inhibit word forms specifically (local demands), a non-target language
globally (global demands) and whether it would differ when bilinguals switch between languages in a mixed language block.

Experiment 1 tested these hypotheses in the context of fully elaborated sentences, when the need to restrict speech output to a specific language was high. Here, we used eye movement measures of semi-spontaneous speech production to assess how individual differences in executive control relate to bilingual language production. While such an approach has been previously established for monolinguals (Griffin, 2001; Griffin & Bock, 2000; Malpass & Meyer, 2010), this study is the first use this paradigm to study bilingual language production, and thus extends prior work using single picture naming to a more naturalistic communicative context (reviewed in Kroll & Gollan, 2014; Linck et al., 2008). In this paradigm, bilinguals produced sentences to describe visually presented picture arrays (Griffin & Spieler, 2006; Malpass & Meyer, 2010), which consisted of three pictures. Participants were instructed to produce the following sentence frames in either their L1 or L2 “The A and the B are above the C” (filler arrays had different picture configurations such that the grammatical frame varied across trials). For the purposes of this paper, we focus our analyses on language production performance for Picture B, which comprised the target picture.

Experiment 2 tested these hypotheses in the context of single picture naming, when the need to restrict speech output to a specific language was only operative at the level of single lexical items. Of note, the exact same pictures were presented in the exact same order as they were named in Experiment 1, thus permitting us to directly compare the effects of individual differences in executive control for two bilingual groups naming the same set of materials but differing only in production output demands elicited by the task (fully elaborated sentences vs. single picture naming).
EXPERIMENT 1

Eye Movement Measures of Sentence Production

Method

Participants

Participants consisted of 48 bilinguals (24 English L1-French L2 and 24 French L1-English-L2, mean age = 21.46, SD = 2.57), with normal or corrected-to-normal vision, and no self-reported speech or hearing disorders. They completed a modified version of the Language Experience and Proficiency Questionnaire to assess self-report L2 proficiency (Marian et al., 2007), and a bilingual animacy judgment task used previously to objectively assess L1 and L2 proficiency (Segalowitz & Frenkiel-Fishman, 2005). During the objective L2 proficiency task, participants decided as quickly and accurately as possible whether a series of nouns were living (cat) or non-living (house), in counterbalanced L1 and L2 blocks. To assess the relative cost of L2 vs. L1 processing, we computed a ratio by dividing the average reaction time on L2 correct trials by those from L1 correct trials for each participant (L2/L1 RT ratio). See Table 1 for averages of self-report and objective L2 proficiency performance.
Table 1

*Self-report and objective L2 proficiency measures for English-French (n=24) and French-English (n=24) bilinguals in Experiment 1*

<table>
<thead>
<tr>
<th>Measure</th>
<th>English-French</th>
<th>French-English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating scales (0-10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall competence***</td>
<td>6.22</td>
<td>8.75</td>
</tr>
<tr>
<td>Sum of self-rating scales (0-100)***</td>
<td>69.13</td>
<td>87.00</td>
</tr>
<tr>
<td>Age of acquisition (yrs. old)*</td>
<td>4.26</td>
<td>6.13</td>
</tr>
<tr>
<td>Choose to speak L2 (%)**</td>
<td>14.76</td>
<td>43.33</td>
</tr>
<tr>
<td>Degree of L1 interference when speaking in L2 (0-5)</td>
<td>2.52</td>
<td>3.00</td>
</tr>
<tr>
<td>Percent of present time spent functioning in each language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1***</td>
<td>82.26</td>
<td>42.04</td>
</tr>
<tr>
<td>L2***</td>
<td>15.09</td>
<td>53.54</td>
</tr>
<tr>
<td>L2 proficiency (L2/L1 RT ratio)***</td>
<td>1.10</td>
<td>0.99</td>
</tr>
</tbody>
</table>

* Two-tailed independent samples t-test significant at 0.05
** Two-tailed independent samples t-test significant at 0.01
*** Two-tailed independent samples t-test significant at 0.001

**Materials**

The entire set of stimulus materials consisted of 324 black and white (300 x 300 pixel) pictures (Alario & Ferrand, 1999; Snodgrass & Vanderwart, 1980; Szekely et al., 2004), none of which were English-French cognates or interlingual homophones. The 324 pictures were arranged into 108 triplets, half of which (n = 54) were named in English and half in French. Of these 108 triplets, 72 were designated as experimental trials (n = 36 for L1 and L2 blocks alone) because they contained the target picture, and 36 were designated as filler trials because they did not contain the target picture (more detail on filler trials are below). Within the experimental set of triplets, English and French pictures were matched on name agreement (Alario et al., 2004; Szekely et al., 2004) and word frequency (Balota et al., 2007; New, Pallier, Ferrand, & Matos, 2001). However, within the set of English and French pictures, the second picture of each array,
which served as the target, varied in name agreement, which we subsequently treat as a continuous rather than categorical variable in all analyses (Appendix G). Name agreement measures were ascertained from prior normative studies (Alario & Ferrand, 1999; Szekely et al., 2004), based upon an H-statistic (Lachman, 1973), which assessed response variability for the proportion of people producing each label for a picture. Thus, lower values of the H-statistic indicate that a large proportion of respondents provided the same name for an image (i.e., low variability), whereas higher values indicate that a small proportion of respondents provided the same name for an image (i.e., high variability). Low and high name agreement pictures did not differ in word frequency within or across English and French. Participants saw each triplet only once in the experiment. Because we wished to assess spontaneous speech production rather than a mixture of speech production and picture label recall, participants were not trained on the pictures prior to the task.

**Procedure**

Participants completed an L1 and L2 (i.e., English or French, depending on the participant’s native language) pure-language block, counterbalanced over participants, which was always followed by a mixed-language block. Across all blocks, each picture was paired with one of two background colors (blue, yellow) that cued which language (L1 or L2) the participant was required to produce on any given trial (see Figure 1). The pairing of color and L1/L2 was also counterbalanced over participants. Thus, in pure language blocks, the background color was always the same, and the background color varied in mixed language blocks. Experimental trials consisted of two pictures above one picture, which participants produced as “The A and the B are above the C” in English (e.g., The duck and the bowl are
above the clock; left panel of Figure 1) or French (e.g., La poupée et la poussette sont au dessus
de la cerise – The doll and the stroller are above the cherry; right panel of Figure 1). Randomly
intermixed filler trials consisted of one picture above two pictures, which participants produced
as “The A is above the B and the C” (e.g., The heart is above the easel and the palm tree; Le
coeur et au dessus du chevalet et du palmier). These filler trials were included to minimize
repetition of grammatical frames across trials (Martin, Crowther, Knight, Tamborello, & Yang,
2010), and thus to increase the spontaneity of speech production.
Figure 1. Example stimuli.

Fixations on each picture were monitored using an EyeLink 1000 tower mounted system (SR-Research, Ontario, Canada) with a sampling rate of 1kHz. Viewing was binocular but we recorded fixations from the right eye only. Participants were instructed to fixate and name the three pictures one at a time, in a left to right, top to bottom sequence. At the beginning of each trial, the second and the third pictures in the display were masked by a black rectangle to block
parafoveal preview, following prior work (Morgan & Meyer, 2005). Thus, when participants first encountered a trial, they saw the first picture in the clear, and the second two pictures were masked with black squares. When the participant’s eyes moved from picture A to picture B, they crossed an invisible boundary, delimited by the masked picture, and the black square was removed, thus making the next picture visible. The same sequence of events happened when participant’s eyes moved from picture B to picture C. As such, participants could not begin planning their speech prior to directly fixating the target image (Malpass & Meyer, 2010). Participants completed each trial by fixating and then using the mouse to click on a small red square presented in the lower right corner of the screen, which triggered onset of the next trial. Ongoing speech output was recorded using an ATR20 microphone, and Experiment Builder (version 1.10.65) was used to synchronize eye movement and vocal responses. The spoken output recorded on each trial was parsed and transcribed by two native-English and two native-French assistants using Transcriber (version 1.5.1).

Following the language production task, participants completed an executive control battery that included non-linguistic Simon and Stroop (Blumenfeld & Marian, 2011) tasks (see Figure 2 for task details, and Table 2 for mean performance). These tasks were implemented using E-Prime (Version 1.0.2, Pittsburgh, PA, USA), and responses were collected using an appropriately labeled button box. The Simon task assessed participants’ ability to ignore a mismatch between the spatial location (the left or right side of the screen) of a stimulus (an up or down arrow) and its required response button (left or right button). Congruent trials in the Simon task occurred when the up or down arrow appeared on the same side of the computer screen as the response; incongruent trials occurred when the up or down arrow appeared on the opposite side of the computer screen as its required response. The Stroop task assessed
participants' ability to ignore a mismatch between semantic meaning assigned to a stimulus (a left or right arrow), its spatial location (the left or right side of the screen) and the response button (the left or right button). Congruent trials in the Stroop task occurred when the arrow pointing to the left or right appeared on the same side of the computer screen as the response; incongruent trials occurred when the arrow appeared on the opposite side of the computer screen as the response. We computed cost scores separately for each task by subtracting the mean correct reaction time (RT) of incongruent trials from that of congruent trials.

Figure 2. Graphic representation of the non-linguistic Simon and Stroop tasks adapted from Blumenfeld and Marian (2011).
Table 2

Means, and standard deviations for reaction times for the congruent and incongruent trials of the Simon and Stroop tasks used in Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Congruent</th>
<th>Incongruent</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Simon task</td>
<td>471</td>
<td>79</td>
<td>504</td>
</tr>
<tr>
<td>Stroop task</td>
<td>468</td>
<td>84</td>
<td>485</td>
</tr>
</tbody>
</table>

Results

We constructed linear mixed effect (LME) models for three dependent variables (DVs) of interest (Baayen, 2008; Bates, 2005): (1) naming accuracy of the target picture, (2) the time spent viewing the target picture (gaze duration), and (3) the time spent viewing the target picture prior to the onset of naming it correctly (gaze-speech latency) (see Table 3).
Table 3

*Means and standard errors of the mean for accuracy, gaze duration, and gaze-speech latency in L1 and L2 in language-pure and language-mixed blocks for pictures with low and high name agreement in Experiment 1*

<table>
<thead>
<tr>
<th></th>
<th>High Name Agreement</th>
<th>Low Name Agreement</th>
<th>Low name agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language-pure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>0.93 0.04</td>
<td>0.64 0.07</td>
<td>0.29</td>
</tr>
<tr>
<td>L2</td>
<td>0.69 0.07</td>
<td>0.42 0.07</td>
<td>0.27</td>
</tr>
<tr>
<td>L2 cost</td>
<td>0.24</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td><strong>Language-mixed</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>0.73 0.06</td>
<td>0.54 0.07</td>
<td>0.19</td>
</tr>
<tr>
<td>L2</td>
<td>0.56 0.07</td>
<td>0.44 0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>L2 cost</td>
<td>0.17</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td><strong>Gaze duration</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language-pure</td>
<td>1093 78</td>
<td>1391 99</td>
<td>-298</td>
</tr>
<tr>
<td>L1</td>
<td>1469 120</td>
<td>1693 131</td>
<td>-224</td>
</tr>
<tr>
<td>L2</td>
<td>-376</td>
<td>-302</td>
<td></td>
</tr>
<tr>
<td>Language-mixed</td>
<td>1159 76</td>
<td>1487 116</td>
<td>-328</td>
</tr>
<tr>
<td>L1</td>
<td>1375 104</td>
<td>1750 121</td>
<td>-375</td>
</tr>
<tr>
<td>L2</td>
<td>-216</td>
<td>-263</td>
<td></td>
</tr>
<tr>
<td>L2 cost</td>
<td>-438</td>
<td>-338</td>
<td></td>
</tr>
<tr>
<td><strong>Gaze-speech latency</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language-pure</td>
<td>1287 80</td>
<td>1514 85</td>
<td>-227</td>
</tr>
<tr>
<td>L1</td>
<td>1725 156</td>
<td>1852 144</td>
<td>-127</td>
</tr>
<tr>
<td>L2</td>
<td>-438</td>
<td>-338</td>
<td></td>
</tr>
<tr>
<td>Language-mixed</td>
<td>1369 90</td>
<td>1659 119</td>
<td>-290</td>
</tr>
<tr>
<td>L1</td>
<td>1610 134</td>
<td>1944 152</td>
<td>-334</td>
</tr>
<tr>
<td>L2</td>
<td>-241</td>
<td>-285</td>
<td></td>
</tr>
</tbody>
</table>
Regarding naming accuracy, responses were considered accurate when participants provided the expected label for a target picture (2066 trials out of 3567 total), or when they provided an acceptable synonym (e.g., stairs versus staircase; 179 trials). When participants did not provide a correct label according to these criteria, they generally provided a superordinate category label (e.g., insect, animal, tool, object), and in very rare cases, they reported not knowing a label (e.g., I don’t know).

Regarding the two latency measures, gaze duration measured how long participants fixated a picture on the first pass before moving their gaze away from the picture, and gaze-speech latency (also known as eye-voice span) measured how long participants took to begin naming a picture starting from the moment it was first fixated. Because people typically move their eyes away from a picture before beginning to name it (Griffin, 2001, 2004; Griffin & Bock, 2000; Griffin & Spieler, 2006; Roelofs, 2007, 2008a, 2008b), we assumed that gaze duration would be less likely than gaze-speech latency to reflect later language production stages that involve articulatory planning.

We computed two sets of models for each DV. The first set examined the basic effects of name agreement on L1 and L2 language production, as a function of pure vs. mixed language blocks. These models included as fixed effects: language of production (L1 versus L2), name agreement (a continuous measure), block (pure- versus mixed-language), and their three- and two-way interactions and main effects. The second set investigated whether individual differences in executive control interacted with these basic effects of L1 vs. L2 language of output, and pure vs. mixed blocks. These were identical to the basic models, but also included an executive control cost score (continuous) as an additional interaction term. To ensure that any effects of executive control could not be attributable to individual differences in L2 proficiency,
we residualized objective L2 proficiency scores (L2/L1 ratio) from the executive control cost score, separately for Simon and Stroop task analyses. Of note, L2 proficiency (L2/L1 RT ratio) did not significantly correlate with Simon (r = -0.16, p > .05) or Stroop inhibition cost (r = -0.10, p > .05) in this young adult sample.

All models included the following control variables: gaze-speech latency and vocalization duration of picture A (i.e., the picture that was named immediately preceding the target picture of interest), participants’ L1 (English versus French), and picture A accuracy. All models also contained maximal random effects (Barr, Levy, Scheepers, & Tily, 2013). Continuous predictor variables were centered and scaled to minimize collinearity (which was < 0.3 across all predictors). Models for accuracy were fitted using logistic regression as they were based on binary data. We removed observations greater than 3 standard deviations from the mean response latency individually for each participant separately for each naming condition (pure-L1/pure-L2/mixed-L1/mixed-L2) for gaze duration (1% of correct trials) and gaze-speech latency (2% of correct trials) to remove extreme outliers. We, further, removed observations less than 200 ms and greater than 6900 ms for gaze duration (1% of correct trials) and gaze-speech latency (1% of correct trials), and then log transformed these data to normalize their distribution. We used a convention of t > 1.96 to report significant effects, and performed Chi Square model comparisons to assess whether the addition of single variables coding for individual differences in executive control was statistically warranted (Baayen, 2008). For the omnibus models, we present results with deviation coded (e.g., -0.5, 0.5) categorical variables, which are interpretable in a manner similar as ANOVA (Barr et al., 2013). However, we followed these omnibus models with treatment coded (e.g., 0, 1; baseline = L1 or pure-language blocks) categorical variables to delineate the source of each interaction, and to create partial effects plots.
We first present the results for the basic models for each dependent variable followed by the executive control (Simon, Stroop) models.

**Effects of Name Agreement, L1 vs. L2 production, and Pure vs. Mixed Language Blocks**

**Accuracy.** Accuracy was reduced for L2 versus L1 production (M = 0.53 versus M = 0.73; b = -1.57, SE = 0.28, z = -5.62); it decreased as name agreement decreased (b = -0.86, SE = 0.17, z = -5.16); and it was reduced for mixed- versus pure-language blocks (M = 0.57 versus M = 0.66; b = -0.72, SE = 0.35, z = -2.04). These main effects were qualified by two significant two-way interactions. First, language and name agreement interacted (b = 0.50, SE = 0.11, z = 4.74), in that the costs of low name agreement were more pronounced for L1 vs. L2 production (Figure 3). Second, language and block interacted (b = 0.80, SE = 0.22, z = 3.58), in that the costs of producing speech in a mixed L1/L2 versus pure-language blocks were again more pronounced for L1 vs. L2 production (Figure 4). There were no other significant interactions or main effects of interest (all z’s < 1.72).
Figure 3. Partial effects plots for an interaction between language of production and name agreement. Greater numbers of name agreement represent multiple distinct labels (lower name agreement among native speakers), whereas lower numbers represent fewer distinct labels (higher name agreement). The negative effects of low name agreement (lower accuracy) were more pronounced for L1 versus L2 production.
**Figure 4.** The effect of language of production and block on naming accuracy. The negative effects of producing language in a mixed L1/L2 versus pure-language blocks were more pronounced for L1 than L2.

**Gaze Duration.** Gaze duration during L2 production was longer than that during L1 production (M = 1572 versus M = 1262 ms; b = 0.26, SE = 0.06, t = 4.68); and it increased as name agreement decreased (b = 0.14, SE = 0.02, t = 5.60). In contrast with the accuracy data, however, gaze duration did not differ between pure- and mixed-language blocks (b = 0.06, SE = 0.05, t = 1.33). No other interactions or main effects were significant (all t’s < 1.22).

**Gaze-Speech Latency.** The results for gaze-speech latency were similar to those for gaze duration. Gaze-speech latency during L2 production was longer than that during L1 production (M = 1782 versus M = 1430 ms; b = 0.23, SE = 0.05, t = 5.10), and increased as name agreement decreased (b = 0.09, SE = 0.02, t = 4.46). Again, gaze-speech latency did not significantly differ
between pure- and mixed-language blocks (b = 0.05, SE = 0.04, t = 1.31). There were no other significant interactions or main effects of interest (all t’s < 1.30).

**Summary.** Taken together, low name agreement negatively affected both L1 and L2 production, however, this effect was greater for L1 vs. L2 production in terms of accuracy. As well, L2 production accuracy was lower overall than L1 accuracy, and naming accuracy was higher overall in pure vs. mixed L1/L2 blocks. Finally, the two measures of speech planning effort (target picture gaze duration and gaze-speech latency) showed that low name agreement increased production effort, as did producing L2 vs. L1 speech, however, name agreement and language did not interact. Interestingly, there were no effects of pure vs. mixed language blocks across both latency measures (there was an expected effect for naming accuracy). This is likely because our target was the second picture in the array rather than the first, a key point to which we later return.

**Effects of Individual Differences in Executive Control**

The models presented here are identical to those above with the one exception that interaction terms are added for individual differences in executive control as measured first by the Simon arrow task, and then the Stroop arrow task. L2 proficiency was residualized from both measures in all models. We focus only on effects involving executive control.

**Simon Task.** There was a significant interaction between production block and Simon cost (b =0.24, SE = 0.10, z = 2.38) (Figure 5) and this model explained significantly more variance than a simpler model including only main effects ($\chi^2(1) = 5.16, p < .05$). To better understand this interaction, we performed a follow-up model that median split the bilinguals into high and low Simon cost groups, and tested whether naming accuracy was greater in the pure vs.
mixed blocks. For bilinguals with low Simon cost scores (i.e., greater executive control), naming accuracy in the pure block was significantly greater than in the mixed block ($b = -1.08$, $SE = 0.35$, $z = -3.11$). However, for bilinguals with high Simon cost scores (i.e., weaker executive control), there was no difference between pure and mixed block naming accuracy ($b = -0.64$, $SE = 0.35$, $z = -1.86$). There were no other significant interactions involving executive control (all $z$’s $< 0.99$).

Figure 5. Partial effects plots for an interaction between performance on a Simon task and production block. Greater numbers on a Simon cost score represent weaker executive control (i.e., greater cost), whereas smaller numbers represent greater executive control. Bilinguals with greater executive control named pictures more accurately in the pure-language blocks than bilinguals with weaker executive control. Performance on a mixed L1/L2 block did not significantly vary as a function of executive control.
There were no significant effects for Simon cost for gaze duration (all t’s < 1.47). However, with respect to gaze-speech latency, Simon cost significantly interacted with language of production (b = 0.08, SE = 0.04, t = 2.03) and this model explained significantly more variance than a simpler model including only main effects ($\chi^2(1) = 4.37, p < .05$). As can be seen in Figure 6, gaze-speech latency was longer for L2 versus L1 production for bilinguals with weaker executive control, however, this effect of language did not differ for bilinguals with greater executive control. No other effects involving Simon cost were significant (all t’s < 0.60).

*Figure 6.* Partial effects plots for an interaction between performance on a Simon task and language of production. Greater numbers on a Simon cost score represent weaker executive control (i.e., greater cost), whereas smaller numbers represent greater executive control.
Bilinguals with weaker executive control named pictures slower in their L2 than L1, whereas bilinguals with greater executive control named pictures as quickly in L2 as in L1.

**Stroop Task.** Unlike the Simon task, there were two significant interactions for the Stroop task. First, for naming accuracy there was an interaction between language and Stroop cost (b = -0.64, SE = 0.29, z = -2.23), in that the cost in L2 vs. L1 naming accuracy was greater for bilinguals who had weaker executive control (Figure 7a). Second, there was a significant interaction between pure vs. mixed block and Stroop cost (b =0.38, SE = 0.17, z = 2.30) (Figure 7b). This model explained significantly more variance than a simpler model including only main effects ($\chi^2(2) = 7.24, p < .05$). To better understand the source of this interaction, we computed follow-up models that median split the bilinguals into high vs. low Stroop cost groups to assess the effect of pure vs. mixed block in each group separately. Similar to what was found for the Simon task, accuracy was greater in the pure- versus mixed blocks for bilinguals with greater executive control (b = -0.83, SE = 0.35, z = -2.39), but this difference was not significant between pure and mixed block performance for the bilinguals with weak executive control (b = -0.39, SE = 0.39, z = -1.00). There were no other significant effects involving Stroop cost for production accuracy (all z’s < 1.31).
Figure 7. Partial effects plots for an interaction between performance on a Stroop task and language of production (left panel; 7a) and block of production (right panel; 7b). Greater numbers on a Stroop cost score represent weaker executive control (i.e., greater cost), whereas smaller numbers represent greater executive control. Bilinguals with weaker executive control named pictures less accurately in their L2 than L1, whereas bilinguals with greater executive control named pictures as accurately in L2 as in L1 (Figure 7a). Bilinguals with greater executive control named pictures in the pure-language blocks more accurately than bilinguals with weaker executive control (Figure 7b). Performance on a mixed L1/L2 block did not significantly vary as a function of executive control.

With respect to production latency, Stroop cost significantly interacted with language for gaze duration ($b = 0.12$, $SE = 0.06$, $t = 2.21$), in that gaze durations were longer for L2 versus L1 production for bilinguals with weaker executive control. The effect of language did not seem to differ for bilinguals with greater executive control (Figure 8). This model explained significantly
more variance than a simpler model including only main effects ($\chi^2(1) = 4.12, p < .05$). No other effects involving Stroop cost were significant for gaze duration (all t’s < 1.65), or for gaze-speech latency (all t’s < 1.44).

Figure 8. Partial effects plots for an interaction between performance on a Stroop task and language of production. Greater numbers on a Stroop cost score represent weaker executive control (i.e., greater cost), whereas smaller numbers represent greater executive control. Bilinguals with weaker executive control named pictures slower in their L2 than L1, whereas bilinguals with greater executive control named pictures as quickly in L2 as in L1.
Discussion

Experiment 1 used eye movement measures of semi-spontaneous speech production to assess how individual differences in executive control relate to bilingual language production. We found that L2 production overall was more effortful than L1, as evidenced by naming accuracy and latency measures (i.e., gaze duration, gaze-speech latency). We also found that production effort increased for low name agreement pictures with multiple distinct labels across all measures. As well, we found that L1 speech production was differentially affected by the mixed-language context, such that it was more effortful in the mixed- versus pure-language contexts (Kroll & Gollan, 2014). Because this effect only appeared in naming accuracy, it possibly reflects an absence of L2 knowledge. However, L1 speech production did not significantly differ across pure- versus mixed-language blocks when bilinguals did access their L2 knowledge on correct trials, as evidenced by latency measures.

Lastly, and of direct relevance to our main experimental question, greater executive control capacity among bilinguals reduced the overall cost associated with L2 production across naming accuracy and latency measures (i.e., gaze duration, gaze-speech latency), as evidenced by Simon and Stroop tasks cost scores interacting with language of production. Greater executive control also reduced the overall cost of language production associated with language-mixing, as evidence by an interaction between executive control (both Simon and Stroop cost scores) and production block for naming accuracy. Interestingly, however, executive control among bilinguals did not relate to local demands of bilingual language production arising from variations in name agreement. However, again it is possible that bilinguals did not know all possible labels for a given target picture (e.g., a bilingual might only know one word for couch in the L2 and not all other alternatives). To evaluate this possibility, we computed new models to
examine whether individual differences in L2 proficiency among bilinguals (after residualizing out executive control capacity) would relate to specific word-form inhibition (local demands) (the details may be found here\(^1\)). The outcome of these new models showed that L2 proficiency did indeed relate to name agreement, such that decreased L2 proficiency led to greater costs for low name agreement pictures, especially during L2 production. However, these effects were found for the latency of correct responses rather than accuracy. Thus, it is unlikely that the absence of knowledge was responsible for the name agreement effects observed here, but rather that speech planning demands increased as name agreement decreased, particularly for low L2 proficiency bilinguals performing the task in their L2.

Taken together, the results of the Experiment 1 advance our understanding of how individual differences in executive control relate to semi-spontaneous bilingual speech production. However, several new questions arise. First, given past work from the single picture naming literature (reviewed in Kroll & Gollan, 2014), it is somewhat surprising that the mixed-versus pure-language blocks did not differ in planning latency, although there were significant differences for naming accuracy. However, as previously mentioned, it is possible that the second picture of the array, which occurred mid-sentence, was protected from language-switching effects usually observed in mixed-language contexts insofar as the first picture of the display bore the brunt of language-switch effect. To evaluate this possibility, we computed

\(^1\) The results revealed that increased L2 proficiency led to reduced costs in L2 production for accuracy \((b = -0.93, SE = 0.28, z = -3.34)\), gaze duration \((b = 0.12, SE = 0.05, t = 2.24)\), and gaze-speech latency \((b = 0.13, SE = 0.04, t = 2.98)\). Moreover, increased L2 proficiency reduced speech planning costs associated with naming pictures with multiple labels (e.g., negative effects of low name agreement) in L2 only, as evidenced by a significant three-way interaction between language, name agreement and L2 proficiency for gaze duration \((b = 0.06, SE = 0.03, t = 2.08)\) and gaze-speech latency \((b = 0.06, SE = 0.03, t = 2.34)\) only. This interaction was not found for accuracy \((b = 0.11, SE = 0.12, z = 0.89)\), suggesting that it was not likely caused by an absence of L2 knowledge but rather by increased effort in accessing knowledge.
models examining productions for the first picture in the display (Picture A)\textsuperscript{2}. Indeed, these analyses clearly show the expected pure versus mixed block effect on the first picture of the array, in that mixed-language context differentially affected L1 but not L2.

A second question is whether domain-general executive control demands would arise for bilingual production tasks that have relatively lower demands associated with the communicative context. That is, target pictures in Experiment 1 were produced in the context of a fully specified grammatical frame, which also involved the production of other pictures in the array. This likely resulted in a highly demanding communicative context that would create a greater demand for the domain general executive control. While it is possible some amount of syntactic priming across similar syntactic frames occurred (Hartsuiker et al., 2004; Kootstra et al., 2010), thereby reducing the overall demands of bilingual language production, the inclusion of filler trials, which required people to produce an alternate syntactic frame, would have minimized any priming benefit. To evaluate the question of whether the need to name pictures in a fully articulated sentence (however stereotyped) was more demanding than naming pictures in isolation, we conducted a single-picture naming study with a new set of bilingual speakers in Experiment 2. Thus, participants named pictures one at a time without a determiner but in the exact same order as the same pictures appeared in triplets in Experiment 1. Participants also completed the same executive control and L2 ability battery.

More specifically, Experiment 2 addressed two questions arising from Experiment 1. The first involves the absence of the expected language by pure vs. mixed block effect (though we did observe this effect for naming accuracy and for speech planning measures when we

\textsuperscript{2} We did observe a significant interaction between language and block for gaze-speech latency of picture A ($b = -0.19, SE = 0.04, t = -4.42$), in that L1 production was more effortful in the mixed vs. pure-language block but L2 production did not significantly differ between pure and mixed-language blocks.
analyzed the first picture in the array (Kroll & Gollan, 2014). The second involves whether executive control demands found in Experiment 1 for semi-spontaneous speech be comparable to when communicative demands only require isolated picture naming. To the extent that single picture naming was less effortful overall than the semi-spontaneous sentence production task used in Experiment 1, we would expect that individual differences in executive control would have a reduced role in modulating L1 vs. L2 production effects.

EXPERIMENT 2

Single picture naming

Method

Participants

Participants consisted of 48 bilinguals (24 English L1-French L2 and 24 French L1-English L2, mean age = 22.23, SD = 2.00), with normal or corrected-to-normal vision, and no self-reported speech or hearing disorders. Participants completed the same battery of tasks assessing their language ability and executive control as in Experiment 1 (see Tables 1 & 2). Participants from Experiment 2 did not significantly differ from participants in Experiment 1 on objective L2 proficiency (L2/L1 RT ratio, t = -1.07, p > .05), residualized Simon (t = 0, p > .05) or Stroop measures (t = 0, p > .05), and self-reported measures, such as L2 speaking ability (t = 0.88, p > .05), L2 overall competence (t = 0.71, p > .05), L2 AoA (t = -0.99, p > .05), how often a bilingual would choose to speak L2 (t = 0.91, p > .05), current L1 exposure (t = -0.07, p > .05), and lastly current L2 exposure (t = 0, p > .05).
Table 1

Self-report and objective L2 proficiency measures for English-French (n=24) and French-English (n=24) bilinguals completing the single picture-naming task in Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>English-French</th>
<th>French-English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rating scales (0-10)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall competence***</td>
<td>6.10</td>
<td>8.31</td>
</tr>
<tr>
<td>Sum of self-rating scales (0-10)***</td>
<td>60.19</td>
<td>81.30</td>
</tr>
<tr>
<td>Age of acquisition (yrs. old)*</td>
<td>4.79</td>
<td>6.88</td>
</tr>
<tr>
<td>Choose to speak L2 (%)**</td>
<td>15.26</td>
<td>34.92</td>
</tr>
<tr>
<td>Degree of L1 interference when speaking in L2 (0-5)*</td>
<td>4.67</td>
<td>3.67</td>
</tr>
<tr>
<td>Percent of present time spent functioning in each</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1***</td>
<td>75.29</td>
<td>48.26</td>
</tr>
<tr>
<td>L2***</td>
<td>19.83</td>
<td>50.22</td>
</tr>
<tr>
<td>L2 proficiency (L2/L1 RT ratio)***</td>
<td>1.18</td>
<td>0.99</td>
</tr>
</tbody>
</table>

* Two-tailed independent samples t-test significant at 0.05
** Two-tailed independent samples t-test significant at 0.01
*** Two-tailed independent samples t-test significant at 0.001

Table 2

Means, and standard deviations for reaction times on congruent and incongruent trials for Simon and Stroop task in Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Congruent</th>
<th>Incongruent</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M  SD</td>
<td>M  SD</td>
<td></td>
</tr>
<tr>
<td>Simon task</td>
<td>462 79</td>
<td>502 85</td>
<td>-40</td>
</tr>
<tr>
<td>Stroop task</td>
<td>448 90</td>
<td>470 105</td>
<td>-22</td>
</tr>
</tbody>
</table>

Materials

The picture materials were identical to those used in Experiment 1.
Procedure

Similar to Experiment 1, participants completed two counterbalanced L1 or L2 (i.e., English or French, depending on participant’s L1) pure-language blocks, followed by a mixed-language block. Across all blocks, participants saw one picture at a time presented in the middle of the screen, paired with one of two counterbalanced background colors (blue, yellow) that cued which language (L1 or L2) to use on a given trial. Picture-color pairing was identical to that of Experiment 1. Participants were instructed to name each picture with the first label that came to their mind that best described the picture, and to then type in their response. Participants were also instructed not use determiners (e.g., the, a), thus, they were only required to produce a single label during L1 or L2 (i.e., English or French) production.

Of note, pictures were presented in exactly the same order that participants had experienced them in Experiment 1. Thus, if a display in Experiment 1 consisted of an A and a B above a C, then participants in Experiment 2 first named Picture A, followed by Picture B, and followed by Picture C. Thus, the target picture of interest (Picture B) always followed a picture that was named in the same language, even during the mixed language block.

Speech output was recorded using ATR20 microphone. Voice onset times of each word were measured using a voice-key. Once participants triggered the voice key, the background color display changed to white and the picture remained on the screen until participants finished typing their response. At the same time all responses and voice-key triggering were also recorded and checked by the experimenter for every trial to ensure that participants triggered the voice key based on the actual response and that they typed their initial response.
Results

There were two DVs of interest: picture naming accuracy, and voice onset RT for correctly named pictures. Similar to Experiment 1, we computed LME models for each dependent variable separately (see Table 3 for descriptive statistics of each DV). We used similar criteria as in Experiment 1 for determining response accuracy, according to which there were 721 incorrect trials out of 3240 total trials across all participants. Voice onset RTs reflected the duration of time between a picture appearing on the screen to the moment a participant began to produce its correct label.
Table 3

Means and standard errors of the mean for accuracy and voice onset RTs in L1 and L2 in language-pure and language-mixed blocks for pictures with low and high name agreement in Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>High name agreement</th>
<th>Low name agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
</tr>
<tr>
<td><strong>Accuracy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language-pure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>0.96</td>
<td>0.03</td>
</tr>
<tr>
<td>L2</td>
<td>0.69</td>
<td>0.07</td>
</tr>
<tr>
<td>L2 cost</td>
<td>0.27</td>
<td>0.26</td>
</tr>
<tr>
<td>Language-mixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>0.90</td>
<td>0.04</td>
</tr>
<tr>
<td>L2</td>
<td>0.74</td>
<td>0.06</td>
</tr>
<tr>
<td>L2 cost</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Voice onset RT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language-pure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>1289</td>
<td>88</td>
</tr>
<tr>
<td>L2</td>
<td>1770</td>
<td>160</td>
</tr>
<tr>
<td>L2 cost</td>
<td>-481</td>
<td>-193</td>
</tr>
<tr>
<td>Language-mixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>1731</td>
<td>127</td>
</tr>
<tr>
<td>L2</td>
<td>1977</td>
<td>176</td>
</tr>
<tr>
<td>L2 cost</td>
<td>-246</td>
<td>-245</td>
</tr>
</tbody>
</table>
Similar to Experiment 1, we first computed models that assessed whether our experimental manipulations were successful, and these included the following as fixed effects: language of production (L1 versus L2), name agreement (a continuous measure), block (pure-versus mixed-language), and their three-way and two-way interactions and main effects. We then computed models that assessed whether individual differences in executive control interacted with these basic effects. The composition of these models was identical to the basic models, but also included an executive control cost score (continuous) as an interaction term (Simon cost, then Stroop cost, as in Experiment 1). Again, we residualized objective L2 proficiency (L2/L1 RT ratio) from the executive control cost score (separately for Simon and Stroop tasks). Unlike Experiment 1, increased L2 proficiency significantly correlated with Simon \( r = 0.42, p < .01 \) but not with Stroop cost scores \( r = 0.09, p > .05 \) in this sample.

Models also included accuracy and a voice onset time for picture A to control for any carry-over difficulties. As in Experiment 1, all models contained maximal random effects structures (Barr et al., 2013). We centered and scaled all continuous predictor variables to minimize collinearity \( (< 0.2) \), and used logistic regression for models involving accuracy. We excluded observations where participants failed to trigger the voice key correctly \( (9\% \text{ of correct trials}) \). We removed observations greater than 3 standard deviations from the mean response latency individually for each participant separately for each naming condition \( \text{pure-L1/pure-L2/mixed-L1/mixed-L2} \) for voice onset RTs \( (1\% \text{ of correct trials}) \) to remove extreme outliers. We, further, removed observations less than 200 ms and greater than 6900 ms for voice onset RTs \( (2\% \text{ of correct trials}) \), and then log transformed these data to normalize their distribution. We used a convention of \( t > 1.96 \) to report significant effects, and performed Chi Square model comparisons to assess whether the addition of single variables coding for individual differences
in executive control was statistically warranted (Baayen, 2008). For the omnibus models we present results with the deviation coded (e.g., -0.5, 0.5) categorical variables to interpret the results in a manner similar as ANOVA. We also followed the omnibus models with treatment codings (e.g., 0,1; baseline = L1 or pure-language blocks) for the categorical variables to clarify the source of each interaction, and to create partial effects plots.

**Effects of Name Agreement, L1 vs. L2 production, and Pure vs. Mixed Language Blocks**

*Accuracy.* Picture naming accuracy was reduced for L2 versus L1 production ($M = 0.62$ versus $M = 0.84$; $b = -2.34$, SE = 0.25, $z = -9.38$); it decreased as name agreement decreased ($b = -1.39$, SE = 0.29, $z = -4.87$), but it did not differ for pure- and mixed-language blocks ($b = -0.48$, SE = 0.50, $z = -0.96$). These main effects were qualified by two significant two-way interactions. First, language and name agreement interacted ($b = 0.58$, SE = 0.18, $z = 3.31$), in that the costs of low name agreement were more pronounced for L1 vs. L2 production (Figure 1). Second, language and block interacted ($b = 0.74$, SE = 0.32, $z = 2.30$), in that the costs of producing speech in a mixed L1/L2 versus pure-language blocks were again more pronounced for L1 vs. L2 production (Figure 2). No other interactions or main effects were significant (all $z$’s < 0.96).
Figure 1. Partial effects plots for an interaction between language of production and name agreement. Greater numbers of name agreement represent multiple distinct labels (lower name agreement among native speakers), whereas lower numbers represent fewer distinct labels (higher name agreement). The negative effects of low name agreement (lower accuracy) were more pronounced for L1 versus L2 production.
Figure 2. The effect of language of production and block on naming accuracy. The negative effects of producing language in a mixed L1/L2 versus pure-language blocks were more pronounced for L1 than L2.
Voice onset RT. Voice onset RTs during L2 production were longer than that during L1 production (M = 1919 versus M = 1583 ms; b = 0.26, SE = 0.04, t = 6.82); they increased as name agreement decreased (b = 0.15, SE = 0.03, t = 5.49); and they were longer in the mixed-versus pure-language blocks (M = 1907 versus M = 1618 ms; b = 0.19, SE = 0.06, t = 3.24). These effects were qualified by several two-way interactions and a significant three-way interaction.

First, language and name agreement interacted (b = -0.08, SE = 0.03, t = -3.00), in that the costs of low name agreement were more pronounced for L1 vs. L2 production. Second, language and block interacted (b = -0.15, SE = 0.05, t = -3.01), in that the costs of producing speech in a mixed L1/L2 versus pure-language blocks was again more pronounced for L1 vs. L2 production. Third, block and name agreement interacted (b = -0.10, SE = 0.05, t = -2.18), in that the costs of low name agreement were more pronounced for the pure-language vs. mixed L1/L2 blocks. Lastly, language, name agreement and block interacted (b = 0.07, SE = 0.04, t = 2.02), in that the costs of low name agreement were more pronounced for L1 vs. L2 production in the pure-language but not in the mixed L1/L2 blocks (Figure 3).
Effects of Individual Differences in Executive Control

*Simon Task.* With respect to naming accuracy, there was a significant main effect of executive control \( (b = -0.46, \ SE = 0.18, \ z = -2.54) \), in that accuracy decreased globally as Simon RT cost increased. This model explained significantly more variance than a simpler model excluding executive control \( (\chi^2(1) = 6.14, \ p < .05) \). With respect to voice onset RTs, there was only a significant main effect of executive control \( (b = 0.07, \ SE = 0.03, \ t = 2.11) \), where voice onset RTs increased as Simon RT cost increased. This model explained significantly more variance than a simpler model excluding executive control \( (\chi^2(1) = 8.27, \ p < .01) \). No other effects involving Simon RT cost were significant.
**Stroop Task.** With respect to naming accuracy, there was again a significant main effect of executive control (b = -0.37, SE = 0.17, z = -2.15), in that accuracy decreased as Stroop RT cost increased. This model explained significantly more variance than a simpler model excluding executive control ($\chi^2(1) = 4.99, p < .05$). With respect to voice onset RT, there were no significant effects involving executive control (all t’s < 1.26).

**Discussion**

In Experiment 2, we assessed how individual differences in executive control relate to bilingual language production of single pictures. Similar to Experiment 1, L2 versus L1 production was more effortful for both naming accuracy and latency (i.e., voice onset RT). We also observed the negative effects of low name agreement in that production effort increased with increases in multiple distinct word labels, as evidenced by naming accuracy and latency. Further, this effect was most pronounced for L1 versus L2 production in the pure-language, such that L2 versus L1 speech planning efforts were overall higher but not different from each other for low name agreement pictures. This suggests that speech planning effort does not differ across L1 and L2 in the presence of multiple competitors. That this effect was most salient during pure- versus mixed L1/L2 blocks, reveals that in the mixed L1/L2 block all pictures behaved like low name-agreement pictures irrespective of language. Of note, we found that L1 speech production was differentially affected by language mixing and, thus, was more effortful in the mixed- versus pure-language contexts. This suggests that bilinguals had greater difficulties in reactivating and accessing L1 knowledge when switching languages in the mixed- versus in the pure-language blocks (Kroll & Gollan, 2014).
With respect to individual differences in executive control, we found that increased executive control reduced the cost associated with the overall language production globally, such that it was not language-specific to the L1 or L2 alone. This effect was evidenced by Simon cost score for voice onset RTs and Simon and Stroop cost scores for accuracy. Similar to Experiment 1, executive control among bilinguals did not relate to local demands of bilingual language production, but rather to the overall demands of language production more generally.

Taken together, results of Experiment 2 are generally consistent with our predictions that (1) L2 production is more effortful than L1; (2) language production effort increases when bilinguals name pictures with multiple distinct labels, (3) executive control capacity related to global aspects of production, but not specifically to language or lexical form selection to overall general successful language production but not to specific language or lexical form selection.

General Discussion

Across two experiments we investigated two questions about bilingual language production and its relation to domain-general executive control. First, we investigated whether there was a link between individual differences in domain-general executive control and bilingual language production in different communicative contexts: in the context of well-formed sentences in Experiment 1, and in the context of single words in Experiment 2. Second, we assessed whether individual differences in executive control modulated language production for specific kinds demands within a particular communicative context, such as local demands (i.e., inhibition of particular lexical forms both within- and across-languages as reflected by name agreement effects) and global demands (i.e., inhibiting activation of an entire language system as reflected by a cost for L2 versus L1 production in general). We tested three predictions: (1) L2
versus L1 production would be more effortful; (2) L1 and L2 production effort would increase as
name agreement decreased (i.e., negative effect of name agreement); and (3) greater executive
control capacity would relate to decreased production effort.

With respect to our first prediction, language production was more effortful during L2
versus L1 production, as reflected by reduced naming accuracy and increased latency measures
across both experiments. In other words, bilinguals named pictures less accurately and slower
than in their L2 than in their L1. This result replicates prior work from single-word picture-
naming studies showing that bilingual name pictures in their less-dominant L2 less accurately
and slower than in their more-dominant L1 (reviewed in Kroll & Gollan, 2014). This result also
replicates prior work on spontaneous speech production in conversations showing that bilinguals
plan and produce speech with less clarity and efficiency in their less-dominant L2 than in their
more-dominant L1 (Pivneva et al., 2012).

With respect to our second prediction, production effort increased as name agreement
decreased, as reflected by reduced naming accuracy and increased latency measures again across
both experiments. In other words, bilinguals named pictures with multiple distinct labels within
a language less accurately and slower than those with fewer multiple distinct labels. This result
also replicates prior studies on native-language production showing that low name agreement
pictures are generally more difficult to produce than high name agreement pictures (Alario et al.,
2004; Griffin, 2001; Kan & Thompson-Schill, 2004). For example, speakers fixate longer
(Griffin, 2001), take longer to begin articulating (Alario, et al., 2004), and recruit to a greater
extent brain areas linked to executive control (Kan & Thompson-Schill, 2004) when naming
pictures with low- versus high-name agreement. Here, we extend these effects to bilingual
language production.
With respect to our third prediction, individual differences in domain-general executive control modulated bilingual language production as a function of (1) communicative context, and (2) specific kinds of production demands within a particular communicative context. First, executive control related to particular communicative contexts in that bilinguals recruited executive control to a greater extent when language-specific demands were higher, such as producing grammatically well-formed sentences in Experiment 1 versus single words in isolation in Experiment 2. In other words, bilinguals with weaker executive control capacity, as measured by performance on Simon and Stroop tasks, named pictures in grammatically well-formed sentences less accurately and slower in L2 than in L1, while bilinguals with greater executive control capacity named pictures as accurately and quickly across their two languages. That this result was observed when producing full sentences in Experiment 1 and not when producing single words in Experiment 2, is consistent with our previous work showing that bilinguals recruit executive control to a greater extent when they need to produce L1 and L2 speech spontaneously and for an extended period of time, such as in a naturally-occurring conversation with an interlocutor (Pivneva, et al., 2012). This result is also consistent with prior work showing that when bilinguals rate in a particular language context the extent to which pictures are similar when they share a label (e.g., house key versus computer keyboard key), bilinguals with greater executive control capacity assign higher similarity ratings (i.e., indices of language-specificity) than bilinguals with weaker executive control capacity (Lev-Ari & Keysar, 2013).

That we did not observe a significant relationship between executive control and single-word production, such as found in picture naming, is not entirely consistent with prior work. For example, a study by Linck and colleagues found that greater executive control related to more successful bilingual language production, as assessed by single picture naming (Linck et al.,
Bilinguals in their study named cognate (e.g., piano shares meaning across English and French) and non-cognate (e.g., language-unique words) pictures in their L2. Participants also completed a Simon task to assess their executive control. Bilinguals with greater executive control named cognates nearly as fast as language-unique words during L2 picture-naming task than bilinguals with weaker executive control, who named cognates faster than language-unique words. These results argue that bilinguals with greater executive control inhibited all irrelevant lexical forms more successfully than bilinguals with weaker executive control. While these findings are important, it is difficult to compare them directly to the results of our study because in our study (1) bilinguals named pictures in both languages, and (2) we eliminated all stimuli with cross-linguistic overlap (e.g., cognates). Such task and stimuli differences could potentially account for differences in results. That aside, our results generally cohere with those observed by Linck and colleagues in that bilinguals with greater executive control in our study similarly inhibited an irrelevant language more successfully than bilinguals with weaker executive control.

Second, executive control related to specific kinds of production demands within a particular communicative context. We found that executive control related more to global demands (i.e., inhibiting activation of an entire language system as reflected by a cost for L2 versus L1 production in general) rather than local demands (i.e., inhibiting activation of particular lexical forms as reflected by multiple object labels both within- and across-languages). This result is consistent with what we have shown in naturally-occurring bilingual monologue and dialogue production (Pivneva et al., 2012). We found that greater executive control related to how efficiently bilinguals planned speech in general (e.g., acoustic-temporal measures of speech planning and production) as opposed to the specific content (e.g., semantic clarity of instruction) of what was produced.
This result is also in part consistent with a recent study by Van Assche and colleagues (Van Assche, Duyck, & Gollan, 2013) that demonstrates that bilinguals engage in a global suppression of a whole language when bilinguals are in particularly potent communicative contexts (e.g., living immersed in the L2). Specifically, Netherlands-based Dutch-English (Experiment 1) and USA-based Chinese-English (Experiment 2) bilinguals generated words that began with a specific letter/phoneme in their L1 and L2. Some letter/phoneme categories were language-specific (e.g., naming words beginning with M in one language and B in the other language), whereas other letter/phoneme categories were shared across languages (e.g., naming words with F in both languages but in separate blocks). While the results revealed that both group of participants engaged in the item-specific suppression, only the Chinese-English bilinguals living in the USA (e.g., immersed in L2) engaged in whole-language suppression. Of note, this study did not assess executive control. Nonetheless, it is evident that potent communicative contexts can aid to engage in whole-language suppression. Thus, when bilinguals in our study produced speech in highly-demanding and potent communicative contexts (e.g., such as producing grammatically well-formed sentences), they globally inhibited all irrelevant lexical information that was associated with a particular language and the need to locally inhibit specific lexical word forms was reduced.

An alternate way of thinking of global versus local demands of producing speech is to link it to production in pure versus mixed blocks. Pure blocks allow bilinguals to be immersed in a communicative context in which a full suppression of an irrelevant task schema (i.e., other language but not L1 or L2 specifically) is optimal, whereas mixed blocks provide a communicative context in which a full suppression of an irrelevant task schema is not optimal and both task schemas must remain constantly active so that successful switching can occur.
Consistent with this conjecture, we show that bilinguals with greater executive control produce speech more accurately in the pure blocks than bilinguals with weaker executive control in Experiment 1. Thus, greater executive control relates to a more successful full suppression of an irrelevant task schema in an immersed communicative context. This result is consistent with prior work by Festman and colleagues (Festman & Munte, 2012; Festman, Rodriguez-Fornells, & Munte, 2010) that demonstrates that bilinguals with greater executive control tend to have fewer unintended switches and better language control when naming pictures.

While we show that executive control is important for bilingual language production as a function of (1) particular communicative contexts, and (2) specific kinds of production demands within a particular communicative context, we also show that executive control is also important for the overall language production efficiency. Recall that bilinguals with greater executive control produced speech more efficiently overall than bilinguals with weaker executive control in Experiment 2. This result is consistent with what others have shown during native language production (Acheson et al., 2011; Acheson & MacDonald, 2009; Hartsuiker & Barkhuysen, 2006; Pivneva et al., 2012). For example, patients with lesions to brain areas associated with executive control name pictures less accurately than healthy controls (Novick et al., 2009).

The link between executive control and successful bilingual production in Experiment 1 is directly supported by the Inhibitory Control model, which assigns an early and central role to executive control in bilingual language production (Green, 1998). Recall that according to the IC model, executive control is recruited to inhibit dominant L1 schema early on, so that L2 schema can be realized and successful L2 production can occur. The amount of inhibition required by the executive control during production is reactive in nature because its strength varies as a function of contextual cues (e.g., language of the conversation). Thus, contextual
cues in Experiment 1 (i.e., producing well-formed sentences with the embedded target picture) were relatively more salient than in Experiment 2 (i.e., producing single words in isolation) for bilinguals with greater versus weaker executive control. As a result, bilinguals with greater executive control could plan and produce speech in the L2 as efficiently as in the L1 in Experiment 1 but this effect was not language-specific in Experiment 2.

The link between recruitment of executive control in bilingual language production is further corroborated by recent neuroimaging studies using similar versions of the Simon and the Stroop tasks (Banich, 2009; Liu, Banich, Jacobson, & Tanabe, 2004; Silton et al., 2010) and the ensuing model, the Cascade-of-Control Model of Executive Function (Banich, 2009). These studies show that participants recruit the dorsolateral prefrontal cortex (DLPFC), the anterior cingulate cortex (ACC), among other areas implicated in executive control while performing these tasks. The DLPFC is thought to impose and maintain top-down attentional resources for task-relevant schemas, such as those found in a Stroop task (e.g., ink-color identification in a classic version of the Stroop task) during early stages of processing, as it activates in the presence of strongly competing task-irrelevant schemas (e.g., reading the printed color word). The ACC is thought to be involved in later stages of processing, such as those typically found in a Simon task. These include response selection, execution, evaluation and conflict monitoring, as the ACC activates in the presence of stimulus-response conflict (e.g., pressing a button when the stimulus spatial location and response button mismatch in a Simon task) (Banich, 2009; Liu et al., 2004; Silton et al., 2010). With respect to bilingual language production, both the DLPFC (i.e., representing early language control) and the ACC (e.g., representing response conflict) have been hypothesized to play a critical role in bilingual language processing as proposed by the extension of the IC model (Abutalebi & Green, 2007; Green, 1998). In support of this, we found
that bilingual language production was modulated by performance on 1) a Stroop task (i.e., reflecting early top-down language control) during earlier stages of planning (i.e., gaze duration in Experiment 1; overall M = 1391); 2) and by that of a Simon task (i.e., reflecting later stages of response-conflict resolution) during later stages of planning such as immediately preceding articulation (i.e., gaze-speech latency in Experiment 1; overall M = 1575; voice onset RT in Experiment 2; overall M = 1717). Thus, while critical for overall production, executive control is particularly important for earlier stages of speech planning, especially when language-specific demands are high, such as producing full sentences versus single words.

The link between executive control and bilingual language production may also be integrated with a recent cognitive theory of executive control, the Dual Mechanisms of Cognitive Control (Braver, 2012). This theory suggests that executive control operates through two independent mechanisms. Early proactive control maintains goal-relevant information in anticipation of executively demanding events, whereas reactive control serves as a “late correction” mechanism that is triggered as high interference stimuli are encountered in real time (Braver, 2012). While we did not specifically test proactive versus reactive types of executive control, our results nonetheless suggest that bilinguals may recruit proactive, top-down attentional control that anticipates linguistic conflict by globally down-regulating activation of a particular language system (e.g., producing a sentence in an exclusive L1 or L2 context), as opposed to reactive, bottom-up attentional control that inhibits conflicting lexical forms within an integrated lexical-semantic memory store as needed.

While our data are most consistent with the Inhibitory Control model (Green, 1998) and the Dual Mechanisms of Cognitive Control model, they can also provide partial support for the language-oriented models of bilingual language production (Costa, 2005; Costa et al., 1999;
Costa & Santesteban, 2004; Costa et al., 2006; De Bot, 1992; Green, 1998; Poullisse & Bongaerts, 1994; see also BIA+ (Dijkstra & Van Heuven, 2002; Runnqvist et al., 2012). Recall that these models are generally agnostic to the role of domain-general executive control during bilingual language production. Instead, they posit that bilinguals encode and represent language-specific cues, which activate with the intention to speak in a particular language. These language cues provide feed-forward control (e.g., mental firewall in the language-specific selection framework proposed by Costa and colleagues) that biases production output to the intended language, while minimizing competition from the irrelevant language. As such, the global control can be achieved via reduced activation from the irrelevant language by heightening activation of the relevant language. Local control is achieved through lateral inhibitory links between representations at any particular level of linguistic processing, or across levels of processing within an integrated core language system. Thus, it is possible that language cues in Experiment 1 (e.g., producing fully formed sentences in L1 or L2) reduced competition from the irrelevant language globally by heightening activation of the sentence-relevant language but this effect was observed only in bilinguals with greater executive control. However, reduced saliency of language cues in Experiment 2 (e.g., producing isolated nouns without determiners) should have increased competition locally, especially in mixed L1/L2 blocks. Yet, the increased local competition, as demonstrated by the strongest negative effect of low-name agreement, occurred during the least effortful pure L1 block production. Thus, what we observed with regards to global control for bilinguals with greater executive control is consistent with language-oriented models but local control is less so.

Interestingly, language-oriented models can often include a form of language control. For example, bilinguals can develop a type of language control through the language-specific
selection mechanisms at high L2 proficiency levels, however, bilinguals would rely on executive control at low L2 proficiency levels (Costa, 2005) and only when speaking the L2 (Schwieter & Sunderman, 2008). Of note, we statistically residualized L2 proficiency out of executive control in the present study. However, we did observe a significant role of L2 proficiency on name agreement effects in Experiment 1. Recall that increased L2 proficiency (after residualizing out executive control) reduced the negative effects of low name agreement (e.g., latency costs associated with naming pictures with multiple distinct labels) during L2 production. Thus, it is possible that high L2 proficiency bilinguals achieved better language control through language-specific selection mechanisms than low L2 proficiency bilinguals in Experiment 1, especially when producing L2 speech - a result generally consistent with language-oriented models.

Because of the potential role of L2 proficiency on the negative effects of low name agreement in Experiment 1 (see footnote 1), we conducted similar analyses in Experiment 2 as in Experiment 1. While increased L2 proficiency (after residualizing out executive control) significantly reduced the overall negative effects of low name agreement, as evidenced by a significant two-way interaction between name agreement and L2 proficiency for accuracy (b = -0.45, SE = 0.20, z = -2.20), this effect was not language specific (the language*name agreement*L2 proficiency interaction was not significant; b = 0.35, SE = 0.24, z = 1.46). Moreover, neither of these effects was present for the voice onset RT (language*name agreement*L2 proficiency interaction: b = -0.004, SE = 0.03, t = -0.13; name agreement*L2 proficiency interaction: b = -0.002, SE = 0.01, t = -0.15). However, increased L2 proficiency significantly reduced the overall costs associated with L2 versus L1 production for accuracy (b = -1.13, SE = 0.34, z = -3.36) and for voice onset RTs (b = 0.13, SE = 0.03, t = 3.73). Thus, high L2 proficiency bilinguals could achieve better performance with respect to knowledge of multiple distinct labels than low L2 proficiency
bilinguals, a result consistent with the local control view of the language-oriented models. That this effect was not language specific argues that there were insufficient language cues during single noun production to engage language-specific selection mechanism for high L2 proficiency bilinguals to achieve global control (but we did observe a general benefit for L2 versus L1 production costs for high L2 proficiency bilinguals). Future research can investigate the role of L2 proficiency and executive control on varying degree of language cues during L2 production in greater detail.

To conclude, individual differences among bilinguals in executive control significantly predict bilingual language production, consistent with prior empirical findings on bilingual language production using single picture (e.g., Kroll & Gollan, 2014; Linck et al., 2008), along with the IC model of bilingual language production (Green, 1998), and with more general executive control frameworks (Banich, 2009; Braver, 2012). Moreover, these results are consistent with other work demonstrating a clear relationship between executive control and bilingual language processing, such as comprehension, and code-switching among others (Blumenfeld & Marian, 2011; Festman et al., 2010; Pivneva, Mercier, & Titone, 2014; Soveri, Rodriguez-Fornells, & Laine, 2011). Whether this relationship between bilingual language production and executive control extends to other populations (e.g., healthy bilingual older adults) that generally exhibit age-related decline in executive controls an open question that we are currently pursuing.
Preface to Chapter 4

The study presented in Chapter 3 investigated the link between executive control and bilingual language production in different communicative contexts with greater experimental control. Across two Experiments the study compared how executive control demands varied when bilinguals produced grammatically well-formed sentences in Experiment 1 and single words in isolation in Experiment 2 in single- and dual-language contexts. The study also compared how executive control demands varied within a particular communicative context, such as locally inhibiting specific lexical candidates (e.g., couch vs. sofa vs. divan) and/or globally inhibiting activation of the entire linguistic system.

The results revealed that executive control is particularly important when producing grammatically well-formed sentences and less so when producing single words. In addition, executive control is also particularly important when globally inhibiting an entire linguistic system as opposed to locally inhibiting specific lexical candidates. Lastly, executive control also reduced the overall costs associated with language-mixing but only in Experiment 1. Strengths of this study included 1) adapting a monolingual eye movement paradigm to study bilingual language production; 2) using a novel measure, name agreement, to assess within- and cross-language activation; and 3) comparing sentence to single word-level production to assess more naturalistic production but still with greater experimental control.

Taken together, studies reported in Chapters 2 and 3 advanced our knowledge in demonstrating that bilinguals recruit executive control to support language production, an essential skill in daily life. To follow up, we set out to extend this to another essential skill integral to daily life, reading. Therefore, in study reported in Chapter 4 we investigated whether individual differences in domain-general executive control can modulate bilingual sentence
reading. Studies conducted by our group and others have consistently found that bilinguals activate both languages when reading for comprehension in L1 (Titone et al., 2011; Van Assche et al., 2009) and L2 (Libben & Titone, 2009; Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker, 2011; Van Assche et al., 2012), and individual differences among bilinguals can modulate cross-linguistic activation during L1 reading (Titone et al., 2011).

Studies demonstrating cross-linguistic activation in bilinguals reading have prompted a development of one of the most influential models of bilingual language processing, the Bilingual Interactive Activation Plus (BIA+, Dijkstra & van Heuven, 2002), which is an extension of the Bilingual Interactive Activation (BIA, Dijkstra & Van Heuven, 1998) and of the monolingual Interactive Activation model of language processing (IA, McClelland & Rumelhart, 1981). Within the framework of this model, bilinguals represent L1 and L2 knowledge in an integrated common mental store (i.e., lexicon), listed as semantics in Figure 4. When bilinguals read for comprehension L1 and L2 knowledge activates automatically and non-selectively when bilinguals visually decode word forms. Activation propagates from sublexical to lexical to semantic levels via excitatory links, often creating competition between activated representations within a particular level. Competition is largely resolved passively via lateral inhibitory connections between representations within a particular level of processing. Importantly, language node activation in the linguistic system classifies language membership but it cannot feedback to other levels to inhibit activated representations. Critically, task demands also cannot directly interact with the linguistic processing system and the bottom up processes involved in word decoding to resolve ongoing competition as they only receive the output (after competition has already been resolved) from the linguistic processing system to adjust activation levels with
respect to task demands. In this regard, the BIA+ is similar to mental firewall accounts of bilingual language production.

Figure 4. Graphical representation of the Bilingual Interactive Activation Plus (BIA+) model of bilingual language processing (Dijkstra & Van Heuven, 2002, p.182).

According to the BIA+ model, situational or individual difference factors from outside of the linguistic processing system should not significantly modulate cross-linguistic activation, especially early on, as the linguistic decoding system is architecturally impermeable to the
outside influences of task schemas or demands associated with it. In contrast, the IC model of bilingual language production suggests that individual differences among bilinguals should modulate cross-linguistic activation to reduce competition throughout language processing levels as the linguistic processing system draws heavily upon the domain-general cognitive system to support language processing (Green, 1998). Interestingly, studies from our group have specifically demonstrated that factors outside of the linguistic processing system can significantly modulate attenuation of cross-language activation. For example, Libben and Titone (2009) used eye movement monitoring technology to demonstrate that while cross-language activation occurs across early and late reading stages, sentence constraint can attenuate cross-language activation during later reading stages. More recently, Titone and colleagues (2011) used the same paradigm to demonstrate that individual differences in L2 ability related to the success with which bilinguals attenuated cross-linguistic activation.

Thus, in the study reported in Chapter 4 we aimed to assess whether and how executive control relates to bilingual sentence reading using an already established eye movement sentence-reading paradigm from Libben and Titone (2009) and Titone and colleagues (2011). The particular aim of the study was to compare whether and how executive control might attenuate cross-linguistic activation across early and late reading stages. To do so, we used eye movement measures of bilingual L2 sentence reading to investigate executive control effects when bilinguals read grammatical sentences that contained interlingual homographs, cognates, or L2-unique control words. Interlingual homographs and cognates have been used in prior work to index cross-linguistic activation by comparing reading times of interlingual homographs or cognates to those of L2-unique control words matched for other important lexical variables (e.g., word length, frequency with which the word appears in print in a particular language, etc.).
Interlingual homographs are a class of words within a language that shares orthography (i.e., spelling) but diverges in semantics (i.e., meaning) across two languages (e.g., *chat* means a casual conversation in English and a cat in French). In contrast, cognates are a class of words that shares orthography and semantics across both languages (e.g., *piano* means a musical instrument across English and French). Evidence from prior studies suggests that generally interlingual homographs interfere with processing (e.g., on average slower than language-unique words) due to conflict in semantics and cognates facilitate processing (e.g., on average faster than language-unique words) due to convergence in semantics.

While interlingual homograph interference and cognate facilitation are thought to commonly reflect automatic non-selective lexical access in cross-linguistic activation, there is an active debate about whether interlingual homographs and cognates are represented similarly or differently within the bilingual mental store (Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010; Titone et al., 2011). If they are represented differently within the mental store, then executive control recruitment might vary with the demands that interlingual homographs and cognates might impose on language processing. Likely, interlingual homographs might require greater executive control recruitment to a greater extent than cognates to attenuate cross-linguistic activation because of semantic conflict.

As will be seen in Chapter 4, the results suggest that individual differences in executive control among bilinguals significantly relate to bilingual reading, similar to what we observed for language production. Interestingly, greater executive control among bilinguals significantly reduced cross-language activation in the presence of semantic conflict, as indexed by interlingual homograph interference but not as indexed by cognate facilitation. More interestingly, cross-
language attenuation occurred during early reading stages, consistent with inhibitory accounts of bilingual language production (e.g., IC model, Green, 1998).
CHAPTER 4:

Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements

(Pivneva, Mercier, & Titone, 2014, Journal of Experimental Psychology: Learning, Memory, & Cognition)
Abstract

Models of bilingual reading, such as BIA+ (Dijkstra & van Heuven, 2002), do not predict a central role for domain-general executive control during bilingual reading, in contrast with bilingual models from other domains, such as production (e.g., the Inhibitory Control Model, Green, 1998). We thus investigated whether individual differences among bilinguals in domain-general executive control modulate cross-language activation during L2 sentence reading, over and above other factors such as L2 proficiency. Fifty French-English bilinguals read L2-English sentences while their eye movements were recorded, and subsequently completed a battery of executive control and L2 proficiency tasks. High and low constraint sentences contained interlingual homographs (chat = a talk in English; a cat in French), cognates (piano in English and French), or L2-specific control words. The results showed that greater executive control among bilinguals but not L2 proficiency reduced cross-language activation in terms of interlingual homograph interference. In contrast, increased L2 proficiency but not executive control reduced cross-language activation in terms of cognate facilitation. These results suggest that models of bilingual reading must incorporate mechanisms by which domain-general executive control can alter the very earliest stages of bilingual lexical activation.
Executive Control Modulates Cross-Language Lexical Activation During L2 Reading: Evidence From Eye Movements

When bilinguals read in their second language (L2), they access orthography, phonology, and semantics from their first language (L1) (reviewed in Van Assche et al., 2012). This also occurs when highly proficient bilinguals read in their L1 (Titone et al., 2011; Van Assche et al., 2009). Bilingual reading is, thus, similar to spoken comprehension (reviewed in Shook & Marian, 2013) and language production (reviewed in Kroll & Gollan, 2014) in revealing how some form of language control is required to resolve cross-language conflict (e.g., Bialystok et al., 2012; De Groot & Christoffels, 2006; Green, 1998). This coheres with models positing an early, central role of executive control during language processing, such as the Inhibitory Control (IC) model (e.g., Abutalebi & Green, 2007; Green, 1998), originally proposed for language production. Whether this extends to bilingual reading is an open question.

The most prominent bilingual reading model is bilingual interactive activation plus (BIA+; Dijkstra & van Heuven, 2002). Accordingly, cross-language competition is resolved in two ways. Within the language system, L1 and L2 representations co-exist within an integrated lexical store, and feed activation forward to nodes coding for language membership (these cannot feed-back to lower levels). Here, lateral inhibitory connections within a level (which would presumably be affected by resting activation), or semantic to orthographic feedback across levels (which would presumably be an interaction of resting activation and contextual constraint) implement passive forms of language control (Bialystok et al., 2012; De Groot, 2011). In contrast, an active form of language control occurs outside the language system through domain-general task schemas that take its outputs as input, and adjust responding based on these continuous activations, the current task, and any relevant decision criteria. Importantly, activity
occurring at the task-schema level is architecturally blocked from the internal dynamics of the language system, and thus, cannot affect bottom-up processes occurring therein.

The BIA+ architecture has implications for the time course of language control during bilingual reading (van Heuven & Dijkstra, 2010). First, passive control processes operative inside the language system (within and across levels) should occur rapidly because they modulate lexical activation directly. Second, active control processes operative outside the language system should occur more slowly, if at all, because they only modulate its outputs rather than ongoing activation. Thus, BIA+ predicts that variables relating to control operations within the language system (e.g., semantic constraint) should directly affect lexical activation. In contrast, variables relating to control operations outside language system (e.g., individual differences in the efficiency of domain-general executive control) cannot affect initial lexical activation directly (Van Heuven & Dijkstra, 2010).

We investigate these predictions using eye movement measures of sentence reading, focusing on two word types – interlingual homographs, which share orthography but conflict semantically across languages (chat – casual conversation in English, cat in French), and cognates, which share both orthography and meaning across languages (jungle). Interlingual homographs usually produce interference (the net result of form overlap and meaning non-overlap), whereas cognates generally produce reading facilitation (the net result of form and meaning overlap) (reviewed in Van Assche et al., 2012). We use eye movement measures to estimate a detailed time-course of processing for these words. First pass measures such as gaze duration (GD: the sum of all first pass fixations on a word) index early comprehension stages (initial lexical activation); later reading measures such as total reading time (TRT: the sum of all
fixations, both first and subsequent passes) index integrative stages subsequent to initial activation (Rayner & Schotter, 2013).

We previously found that highly proficient French-English bilinguals reading in their L2 (English) showed classic homograph interference and cognate facilitation effects for GD, irrespective of whether an L2 (English) carrier sentence semantically biased a target word (Libben & Titone, 2009), consistent with single-word decision studies showing non-selectivity (reviewed in Van Assche et al., 2012). However, high semantic constraint attenuated homograph interference and cognate facilitation for TRT. Within BIA+, these effects would have arisen from lexical activation and lateral inhibition occurring at the orthographic level, and feedback from the semantic to the orthographic level within the language system.

We also investigated L1 bilingual reading and issues related to language control by manipulating task demands in highly proficient English-French bilinguals (Titone et al., 2011). Specifically, we interspersed L2 (French) non-target-language filler sentences on 1/3 of trials, which resulted in enhanced cross-language activation during L1 (English) target-language trials, particularly for GD. Thus, task demands modulated initial lexical activation, though it remains unclear whether fillers modulated L2 (French) resting activation of phonology/orthography within the language system, consistent with BIA+, or task expectations occurring outside the language system, inconsistent with BIA+. Thus, an open question is whether domain-general cognitive processes modulate early lexical activation, in contrast with predictions following from BIA+.

We thus investigated this issue, and were particularly interested in whether executive control helps to reduce cross-language competition, or plays other roles with respect to speed and fluency of language processing (for the role of executive control in general sentence processing
in monolinguals see Loncke, Desmet, Vandierendonck, & Hartsuiker, 2011). BIA+ has no clear mechanism by which executive control affects early reading stages, as such effects reflect processing efficiency outside rather than inside the language system. However, other models heretofore not specific to reading (e.g., Abutalebi & Green, 2007; Green, 1998), imply that differences among bilinguals in domain-general executive control should penetrate the very earliest stages of reading, as they are fundamental to bilingual processing generally.

Thus, if domain-general executive control operates outside the language system, there should be no relation between executive control and cross-language competition during early reading stages, as reflected by measures such as GD (which unlike first fixation duration, reflects complete first pass word processing). However, if domain-general executive control is central to bilingual language processing (Abutalebi & Green, 2007; Green, 1998), increased executive control should pattern with less cross-language conflict during early reading. Executive control should also be particularly important for interlingual homographs, which require active cross-language suppression (as hypothesized by Macizo, Bajo, & Martin, 2010), versus cognates, which do not. Executive control should also be particularly important when there is limited semantic to orthographic feedback to help resolve cross-language conflict, as would be true for low constraint sentences.

**Method**

**Participants**

Fifty French-English bilinguals (36 females; $M_{age} = 24.02$, $SD_{age} = 3.63$), with normal/corrected vision, and no self-reported history of speech/hearing disorders, participated for course credit or monetary compensation ($10/hour).

Participants completed a language questionnaire, and an objective language proficiency
task where they decided whether a visually presented noun was animate or inanimate in separate counterbalanced L1- and L2-only blocks (Segalowitz & Frenkiel-Fishman, 2005). From this, we divided L2 by L1 mean correct reaction times (RT) to obtain a continuous, objective measure of L2 versus L1 proficiency, L2/L1 ratio (see also Van Assche et al., 2012) (Table 1).
Table 1

*Self-report English proficiency and objective L2 ability (L2/L1 RT ratio)*

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rating Scales (0-10)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking Ability</td>
<td>4</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Reading Ability</td>
<td>4</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Writing Ability</td>
<td>3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Translating Ability</td>
<td>3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>4</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Pronunciation</td>
<td>3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Fluency</td>
<td>3</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>Grammatical Ability</td>
<td>3</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Overall Competence</td>
<td>4</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td><strong>Sum of Rating Scales (0-100)</strong></td>
<td>37</td>
<td>100</td>
<td>70</td>
</tr>
</tbody>
</table>

**Age of L2 acquisition (yrs. old)**

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Began Acquiring English</td>
<td>0</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Became competent in English</td>
<td>5</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>Began reading in English</td>
<td>3</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>Became competent in reading</td>
<td>5</td>
<td>22</td>
<td>16</td>
</tr>
</tbody>
</table>

**Degree of French interference when**

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Reading</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
</tbody>
</table>

**Percent of present time spent functioning**

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>2</td>
<td>83</td>
<td>27</td>
</tr>
<tr>
<td>French</td>
<td>16</td>
<td>98</td>
<td>70</td>
</tr>
</tbody>
</table>

**L2/L1 RT ratio**

<table>
<thead>
<tr>
<th></th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.86</td>
<td>1.24</td>
<td>1.01</td>
</tr>
</tbody>
</table>
Materials

Materials, taken from Libben and Titone (2009), consisted of 31 form-identical cognates (jungle), 27 interlingual homographs (chat – French for cat), and English-only control words matched to the cognates and homographs on word length, frequency, and neighborhood density. Words occurred sentence-medially, and high/low semantic constraint was manipulated and verified by semantic bias and cloze probability ratings from another group of English-L1 participants (Table 2).
Table 2

*Sample stimuli used in the sentence reading task (Libben & Titone, 2009) and their corresponding mean semantic bias and mean cloze probability ratings*

<table>
<thead>
<tr>
<th>Word type</th>
<th>High-constraint</th>
<th>Semantic bias (0-10)</th>
<th>Cloze probability (0-1)</th>
<th>Low-constraint</th>
<th>Semantic bias (0-10)</th>
<th>Cloze probability (0-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interlingual homographs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target word</td>
<td>Because she knew the change was counterfeit, the brown colored coin was thrown out.</td>
<td>8.34</td>
<td>0.4</td>
<td>Because it was completely worthless, the brown colored coin was thrown out.</td>
<td>2.90</td>
<td>0.04</td>
</tr>
<tr>
<td>Matched control</td>
<td>Because it didn't clean and lather well, the brown colored soap was thrown out.</td>
<td>8.53</td>
<td>0.37</td>
<td>Because it smelled really bad, the brown colored soap was thrown out.</td>
<td>2.87</td>
<td>0.02</td>
</tr>
<tr>
<td>Cognates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target word</td>
<td>When they were on the safari, they saw an enormous jungle that was dark and scary.</td>
<td>8.86</td>
<td>0.48</td>
<td>When they were on their trip, they saw an enormous jungle that was dark and scary.</td>
<td>3.05</td>
<td>0.04</td>
</tr>
<tr>
<td>Matched control</td>
<td>When she was chewing her gum, she blew an enormous bubble that was pink and shiny.</td>
<td>8.85</td>
<td>0.49</td>
<td>When she waited for her friend, she blew an enormous bubble that was pink and shiny.</td>
<td>2.82</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Procedure

Eye movements were recorded with an Eye-Link 1000 (SR-Research, Ontario, Canada) from the right eye (binocular viewing). Sentences were presented on a 20-inch CRT (71 cm away). Comprehension questions occurred on 25% of trials. Following the reading task, participants completed an executive control battery, which included anti-saccade (Hallett, 1978), non-linguistic Simon and Stroop (Blumenfeld & Marian, 2011), and Number Stroop tasks (Table 3; for detailed descriptions, see Mercier et al., 2013; Pivneva et al., 2012). The anti-saccade task assessed the ability to inhibit eye movements away from peripheral targets. The non-linguistic Simon task assessed the ability to inhibit a mismatch in stimulus spatial location and response button. The non-linguistic Stroop and number Stroop tasks assessed the ability to inhibit a mismatch in semantic meaning assigned to a stimulus and the response button. To estimate executive control capacity common to all tasks (the unity of executive control, Miyake & Friedman, 2012), we created a composite score for each participant (McMurray et al., 2010). All sub-measures correlated significantly with the composite score (Table 4), though Simon task performance had a greater influence than others. Participants then completed the L2 questionnaire and objective proficiency test.
Table 3
Mean correct reaction time values, and standard deviations for executive control measures and the composite executive control cost score

<table>
<thead>
<tr>
<th>Task</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Simon task</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent</td>
<td>482</td>
<td>98</td>
</tr>
<tr>
<td>Incongruent</td>
<td>531</td>
<td>94</td>
</tr>
<tr>
<td>Cost</td>
<td>-49***</td>
<td>4</td>
</tr>
<tr>
<td><strong>Stroop task</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent</td>
<td>489</td>
<td>125</td>
</tr>
<tr>
<td>Incongruent</td>
<td>499</td>
<td>103</td>
</tr>
<tr>
<td>Cost</td>
<td>-10</td>
<td>23</td>
</tr>
<tr>
<td><strong>Number Stroop task</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congruent</td>
<td>617</td>
<td>94</td>
</tr>
<tr>
<td>Incongruent</td>
<td>676</td>
<td>102</td>
</tr>
<tr>
<td>Cost</td>
<td>-59***</td>
<td>-8</td>
</tr>
<tr>
<td><strong>Anti-saccade task</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-saccadic (Pure)</td>
<td>287</td>
<td>37</td>
</tr>
<tr>
<td>Anti-saccadic (Pure)</td>
<td>379</td>
<td>44</td>
</tr>
<tr>
<td>Pro-saccadic (Mixed)</td>
<td>282</td>
<td>41</td>
</tr>
<tr>
<td>Anti-saccadic (Mixed)</td>
<td>380</td>
<td>44</td>
</tr>
<tr>
<td><strong>Total Anti-saccade Cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro-saccadic (Pure + Mixed)</td>
<td>285</td>
<td>36</td>
</tr>
<tr>
<td>Anti-saccadic (Pure + Mixed)</td>
<td>380</td>
<td>40</td>
</tr>
<tr>
<td>Cost</td>
<td>-95***</td>
<td>-4</td>
</tr>
<tr>
<td><strong>Task Switch Cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pure (Pro-saccadic + Anti-saccadic)</td>
<td>333</td>
<td>33</td>
</tr>
<tr>
<td>Mixed (Pro-saccadic + Anti-saccadic)</td>
<td>331</td>
<td>33</td>
</tr>
<tr>
<td>Cost</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td><strong>Executive Control Cost Score</strong></td>
<td>-0.0002</td>
<td>0.0051</td>
</tr>
</tbody>
</table>

*** Two-tailed paired samples t-test significant at p < 0.001
Table 4

Correlations between individual executive control measures, L2 proficiency. WASI and the composite executive control measure

<table>
<thead>
<tr>
<th></th>
<th>Executive control (composite score)</th>
<th>Simon cost</th>
<th>Stroop cost</th>
<th>Number Stroop cost</th>
<th>Antisaccade cost</th>
<th>Antisaccade task switch cost</th>
<th>L2 proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive control measures</td>
<td>0.54 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simon cost</td>
<td>0.54 **</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroop cost</td>
<td>0.28 *</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number Stroop cost</td>
<td>0.12</td>
<td>0.11</td>
<td>-0.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antisaccade cost</td>
<td>0.46 **</td>
<td>-0.11</td>
<td>0.25</td>
<td>-0.18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antisaccade task switch cost</td>
<td>0.31 *</td>
<td>0.25</td>
<td>-0.28</td>
<td>-0.19</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 proficiency</td>
<td>-0.04</td>
<td>-0.12</td>
<td>0.09</td>
<td>-0.21</td>
<td>-0.01</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>WASI</td>
<td>-0.17</td>
<td>-0.20</td>
<td>-0.07</td>
<td>0.02</td>
<td>-0.32 *</td>
<td>-0.17</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

*p < .05, **p < .01

1 L2 ability is not residualized out of this variable
Results

We removed observations < 100 ms (2.43% of trials), and where targets received no fixations (8.53% of trials). The first dependent variable (DV) of interest was GD (gaze duration), the sum of all fixation durations starting when the eye first lands on a word until it moves away, which reflects early lexical access as it is influenced by variables such as frequency (Rayner & Schotter, 2013). The second DV was TRT (total reading time), the sum of all fixation durations on a word, which reflects later stages of comprehension as it includes second-pass fixations and is influenced by sentence- or discourse-level factors (Rayner & Schotter, 2013). Table 5 presents GD and TRT for homographs, and Table 6 presents the same for cognates. Table 7 presents correlations between composite executive control scores, individual executive control scores, L2 proficiency, WASI and the homograph interference effect for GD and TRT.
Table 5
Mean values (ms) and standard errors of the mean for first pass gaze duration (GD), and total reading time (TRT) for interlingual homographs and matched control words overall for all bilinguals and split by quartiles on executive control cost score. Of note, all analyses reported in the paper used continuous measures of executive control, and are split into quartiles here to facilitate inspection of the means.

<table>
<thead>
<tr>
<th>Executive Control (EC)</th>
<th>All participants (n=50)</th>
<th>Greatest EC</th>
<th>Greater EC</th>
<th>Weaker EC</th>
<th>Weakest EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homograph</td>
<td>293</td>
<td>19</td>
<td>274</td>
<td>16</td>
<td>286</td>
</tr>
<tr>
<td>Control</td>
<td>301</td>
<td>18</td>
<td>296</td>
<td>18</td>
<td>300</td>
</tr>
<tr>
<td>Difference</td>
<td>-8</td>
<td>-22</td>
<td>-14</td>
<td>-1</td>
<td>8</td>
</tr>
<tr>
<td>High Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homograph</td>
<td>286</td>
<td>16</td>
<td>270</td>
<td>15</td>
<td>283</td>
</tr>
<tr>
<td>Control</td>
<td>288</td>
<td>19</td>
<td>271</td>
<td>19</td>
<td>278</td>
</tr>
<tr>
<td>Difference</td>
<td>-2</td>
<td>-1</td>
<td>5</td>
<td>-8</td>
<td>-7</td>
</tr>
<tr>
<td>Total Reading Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Constraint</td>
<td>479</td>
<td>45</td>
<td>443</td>
<td>43</td>
<td>402</td>
</tr>
<tr>
<td>Homograph</td>
<td>442</td>
<td>43</td>
<td>457</td>
<td>45</td>
<td>413</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>37</td>
<td>-14</td>
<td>-11</td>
<td>106</td>
<td>65</td>
</tr>
<tr>
<td>High Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homograph</td>
<td>428</td>
<td>41</td>
<td>443</td>
<td>42</td>
<td>382</td>
</tr>
<tr>
<td>Control</td>
<td>406</td>
<td>36</td>
<td>367</td>
<td>30</td>
<td>357</td>
</tr>
<tr>
<td>Difference</td>
<td>22</td>
<td>76</td>
<td>25</td>
<td>-14</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 6

Mean values (ms) and standard errors of the mean for first pass gaze duration (GD), and total reading time (TRT) for cognates and matched control words overall for all bilinguals and split by quartiles on executive control cost score. Of note, all analyses reported in the paper used continuous measures of executive control, and are split into quartiles here to facilitate inspection of the means.

<table>
<thead>
<tr>
<th></th>
<th>All participants (n=50)</th>
<th>Greatest EC</th>
<th>Greater EC</th>
<th>Weaker EC</th>
<th>Weakest EC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaze Duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognate</td>
<td>M  287  16 SE</td>
<td>M  261  16 SE</td>
<td>M  298  18 SE</td>
<td>M  290  13 SE</td>
<td>M  299  16 SE</td>
</tr>
<tr>
<td>Control</td>
<td>307  20</td>
<td>292  16</td>
<td>297  20</td>
<td>316  17</td>
<td>323  24</td>
</tr>
<tr>
<td>Difference</td>
<td>-20</td>
<td>-31</td>
<td>1</td>
<td>-26</td>
<td>-24</td>
</tr>
<tr>
<td>High Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognate</td>
<td>276  15</td>
<td>260  13</td>
<td>276  18</td>
<td>280  16</td>
<td>288  14</td>
</tr>
<tr>
<td>Control</td>
<td>291  18</td>
<td>279  14</td>
<td>264  12</td>
<td>300  18</td>
<td>316  23</td>
</tr>
<tr>
<td>Difference</td>
<td>-15</td>
<td>-19</td>
<td>12</td>
<td>-20</td>
<td>-28</td>
</tr>
<tr>
<td>Total Reading Time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognate</td>
<td>404  34</td>
<td>402  35</td>
<td>367  26</td>
<td>413  30</td>
<td>430  41</td>
</tr>
<tr>
<td>Control</td>
<td>452  38</td>
<td>458  37</td>
<td>391  33</td>
<td>487  35</td>
<td>467  45</td>
</tr>
<tr>
<td>Difference</td>
<td>-48</td>
<td>-56</td>
<td>-24</td>
<td>-74</td>
<td>-37</td>
</tr>
<tr>
<td>High Constraint</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognate</td>
<td>367  30</td>
<td>368  32</td>
<td>339  26</td>
<td>367  27</td>
<td>394  35</td>
</tr>
<tr>
<td>Control</td>
<td>404  39</td>
<td>390  31</td>
<td>329  25</td>
<td>414  38</td>
<td>478  52</td>
</tr>
<tr>
<td>Difference</td>
<td>-37</td>
<td>-22</td>
<td>10</td>
<td>-47</td>
<td>-84</td>
</tr>
</tbody>
</table>
Table 7

Correlations between individual executive control measures, L2 proficiency. WASI, the composite executive control measure and the homograph interference effect as measured by Gaze Duration (GD) and Total Reading Time (TRT)

<table>
<thead>
<tr>
<th>Executive control measures</th>
<th>GD High-constraint</th>
<th>GD Low-constraint</th>
<th>TRT High-constraint</th>
<th>TRT Low-constraint</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composite measure (L2 ability residualized out)</td>
<td>-0.03</td>
<td>0.32 *</td>
<td>-0.22</td>
<td>0.34 *</td>
</tr>
<tr>
<td>Composite measure (L2 ability not residualized out)</td>
<td>-0.11</td>
<td>0.29 *</td>
<td>-0.22</td>
<td>0.31 *</td>
</tr>
<tr>
<td>Simon cost</td>
<td>0.05</td>
<td>0.29 *</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Stroop cost</td>
<td>-0.09</td>
<td>-0.01</td>
<td>0.24</td>
<td>-0.14</td>
</tr>
<tr>
<td>Number Stroop cost</td>
<td>-0.05</td>
<td>0.26</td>
<td>-0.27</td>
<td>0.26</td>
</tr>
<tr>
<td>Antisaccade cost</td>
<td>-0.11</td>
<td>0.12</td>
<td>0.03</td>
<td>0.14</td>
</tr>
<tr>
<td>Antisaccade task switch cost</td>
<td>-0.12</td>
<td>0.13</td>
<td>-0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>L2 proficiency</td>
<td>0.27</td>
<td>-0.12</td>
<td>0.20</td>
<td>-0.24</td>
</tr>
<tr>
<td>WASI</td>
<td>-0.07</td>
<td>-0.20</td>
<td>-0.14</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*p < .05

*Correlations are significant at the .05 level (2-tailed).
**Linear mixed-effects regression**

We computed linear mixed-effects (LME) models for homographs and cognates separately using the lme4 library, version 0.999375-32 (Bates, 2005) in R (Baayen, 2008). Fixed effects included word type (Control versus Target), sentence constraint (Low versus High), and executive control cost after residualizing out L2 ability (continuous; centered and scaled), and their three-way interaction. Wechsler Abbreviated Scale of Intelligence (WASI) L1 vocabulary served as a control variable that assessed verbal intelligence or L1 proficiency (Wechsler, 1999). Executive control and L2 proficiency did not significantly correlate ($r = -0.04, p > .05$), nor did executive control and WASI ($r = -0.17, p > .05$). We included random intercepts and correlated random slopes for word type, sentence constraint, and their interaction for participants and items (Barr et al., 2013). Collinearity was < 0.26 for all variables. DVs were natural log-transformed. We used a convention of $t > 1.96$ to report significant effects, and Chi square to assess whether models containing three-way interactions explained more variance than models containing only two-way interactions. For omnibus models, we deviation coded (-0.5, 0.5) categorical variables (Barr et al., 2013). To identify source of interactions in sub-models and plot partial effects, we treatment coded (0, 1; baseline = control words) categorical variables.

**Analysis I: Interlingual Homographs**

If domain-general executive control modulates initial lexical activation, increased executive control should relate to less homograph interference. Moreover, this effect should be stronger for low-constraint sentences, which do not benefit from semantic-to-orthographic feedback to the same extent as high-constraint sentences.

---

3 The overall pattern of results was the same using F1 ANOVA for both GD and TRT after median splitting the group into bilinguals who had high vs. low executive control costs.
**Gaze Duration (GD).** The three-way interaction between word type, sentence constraint, and executive control cost was significant \((b = -0.06, SE = 0.03, t = -2.18)\) (Figure 1; upper panels), and this model explained significantly more variance than a simpler model including only two-way interactions \((\chi^2(1) = 4.74, p < .05)\). As executive control decreased, homographs became slower than language-unique words, for low-constraint sentences only. In follow-up analyses decomposing this interaction, word type by executive control interacted in low- \((b = 0.04, SE = 0.02, t = 2.08)\) but not high- constraint \((b = -0.02, SE = 0.02, t = -0.91)\) sentences. There was only a significant main effect of executive control for high-constraint sentences \((b = 0.05, SE = 0.02, t = 2.10)\), where bilinguals with weaker vs. greater executive control exhibited longer GD.
Figure 1. Partial effect of standardized executive control (L2 proficiency residualized out) on homograph interference in low- and high-constraint sentences for gaze duration (upper panels) and total reading time (lower panels).
**Total Reading Time (TRT).** Like GD, there was a three-way interaction between word type, sentence constraint, and executive control ($b = -0.01, SE = 0.04, t = -2.84$) (Figure 1; lower panels), and this model explained significantly more variance than a simpler model including only two-way interactions ($\chi^2(1) = 11.03, p < .001$). Again, as executive control decreased, homographs became slower than language-unique words, for low-constraint sentences only. This was confirmed in analyses of low constraint sentences alone by an interaction between word type and executive control ($b = 0.06, SE = 0.03, t = 2.05$), showing that homograph TRT increased as executive control cost increased. However, for high constraint sentences alone, the interaction between word type and executive control showed an opposite pattern ($b = 0.06, SE = 0.03, t = -2.23$). TRT for control words increased relative to homographs as executive control increased. To further examine this pattern, we computed models splitting the original three-way by word type rather than sentence constraint. Sentence constraint (low- vs. high; baseline = low) interacted with executive control for control words ($b = 0.07, SE = 0.03, t = 2.53$), but not homographs ($b = -0.05, SE = 0.03, t = -1.77$).

**Summary of Interlingual Homograph Results.** Consistent with the IC model, and contrary to BIA+, greater executive control (after residualizing out L2 ability, though the results were identical if we did not do so) related to less homograph interference, particularly in low-constraint sentences during early (i.e., indexed by GD) and late (i.e., indexed by TRT) reading stages. However, bilinguals with greater executive control also seemed less likely to use high contextual constraint to dampen homograph interference during late reading stages (e.g., TRT), despite the fact that they used L2 contextual constraint for control words more effectively than bilinguals with
weaker executive control (for other work suggesting reduced use of context during L2, see Martin, Thierry, Kuipers, Boutonet, Foucart, & Costa, in press). Thus, we speculate that bilinguals with greater executive control read the control words in high constraint sentences during later stages at a speed that could not be sustained for homographs, particularly given that cross-language semantic mismatches could have been detected post-lexically.

**Analysis II: Cognates**

Here, we investigate whether domain-general executive control modulates early lexical activation for cognates. If cognate facilitation arises from cross-language competition, we would expect a similar pattern as that seen for homographs. However, if facilitation arises because cognates are functionally more frequent than language-unique control words (Titone et al., 2011), we would expect no effect of executive control.

**Gaze Duration.** The three-way interaction between word type, sentence constraint, and executive control was not significant, nor were any two-way interactions ($b = -0.03$, $SE = 0.04$, $t = -1.36$). However, a word type main effect indicated that bilinguals read cognates faster than language-unique words ($M = 280$ versus 294 ms; $b = -0.03$, $SE = 0.02$, $t = -1.98$). Bilinguals also read all words faster in high- versus low-constraint sentences ($M = 281$ versus 293 ms; $b = -0.05$, $SE = 0.01$, $t = -3.77$), and all words more slowly as executive control decreased ($b = 0.04$, $SE = 0.02$, $t = 2.18$).
**Total Reading Time.** Like GD, the three-way interaction between word type, sentence constraint and executive control for TRT was not significant, nor were any two-way interactions ($b = -0.03, SE = 0.03, t = -0.99$). However, there was a word type main effect indicated that bilinguals read cognates faster than language-unique words ($M = 386$ versus $428$ ms; $b = -0.08, SE = 0.03, t = -2.91$). Bilinguals also read all words faster in high- versus low-constraint sentences ($M = 386$ versus $428$ ms; $b = -0.11, SE = 0.02, t = -5.28$). Unlike GD, executive control did not significantly modulate TRT ($b = 0.05, SE = 0.03, t = 1.35$).

**Summary of Cognate Results.** Bilinguals showed cognate facilitation for both GD and TRT, however, this did not interact with any other factor. Thus, greater executive control did not modulate cognate facilitation. Moreover, in contrast with Libben and Titone (2009), context did not modulate cognate facilitation for TRT, which may be because the present sample was overall less L2 proficient (see also Van Assche et al., 2011), a point we discuss below.

**Analysis III: L2 Proficiency Effects**

Here, we isolate effects of L2 proficiency after residualizing out executive control. We do so for two reasons. First, similar effects might occur for any variable indexing processing difficulty, and thus not be specific to executive control. Second, if homograph and cognate effects arise for different reasons (respectively, cross-language conflict vs. frequency), only cognate facilitation should relate to L2 proficiency (Van Hell & Dijkstra, 2002).
For homographs, there was a main effect of L2 proficiency in that GD and TRT increased overall as L2 proficiency increased (respectively, \( b = 0.07, SE = 0.02, t = 4.45 \); \( b = 0.08, SE = 0.01, t = 9.90 \)). L2 proficiency did not interact with any other variable. For cognates, there was a three-way interaction between word type, sentence constraint and L2 proficiency for GD (\( b = 0.06, SE = 0.03, t = 2.50 \)), and this model explained more variance than one that included two-way interactions only (\( \chi^2(1) = 5.49, p < .05 \)). As Figure 2 (upper panels) show, bilinguals exhibited more cognate facilitation as L2 proficiency decreased for low constraint sentences only. Follow-up analyses showed an interaction between word type and L2 proficiency for low constraint sentences (\( b = -0.07, SE = 0.02, t = -4.33 \)), but not high constraint sentences (\( b = -0.00, SE = 0.02, t = -0.28 \)).
Figure 2. Partial effect of standardized L2 proficiency (L2/L1 RT ratio; executive control residualized out) on cognate facilitation in low- and high-constraint sentences for gaze duration (upper panels) and total reading time (lower panels).
Like GD, TRT showed a three-way interaction between word type, sentence constraint and L2 proficiency \((b = 0.08, SE = 0.03, t = 3.01)\), and this model explained more variance than a model that included two-way interactions only \((\chi^2(1) = 5.65, p < .05)\). As Figure 2 also shows (lower panes), bilinguals exhibited more cognate facilitation as L2 proficiency decreased, for low constraint sentences only. Follow-up analyses again showed an interaction between word type and L2 proficiency for low constraint \((b = -0.08, SE = 0.02, t = -4.17)\), but not high constraint sentences \((b = 0.01, SE = 0.02, t = 0.31)\).

Thus, L2 reading was faster as L2 proficiency increased. However, the magnitude of homograph interference did not change. In contrast, similar to van Hell and Dijkstra (2002), cognate facilitation was greater as L2 proficiency decreased, and this interaction was reduced in high constraint contexts, consistent with Libben and Titone (2009).

**Discussion**

We investigated whether individual differences in domain-general executive control related to cross-language activation during bilingual reading. Greater executive control among bilinguals led to less cross-language activation during early reading stages, in terms of interlingual homograph interference for low constraint sentences. In contrast, greater executive control did not modulate cognate facilitation, although reduced L2 proficiency led to greater cognate facilitation across all measures. Like prior work (Libben & Titone, 2009), cognate facilitation was driven by language-unique control words rather than cognates themselves. This suggests that cognate facilitation
partially arises because control words have a lower functional frequency than cognates, which overlap orthographically, semantically and often phonologically across languages. Consequently, bilinguals may be more sensitive to the lower functional frequency of language-unique words because of divided L1/L2 exposure (Gollan et al., 2008; Gollan et al., 2011; Titone et al., 2011; Whitford & Titone, 2012).

Greater executive control among bilinguals also led to enhanced use of contextual constraint for control words during later L2 reading stages (TRT), consistent with work on monolinguals (e.g., Chiappe, Siegel, & Hasher, 2000; Prat, 2011). Given that L2 vs. L1 processing is generally less practiced and fluent, bilinguals devote more resources to low-level word decoding and may have fewer left over to perform other language operations, such as integrating words in an unfolding context (reviewed in Segalowitz, 2010). Future work can assess whether executive control affects low-level word decoding during single-word processing.

Compared to Libben and Titone (2009), there were similarities and differences with the present study, which we attribute to the overall lower L2 ability in the present sample. First, Libben and Titone showed contextual facilitation for control words, and less homograph interference in high vs. low constraint sentences for TRT (here, we only show this for high executive control bilinguals in GD; Libben and Titone did not assess executive control). However, the bilinguals tested here reported significantly less daily usage of English than those tested in Libben and Titone (27 versus 50%, $p < .01$) and more of French (70 versus 47%, $p < .01$). Consequently, they may have been less likely as a group to use context efficiently (e.g., Martin et
al., in press), and less likely to use context to modulate cross-language activation, similar to prior work using a relatively less L2 proficient sample (van Assche et al., 2011). Moreover, compared to bilinguals tested in Libben and Titone, the current sample may have weighted the benefits arising from form-overlap more heavily than the costs arising from meaning overlap, thus showing more initial (i.e., for GD) facilitation for homographs (e.g., Dijkstra, Grainger, & Van Heuven, 1999), which decreased over time (i.e., for TRT).

However, the GD results, which more reflect early lexical activation, show that increased executive control and contextual constraint both attenuated interlingual homograph interference (see also Libben & Titone, 2009). This is most consistent with models ascribing a central, early role to executive control (Green, 1998), but do not follow from models positing that domain-general language control is architecturally blocked from core language operations (e.g., BIA+).

Presumably, cross-language semantic conflict makes homographs more likely than cognates to require domain-general executive control (Macizo et al., 2010), except perhaps when cognate status increases the saliency of contextually irrelevant meanings (Fontes & Schwartz, 2011) or phonological mismatches occur in tasks emphasizing phonology (e.g., speaking; Linck et al., 2008).

Indeed, given prior work showing executive control influences on cognates during production (Linck et al., 2008), and phonological overlap effects on cognates during L2 reading (Dijkstra et al., 2010; Schwartz, Kroll, & Diaz, 2007), we computed models assessing the impact of cross-language phonological conflict. Homographs showed no effects, which either means that semantics trumps phonology, or that the homographs used here were relatively short, and failed to
vary sufficiently in phonological overlap. However, bilinguals with high executive control read
cognates in low constraint sentences faster when overlap was high (piano but not jungle, in English
and French; word type * executive control * overlap interaction, $b = 0.03, SE = 0.02, t = 2.05$).
Sub-models verified that bilinguals with high executive control showed facilitation only when
cognates were phonologically similar across English and French, whereas bilinguals with low
executive control showed facilitation irrespective of overlap. Thus, cross-language phonological
overlap may interact with executive control, and its impact may be even greater for tasks
encouraging greater phonological processing than ours.

To conclude, individual differences among bilinguals in domain-general executive control
modulate the initial stages of cross-language activation during L2 reading, similar to production
(Festman et al., 2010; Linck et al., 2008; Pivneva et al., 2012; Pivneva & Titone, under review),
and spoken comprehension (Blumenfeld & Marian, 2011; Mercier et al., 2013). Thus, bilingual
reading models (i.e., BIA+) ought to better specify how domain-general executive control directly
affects operations occurring within the language system. One possibility is to enable two sources
of feedback not currently present in BIA+: one from the task/decision level to the language nodes
(which already exist within the language core), and one from the language nodes to lower levels.
Such modifications might optimally integrate this important model of bilingual reading with the IC
model, which addresses bilingual production, and better reflect the neural underpinnings of
bilingual language control (e.g., Abutalebi & Green, 2008; Van Heuven & Dijkstra, 2010).
CHAPTER 5:

Summary and discussion
The goal of this thesis was to investigate whether and how executive control relates to bilingual language production and reading across different communicative contexts. As detailed in the introduction, many studies investigating bilingual language processing, across virtually all language domains, have demonstrated that bilinguals activate both of their languages automatically and non-selectively (reviewed in Kroll & Gollan, 2014), when speaking or reading in a less-dominant L2. Thus, non-selective activation presumably promotes cross-linguistic interactions between L1 and L2 (Dussias, 2003; Gollan et al., 2008; Jared & Kroll, 2001; Kroll, Dussias, et al., 2012; Kroll & Gollan, 2014; Whitford & Titone, 2012), which involve cortical and subcortical brain areas involved in both language and general cognition, particularly executive control (e.g., Abutalebi, 2008; Abutalebi et al., 2012; Abutalebi & Green, 2007, 2008). Of relevance here, although multiple known languages activate and interact with each other within the bilingual brain, bilinguals have the ability to restrict their production to one language when conversing with monolinguals, switch between languages when conversing with bilinguals, or read for comprehension in both languages. It is thought that bilinguals achieve this linguistic flexibility through the application of control mechanisms (e.g., Abutalebi & Green, 2007; Green, 1998; reviewed in Kroll & Gollan, 2014).

Also, as detailed in the introduction of this thesis, theoretical accounts vary with respect to how language control mechanisms are implemented. Mental firewall accounts argue that language control mechanism, namely the language-specific selection mechanism, resides within the linguistic system with minimal to no influence from the domain-general cognitive processing system (e.g.,
Costa, 2005; Costa & Caramazza, 1999; Costa et al., 1999). Inhibitory accounts argue that language control is recruited from the domain-general cognitive processing system, while the extent to which recruitment occurs varies with the communicative context and bilingual behavioral ecology (e.g., Abutalebi & Green, 2007; Green, 1998; Green & Abutalebi, 2013; Kroll & Bialystok, 2013; Kroll & Gollan, 2014). The idea here is that language control is part and parcel of executive control generally, though, some have argued only for partial relationship between the two (e.g., Calabria et al., 2012).

This thesis makes novel contributions to this theoretical debate by assessing whether individual differences in non-linguistic, domain-general executive control among bilingual speakers relate to bilingual language production across different communicative contexts, such as 1) spontaneous speech produced in L1 and L2 monologues and dialogues (Chapter 2), 2) production of grammatically well-formed sentences and single words across single- and dual-language contexts (Chapter 3). The present thesis also makes a novel contribution by assessing the role of individual differences in executive control among bilinguals during bilingual sentence reading (Chapter 4), a linguistic process thought to be impermeable to domain-general cognitive process, given currently accepted bilingual reading models (e.g., BIA+, Dijkstra & van Heuven, 2002).

With respect to Chapters 2 and 3 of the thesis, which focused on production, one goal was to assess whether the link is observed across different communicative contexts, such as semi-spontaneous dialogue/monologue speech, grammatically well-formed sentences, and isolated single words.

With respect to Chapter 4, which focused on reading, one goal was to assess whether the link
extends to comprehension. Across studies reported in this thesis, bilingual adults completed L1 and L2 production tasks (Chapters 2, 3) and an L2 sentence-reading task (Chapter 4). They also completed a task battery assessing individual differences in L1 and L2 proficiency/knowledge and executive control across all studies reported in this thesis.

The following section summarizes the goals and the results of each study followed by a discussion of how these results together can be used to support recently proposed models of bilingual language production. The last section of this thesis will discuss possible limitations and future directions for research.

**Summary of studies**

**Chapter 2 – Inhibitory control and L2 proficiency modulate bilingual language production: Evidence from spontaneous monologue and dialogue speech (Pivneva, Palmer, & Titone, 2012, Frontiers in Psychology: Bilingualism & Cognitive Control Special Issue).**

The central goal of the study presented in Chapter 2 was to investigate whether and how individual differences in executive control related to bilingual language production in naturalistic communicative contexts, specifically spontaneous L1 and L2 monologue and dialogue speech. Additionally, this study assessed whether individual differences in L2 proficiency also modulate bilingual language production. Participants completed a modified version of the Map Task to index bilingual language production, with respect to the content of what was produced and/or the
acoustic-temporal efficiency with which speech was produced. Participants also completed an executive control and L2 proficiency task battery.

Greater executive control among bilinguals related to greater efficiency (as evidenced by acoustic-temporal index of a vocalization duration divided by its preceding pause duration ratio) in speech production, especially when bilinguals conversed in a dialogue in their less-dominant L2. In contrast, executive control did not relate to the content (as evidenced by clarity of instructions, speaker fluency and nativeness) of what was produced. Of note, individual differences in L2 proficiency related to the content of what was produced and also to the efficiency with which speech was produced. Altogether, the results suggest that L2 knowledge significantly affects all aspects of bilingual monologue and dialogue speech production, while domain-general cognitive processes mostly affect the efficiency with which bilinguals produce speech to attain fairly uniform content production. These findings replicate prior single-word production work showing that greater executive control relates to more efficient management of cross-language activation during L2 production (Linck et al., 2008; reviewed in Gollan & Kroll, 2014). Here, we extended these results to language production in more naturalistic communicative contexts found in daily life.

Chapter 3 – Bilingual language production and individual differences in executive control: Differential effects of producing words in sentences vs. isolation (Pivneva, & Titone, under review, Cognition).

The central goal of the study presented in Chapter 3 was to investigate, with a greater degree of experimental control than my previous work (Chapter 2), whether executive control
demands in bilingual language production differ between different communicative contexts across two Experiments. In Experiment 1, we used a novel method, eye movement monitoring, to assess production of L1 and L2 grammatically well-formed sentences, while in Experiment 2, we used classically-used behavioral methods to assess production of L1 and L2 words in isolation. Across both Experiments we again assessed individual differences in non-linguistic executive control.

Within each Experiment of Chapter 3, a secondary goal was to assess whether executive control demands also differed when bilinguals produced sentences and single words in single- and dual-language communicative contexts.

Here, we found that greater executive control significantly predicted the success with which bilinguals produced grammatically well-formed sentences in Experiment 1 and to a lesser extent single words in Experiment 2. Greater executive control also predicted the success with which bilinguals inhibited globally the entire linguistic system versus locally specific words. Lastly, executive control predicted the success with which bilinguals had managed production in a single-language context, where it is optimal to inhibit an irrelevant language (e.g., speaking to a monolingual friend), versus dual-language context, where it is optimal to keep both languages active. These results replicate prior single-word production findings demonstrating that potent communicative contexts facilitate inhibition of an entire linguistic system (e.g., Van Assche et al., 2013). Furthermore, these results replicate prior findings demonstrating that greater executive control bilinguals have fewer unintended language switches in single-language contexts (e.g., Festman & Munte, 2012; Festman et al., 2010). Importantly, the study reported in Chapter 3
extends these results to semi-spontaneous language production but in a presence of greater experimental controls.

**Chapter 4 – Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements (Pivneva, Mercier, & Titone, 2014, Journal of Experimental Psychology: Learning, Memory, & Cognition).**

The central goal of the study presented in Chapter 4 was to assess whether the link between executive control and bilingual language processing extends to bilingual language comprehension. We focused specifically on written comprehension, because it is thought that cross-language activation observed in reading and understanding sentences should be relatively immune to executive control effects (Dijkstra & Van Heuven, 2002), especially during early processing stages (e.g., word form decoding). We used a previously established eye movement monitoring paradigm (Libben & Titone, 2009) to assess cross-language activation during reading. Specifically, bilinguals read high- and low-constraint English sentences containing interlingual homographs, cognates, or control words. Cross-language activation was indexed via interlingual homograph interference, which is the difference in reading times between interlingual homographs and matched control words, and cognate facilitation, which is the difference in reading times between cognates and matched control words. Participants also completed an executive control and L2 proficiency battery.

We found that greater executive control among bilinguals related to reduced cross-language activation, as indexed by interlingual homograph interference but not as indexed by cognate
facilitation. In contrast, increased L2 proficiency among bilinguals related to reduced cross-language activation, as indexed by cognate facilitation but not as indexed by interlingual homograph interference. It is likely that increased semantic conflict associated with processing interlingual homographs versus cognates triggered greater executive control recruitment to inhibit the non-target meaning and to facilitate processing early on, consistent with predictions by Macizo and colleagues (2010). Moreover, greater executive control among bilinguals also related to more efficient sentence constraint processing, especially during later processing stages, consistent with findings from monolingual (Chiappe et al., 2000; Prat, 2011) and bilingual (Martin et al., in press; Segalowitz, 2010) studies of context processing. Lastly, executive control significantly related to sentence-level processing in reading, similar to what we observed in bilingual language production (Pivneva & Titone, under review).

With respect to theoretical frameworks, the finding that that executive control related to early reading stages is consistent with the IC model of bilingual language production as applied to comprehension. The IC model proposes a critical role for executive control in language processing and, thus, early effects are to be expected. In contrast, the BIA+ model of bilingual language processing does not have a built-in mechanism to account for the executive control influences observed at the very early stages of comprehension because the language processing system is thought to be architecturally impenetrable to influences of the domain-general cognitive system during earliest stages of processing (as shown in Figure 4, p. 159 of this thesis). This assumption
cannot account for the findings that bilinguals rely heavily on domain-general executive control in sentence reading, an effect similar to bilingual language production.

**General Discussion**

This thesis presented three studies containing four experiments that commonly investigated whether and how executive control capacity among bilinguals related to bilingual language processing, namely production and reading, across different communicative contexts. Greater executive control related to more successful and efficient language production, such as engaging in conversation, or producing or reading grammatically well-formed sentences. Indeed, executive control capacity was particularly important for the efficiency of unfolding real-time dynamics of spontaneous L1 and L2 speech planning processes to yield a fairly uniform percept of the produced content, as evidenced by significant effects of executive control on greater speech planning ratio but not on independent ratings of content measures across monologues and dialogues in Chapter 2. Executive control capacity was also particularly important for managing cross-linguistic activation at the global rather than local levels of processing, such as producing grammatically well-formed sentences versus single words and optimally inhibiting an entire linguistic system in pure versus mixed-language blocks in Chapter 3. Lastly, executive control capacity was particularly important for reading sentences with words (i.e., interlingual homographs), which created cross-linguistic conflict, and for enhanced processing of contextual information in general in Chapter 4. Taken together the results suggest that while executive control capacity is generally important for
bilingual language processing, the extent to which it is recruited varies as a function of processing demands of particular communicative tasks. More importantly, executive control might, in fact, be “the language control” that bilinguals recruit to support their bilingual language production and comprehension processes, which include cross-linguistic activation, among others.

**Application to bilingual language production models**

The results of this thesis provide additional evidence in support of Green’s (1998) IC model of bilingual language production, which posits that bilinguals recruit their domain-general executive control to facilitate bilingual language processing, especially in demanding communicative contexts. Furthermore, the studies reported here are potentially significant for future refinement of bilingual language processing models that currently remain agnostic to the interactions between domain-general cognitive and language processing systems, such as the mental firewall accounts in production (e.g., Costa, 2005; Costa & Caramazza, 1999; La Heij, 2005), and the BIA+ model in reading (Dijkstra & van Heuven, 2002).

More specifically, these alternative accounts do not acknowledge executive control involvement in language processing. While certain proposed mechanisms of local control to reduce cross-language activation are specified within a particular processing level (i.e., lateral inhibition of non-target linguistic representations at the level of word forms or orthography/phonology), the control mechanisms are specific to the internal functioning of the language processing system, and, thus, are architecturally blocked from the domain-general cognitive system. That we observed an
association between performance on tasks measuring non-linguistic executive control ability and the efficiency with which bilinguals produce and understand language, demonstrates that language processing does, in fact, rely on control mechanisms from the domain-general cognitive system outside the language processing system. One way to improve the alternative models is to allow for full interactions between domain-general cognitive mechanisms and the language-processing system throughout different processing levels. Another way is to allow for interactions in situations with increased cognitive demands, or at least early on in processing. Both of these proposed alternatives are currently supported by the IC model (Green, 1998) and its extension (Abutalebi & Green, 2007). Consistent with this idea, Runnqvist and colleagues (2014, p.183), from the mental firewall accounts, have recently begun to define the language-specific selection mechanism as “a lexicon-external monitoring device capable of restricting lexical search exclusively to the intended language, while ignoring activated words in the nonintended language.” While it is unclear whether the said monitoring device is, in fact, executive control, we can allude that it can possibly reflect domain-general cognitive mechanisms (possibly selective attention), which are external to the language-processing system.

Although our results do not allow us to differentiate between specific functions of control mechanisms recruited from domain-general cognitive processing, we hypothesize that control mechanisms largely represent inhibitory functions, consistent with most recent models of executive control (Miyake & Friedman, 2012) and bilingual language processing (Kroll et al., 2014; Kroll & Gollan, 2014). Therefore, the alternative accounts would need to consider modifications that
represent largely inhibitory functions (and less likely selective attention) in the interactions between executive control and language processing system to account for the results reported in this thesis. The proposed modifications could then explain global early proactive-type of inhibitory control mechanisms that we observed across studies reported here and still allow for more local later reactive-type of control mechanisms to engage when necessary. However, we cannot speculate with certainty whether the reactive control mechanisms would be governed by the domain-general executive control or reflect inhibitory functions, as these were not observed in the results reported here.

**Comprehension versus production processes**

Given the reported empirical demonstration that executive control is involved in bilingual production and reading, it is important to discuss possible processing differences and similarities across two modalities. Generally, production is thought to be the flip side of comprehension with respect to information processing and vice versa. Specifically, production is generally thought to involve top-down processes (e.g., involving prior knowledge, context, expectations), while reading is generally thought to involve bottom-up processes (e.g., combining information based on sensory input). At its extreme, production and comprehension processes can be thought to be independent of each other and, thus, fundamentally different from each other (e.g., Shea, 2014). However, perception, comprehension, and thought cannot exist without sensory input involving bottom-up processes and nor can they also exist without memory or knowledge involving top-down processes.
(Kintsch, 2005). To that effect, recent work has now shown that both production and comprehension commonly rely on top-down and bottom-up processing (e.g., Angosto, Sánchez, Álvarez, Cuevas, & León, 2013; Goldrick, 2014; Kintsch, 2005) and, thus, share processing resources (e.g., Awh, Belopolsky, & Theeuwes, 2012; Fedorenko & Thompson-Schill, 2014; Geranmayeh et al., 2014), especially during earliest stages of processing (Shuai & Gong, 2014). Hence, “…the question for the theorist is not top-down or bottom-up, but how do these processes interact to produce fluent comprehension,” stated by Kintsch (2005, p.126) with respect to reading.

In that regard, the results presented in this thesis are consistent with prior studies in that bottom-up and top-down processes are both integral to successful bilingual language production and, especially, to bilingual reading. Specifically, in the sentence reading study reported in Chapter 4 executive control (top-down process, if we may speculate) likely interacted with (or maybe even guided) bottom-up decoding processes to allow faster and/or more efficient processing, including 1) faster overall reading, especially within the first 300 ms; 2) earlier cross-language attenuation when the decoded words presented semantic conflict (e.g., interlingual homographs); and 3) an enhanced use of context when reading sentences with language-unique words. Similarly, in the sentence and single word production studies in Chapter 3, bottom-up processes involved in perception and picture decoding likely interacted with executive control once again to allow faster and/or more efficient processing, including 1) faster overall naming for greater executive control bilinguals when naming single words in Experiment 2; 2) greater and more consistent access to L1 and L2 knowledge (e.g., accuracy) when naming sentences in Experiment 1; 3) more efficient
cross-language attenuation when producing sentences in Experiment 1; and 4) enhanced use of bottom-up task cues (e.g., yellow background cue to speak L2) to adhere to task demands when producing sentences in Experiment 1. Lastly, in bilingual conversation study in Chapter 2 involved an orchestrated integration of production (e.g., producing instructions) and auditory (e.g., understanding the collocutor) and written (e.g., reading map labels) comprehension processes as they unfolded dynamically in real-time. To that effect, bottom-up processes interacted with executive control to facilitate 1) decoding images and words printed on a map; 2) perceiving and decoding auditory incoming signal; 3) adjusting planning efficiency in real-time to adhere to the unfolding task demands and to the dynamics of conversational timing rhythms. To sum up, across studies reported in this thesis, top-down and bottom-up processes likely interact and guide production and comprehension processes in young adult bilinguals, especially in cognitively demanding situations often encountered in daily life.

**Current developments in bilingual language processing**

Over decades our knowledge about bilingual language processing and general cognition has improved significantly with important theoretical contributions based on multiple studies. However, it is unclear if any bilingual model is the panacea to account for different result patterns found across various studies (Runnqvist et al., 2014). Often, differences in result patterns can be attributed to methodological inconsistencies with respect to participants and tasks rather than specific models. For example, studies are often inconsistent in task types and administration, such
as pre-training on stimuli (which in itself could change production processes from retrieval to recall); the number of repetitions per stimulus; or whether single-word production and reading, in fact, reflect the complexity of speech planning and comprehension processes occurring in real-time. Recently, Runnqvist and colleagues (2014, p.193) eloquently proposed that “it is essential to generate novel approaches” to study bilingual language processing besides picture-word interference and language-switching paradigms. To that effect, we and others have shifted towards newer methods, such as methods measuring cumulative semantic interference within- and cross-linguistically (Runnqvist et al., 2012), name agreement and semi-spontaneous methods of production and eye movement measures of production and comprehension reported in this thesis.

With respect to differences attributed to participants, some studies ignore or remain agnostic to important characteristics across bilingual participants, such as conflating L1 and L2 processing across different participants; recruiting balanced or reverse-dominant bilinguals for whom differences across two languages are minimal or reveal unclear dominance patterns; or failing to capture important linguistic differences associated with different behavioral ecologies (Green, 2011). With statistical methods recently adopted by psycholinguistics, such as linear mixed-effects regression, it is now possible to capture continuously the multidimensionality of bilingual differences within- and across-participants to provide more-refined theoretical contributions. More importantly, new statistical methods allow researchers to assess with greater precision the dynamics of adaptive behaviors in bilinguals across trials within an experimental session (Pivneva, Free, & Titone, 2014; Wu & Thierry, 2013).
In addition, with respect to theoretical innovations, the most recent development among the bilingual language processing models, namely the Adaptive Control (AC) hypothesis (Green & Abutalebi, 2013), now emphasizes the adaptive role of executive control in language production across various communicative contexts. According to the AC hypothesis, there are three different interactional contexts: single-language (e.g., speaking one language only), dual-language (e.g., switching languages between sentences in a conversation) and dense code-switching (e.g., switching mid-sentence, mixing and/or blending two languages). Executive control within the AC hypothesis is represented via eight control processes: goal maintenance, interference control composed of conflict monitoring and interference suppression, salient cue detection, selective response inhibition, task disengagement, task engagement, opportunistic planning. The central tenet of the AC hypothesis is that executive control and language production are dynamic processes, which interact differently and to a different extent as a function of different communicative contexts. Largely, languages are thought to compete with each other. To resolve competition, bilinguals recruit executive control to a greater extent (e.g., greater number of processes) in dual-language contexts, where two languages must remain active (e.g., eight control processes), versus single-language contexts where one language must remain active and the other inhibited (e.g., three control processes). In contrast, languages are thought to cooperate with each other in dense code-switching contexts. Within the dense code-switching contexts bilinguals recruit only opportunistic planning, which Green and Abutalebi (2013, p.519) defined as “making use of whatever comes most readily to hand in order to achieve a goal” (e.g., adapting words from
one language to fit into the sentence-frame in another language; using gestures to convey meaning when having difficulties in word retrieval). To summarize, the field has started shifting to regard language production processes as a dynamic goal-directed behavior that bilinguals adaptively regulate using domain-general executive control.

**Bilingualism and consequences to cognition**

Given that bilinguals recruit executive control to support bilingual language processing, one can begin to think of language as one of many skills (e.g., driving, juggling, playing a musical instrument) that humans can learn and perfect (Draganski et al., 2004; Maguire et al., 2000; Munte, Altenmuller, & Jancke, 2002; Schlaug, 2001). If language is a skill governed by the domain general cognitive system, then individual differences in a particular skill significantly matter to the success with which the skill is accomplished (Ackerman, 1988, 1992; Boyle & Ackerman, 2004). More interestingly, overlearned experiences with specific skills often carry cognitive consequences for brain and behavior, such as enlarged hippocampi in London taxi drivers (e.g., Maguire et al., 2000), and grey matter differences associated with musical training (e.g., Munte et al., 2002; Schlaug, 2001) or juggling (e.g., Draganski et al., 2004).

With respect to bilingualism, Bialystok and colleagues have argued for the effects of bilingualism on domain-general executive control for the last 20 years (reviewed in Bialystok, 2009, 2010; Bialystok & Craik, 2010; Bialystok, Craik, Green, et al., 2009; Bialystok et al., 2012; Kroll & Bialystok, 2013). From this work, we now know that there are often behavioral
differences between bilinguals and monolinguals on tasks assessing executive control with bilinguals revealing benefits in inhibition, selective attention and conflict monitoring, especially in samples of young children and healthy older adults for whom executive control is sub-optimal (reviewed in Bialystok, 2009; Bialystok, Craik, Green, et al., 2009; Bialystok et al., 2004; Bialystok et al., 2012; Bialystok et al., 2006; Chertkow et al., 2010; Craik & Bialystok, 2006). The results in younger adults have been mixed, presumably because young adults are at the peak of their cognitive processing Bialystok, 2009; Bialystok, Craik, & Luk, 2009; Hilchey & Klein, 2011; but see Kousaie & Phillips, 2012; Paap, 2014; Paap & Greenberg, 2013; Paap & Liu, 2014). Most recent neuroimaging research now reveals that there are structural changes brought on by bilingualism (Bialystok et al., 2012; Kroll & Bialystok, 2013) and, in effect, different types of bilingualism (Klein et al., 2013). Importantly, bilingualism carries cognitive consequences to brain and behavior, like any other cognitively challenging skill (Baum & Titone, in press; Bialystok et al., 2012; Kroll & Bialystok, 2013; Kroll et al., 2014; Kroll & Gollan, 2014; Titone & Baum, in press; Titone, Pivneva, Sheikh, Webb, & Whitford, accepted). Whether cognitive consequence brought on by bilingualism vary among bilinguals remains open to investigation.

To that effect, my graduate work also included a project in which my colleagues and I collected data from a task battery assessing executive control and L2 proficiency in 967 healthy young adult bilinguals (Pivneva, Sheikh, et al., 2014). We quantified bilinguals with respect to their formative language experience (e.g., varying on a continuum of age of L2 acquisition from language-mixing to language-exclusive formative contexts) and current language experience (e.g.,
whether it matched or mismatched formative language experience). Of special importance, we characterized and capitalized on varied bilingual behavioral ecology (Green, 2011) often found in bilingual cities like Montreal, such that bilinguals can be in language-exclusive (L1-only; L2-only) or language-mixing (L1/L2) communicative contexts at home, at work, or in social situations.

Thus, we hypothesized that bilinguals might reveal executive control benefits when their adult socio-linguistic context mismatched their formative socio-linguistic context, which represents an ecological linguistic niche, specialized for a specific type of executive control recruitment from an early age (Green, 2011). For example, sequential bilinguals who learned their languages one at a time in a language-exclusive manner might work exceedingly hard with respect to executive control recruitment when their current socio-linguistic environment involved active daily use of two languages at home, at work or in social situations, because their formative socio-linguistic niche involved greater practice inhibiting the non-target language rather than keeping both languages active. Generally, earlier age of L2 acquisition related to greater executive control benefits (i.e., smaller Simon RT cost), because bilinguals who learned two languages from birth likely had longer experience recruiting executive control to manage cross-linguistic activation at the time of testing (see also Luk, De Sa, & Bialystok, 2011). Consistent with our prediction, we found that a mismatch in the current versus formative socio-linguistic contexts related to greater executive control benefits among young adult bilinguals. These results suggest that individual differences in bilingual behavioral ecology are an important predictor for executive control
advantages and likely neurocognitive change (for the argument that bilinguals cognitive benefits are a product of sustained context monitoring for linguistic cues see Costa et al., 2009).

Recently, Stocco and colleagues have proposed an elegant neurobiological framework that posits how bilingual experience can enhance executive control (Stocco, Yamasaki, Natalenko, & Prat, 2014). According to this framework, bilinguals begin to exercise their fronto-striatal network, a neural system implicated in language processing and executive control, from an early age by managing cross-linguistic interactions (reviewed in Stocco et al., 2014). Training of the fronto-striatal network depends heavily on the basal ganglia, which controls how information is routed to the prefrontal cortex (e.g., Conditional Routing Model; Andrea Stocco, Lebiere, & Anderson, 2010) and which also supports learning and skill acquisition (reviewed in Packard & Knowlton, 2002). In situations when the skill is automatic (e.g., speaking L1), the information is routed using the cortico-cortical networks with no basal ganglia involvement. The intensity with which the information flows is shaped by previous experience/context. However, in situations when a novel skill is acquired (e.g., learning L2), bilinguals recruit basal ganglia to mediate the flow of information based on the changes in the current environment/context. Thus, bilingual executive control advantages can be a product of strengthening of the fronto-striatal network with repeated L2 practice.

While this thesis cannot directly address the role of basal ganglia in the fronto-striatal network training nor in bilingual cognitive advantages, we can highlight similarities across 1) the results presented in this thesis with respect to early global inhibition; 2) Stocco’s neurobiological
framework with the fronto-striatal network including basal ganglia with respect to its role in information processing (see also Wiecki & Frank, 2013); 3) the role of basal ganglia in prior work on bilingual language processing (Abutalebi, Della Rosa, Ding, et al., 2013; Abutalebi, Della Rosa, Gonzaga, et al., 2013; Abutalebi et al., 2012; Abutalebi & Green, 2007, 2008); and 4) the subcortical-prefrontal network for global directed inhibition within the unified framework of executive control (Munakata et al., 2011).

Limitations

While the studies reported in this thesis make important contributions to the research on bilingualism and cognition, there are naturally several limitations common worth noting.

Cross-linguistic competition versus greater general L2 difficulties

It is unclear if we have successfully observed competition arising from cross-linguistic activation across studies reported here (particularly Chapters 2 and 3 where stimuli did not overlap cross-linguistically). Decades of research have consistently argued for cross-linguistic activation and likely competition across production and comprehension (reviewed in Kroll et al., 2014; Kroll & Gollan, 2014; Van Assche et al., 2012). However, we cannot refute the idea that all effects observed here might be due to general L2 processing difficulties (e.g., slower lexical access; greater difficulties with L2-only competitors; less skilled at L2 phonology/orthography, etc.) and not
necessarily a consequence of cross-linguistic activation and competition (reviewed in Kroll & Gollan, 2014; Runnqvist, Strijkers, Sadat, & Costa, 2011).

It is thought that L2 processing deficits could be attributed to less frequent L2 versus L1 use, as per frequency-lag or weaker links hypothesis (Gollan et al., 2008; Gollan, Montoya, et al., 2005; Gollan et al., 2011). More specifically, bilinguals often use one language (e.g., often the L2) less frequently than the other (and the frequency of use is even “lesser” compared to monolinguals). The reduced practice weakens the links between conceptual and lexical representations (e.g., word forms) and delays lexical access/retrieval, which increases difficulties in L2 processing or even difficulties in L1 processing (Gollan et al., 2008; Gollan, Montoya, et al., 2005; Gollan et al., 2011; Whitford & Titone, 2012). Similarly, using a particular lexical representation (e.g., word form) strengthens the link between conceptual and lexical representations but weakens the links of all other semantically related representations within- and across languages, as per bilingual version of the cumulative semantic interference effects view (Runnqvist et al., 2012; Runnqvist et al., 2014) or monolingual Dark Side model (Dell, Nozari, & Oppenheim, 2014; Oppenheim, Dell, & Schwartz, 2010). While these explanations are consistent with what we observed with respect to general L2 difficulties across all studies or deficits in sentence context processing in Chapter 4, they cannot fully capture the interactions between executive control and bilingual language production and comprehension reported in this thesis. More importantly, we statistically controlled for L2 proficiency across all studies of this thesis.
Specific linguistic characteristics of Montreal

All studies presented in this thesis occurred in Montreal, Canada within an English and French linguistic milieu. Thanks to the official statuses for English and French in Canada, the socio-linguistic environment of Montreal offers varied bilingual experiences (e.g., language-exclusivity; language-switching; language-blending) and provides ample opportunities to advance linguistic knowledge in one or all known languages across these varied communicative contexts (Baum & Titone, in press; Pivneva, Sheikh, et al., 2014; Titone & Baum, in press; Titone et al., accepted). Thus, the results obtained within Montreal’s socio-linguistic environment might extend well to countries, such as Spain (e.g., San Sebastian, Barcelona) or Belgium among others, where two languages have a large public presence, if not an official status. However, it is unclear whether the results observed in this thesis will extend equally well to socio-linguistic environments where only one language has a larger public presence, while the other is a minority language or where varied bilingual experiences are less accepted publicly (e.g., language-switching, language-blending).

Specific linguistic characteristics of bilinguals

A large proportion of participants in the studies reported in this thesis involved bilingual Montrealers. As noted in the introduction, more than half of Montrealers are bilingual and nearly one out of five is trilingual. Montreal’s enriching socio-linguistic environments allow high levels of functioning with respect to L2, notwithstanding motivational aspects, of course. Yet, it is unclear
whether the results reported here will extend to all bilinguals for whom L2 knowledge increase might be driven partly by fulfillment of academic requirements or out of post-migratory necessity to integrate into a new community.

Specific linguistic characteristics of languages at study

Within this thesis, language materials were presented in English and French. We must acknowledge that we selected English and French because a large proportion of Montreal population speaks one or both languages from birth. However, it is unclear if the results reported in this thesis will extend to different language-pairs that could provide alternate linguistic cues that might facilitate or interfere with language production and comprehension processes. For example, conceptual representations within a bilingual lexicon might include cultural cues, such that bilinguals name culturally-biasing pictures (e.g., typical Chinese versus Western lantern) faster in culturally-congruent versus incongruent contexts even in their less-dominant language (Jared, Poh, & Paivio, 2013). In addition, it is unclear if some conceptual representations might be language-specific (or perhaps better represented) when a word label for that concept exists only in one language (e.g., no single word form for privacy in Russian; Pavlenko, 2011) or if the word label in one language refers to multiple objects cross-linguistically but the consistency in L1 naming changes with greater L2 exposure (e.g., krujka, chashka, bokal, stakan, ryumka, fujer as referents for English cup and glass; Pavlenko & Malt, 2011). With respect to reading, different orthographies for a given language-pair (e.g., English and Mandarin) might, in turn, provide less
confounded views of cross-linguistic activation (Thierry & Wu, 2007; Wu & Thierry, 2012a, 2012b).

**No neural evidence for executive control activation**

Given the behavioral nature of the studies reported in this thesis, we do not have neural evidence for executive control recruitment during bilingual language processing. Thus, it is unclear whether bilinguals actively engaged brain areas and networks associated with executive control when completing production and comprehension tasks reported here. However, prior language processing studies have found that bilinguals do recruit brain areas (e.g., DLPFC, ACC, among others) and networks associated with executive control when naming single words of varying difficulty in monolinguals (e.g., Kan & Thompson-Schill, 2004; Loncke et al., 2011) and bilinguals, especially when switching languages (e.g., Abutalebi, 2008; Abutalebi et al., 2012; Abutalebi & Green, 2007, 2008). In addition, prior executive control studies have found recruitment of similar brain areas (e.g., DLPFC, ACC, among others) and networks when participants performed versions of executive control tasks (e.g., Simon and Stroop tasks) employed by the studies reported in this thesis (Banich, 2009; Liu et al., 2004; Silton et al., 2010). Thus, we hypothesize that bilinguals likely actively engaged brain areas and networks associated with executive control when completing production and comprehension tasks reported in this thesis. However, to test this hypothesis future research using neuroimaging methods can assess whether
bilinguals activate similar brain areas and networks when completing executive control tasks and language production and comprehension tasks.

**Future directions**

The work presented in this thesis has enhanced our understanding of the role of executive control in bilingual language production and reading. It also opens many opportunities for future research. For example, research in bilingualism has started to shift from group comparisons to an individual differences approach thanks, in part, to more powerful statistical methods (e.g., Linck, et al., 2008; Linck, et al., 2012; Mercier, et al., 2013; Mercier, et al., accepted; Pivneva, Sheikh, et al., 2014; Pivneva, Free, & Titone, 2014). We encourage researchers to continue to integrate continuous measures of processing, where possible.

Ultimately, language production and comprehension are dynamic ongoing processes that unfold over time. To capture the dynamics involved in natural conversations, future research might consider venturing more “into the wild” by extending single word production to more complex language production as it unfolds in natural conversations. The advent of eye movement monitoring technology, and ever-increasing mobile technology, would likely provide such options to assess real-time processing. In addition, recent advances in statistical methods now allow to employ multivariate statistics (e.g., multiple regression) when studying neural correlates of monolingual language production (Miozzo, Pulvermüller, & Hauk, 2014). These methods can
likely be integrated to investigate the dynamics of neural correlates of bilingual language production processes using magnetoencephalography (MEG).

With respect to replicating findings of this thesis, future studies might consider using spatially sensitive neuroimaging methods (e.g., functional Magnetic Resonance Imaging) to assess whether executive control is recruited during bilingual language production and comprehension tasks reported in this thesis and thus providing the missing neural evidence. In addition, future studies might also consider using temporally sensitive neuroimaging methods (e.g., MEG) to assess whether executive control recruitment, in fact, occurs early in language processing as reported in this thesis. Recent developments in MEG methods (e.g., Brainstorm; Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011) now allow researchers to merge temporal and spatial sensitivity to investigate the dynamics of the relationship between executive control and bilingual language processing.

In addition, future work can investigate the dynamics of language production and comprehension processes when bilinguals encounter a forced or an optional language switch in the middle of an utterance when speech planning and comprehension processes have already been deployed. For example, Gollan and colleagues have demonstrated that providing an option to switch languages during single-word production yields different pattern of switching costs versus forcing bilinguals to switch languages or stay within a particular language (e.g., Gollan & Ferreira, 2009). Future work can conduct similar studies using more naturalistic measures of production, such as eye movement method presented here. More interestingly, future work can also investigate
whether the dynamics of cueing language-switching can be modulated by classes of words (e.g.,
cognates) that share lexical information cross-linguistically and likely facilitate or even prompt a
language-switch. For example, prior studies have shown that bilinguals are likely to switch
languages in the presence of a cognate (e.g., Broersma, 2009; Broersma & De Bot, 2006; Clyne,
1980; Kootstra, Van Hell, & Dijkstra, 2012). However, the overarching interest, at least for the
author of this thesis, would concern the role of executive control across these different switching
scenarios.

Future research might also consider shifting to view bilinguals and/or executive control as
dynamic systems that adapt to ongoing environmental demands (Cocchi, Zalesky, Fornito, &
Mattingley; Green, 2011; Green & Abutalebi, 2013; Wu & Thierry, 2013; Yeung, 2013). For
example, some researchers have started to focus on whether training executive control carries
consequences to monolingual (e.g., Hussey & Novick, 2012) and bilingual (Prior & Gollan, 2013)
language processing. While promising, the results from bilingual studies (Prior & Gollan, 2013)
have shown only limited transfer effects (e.g., from executive control to the less-dominant language
only). Thus, future research can continue to explore the age-old argument of what came first, the
chicken or the egg. Is it the case that active knowledge and usage of the second language (or of any
other skill in that case) presents cognitive advantages? Or is it individuals who have greater
executive control also have greater cognitive resources to seek out willingly enriching experiences
(e.g., juggling, bilingualism, musical training, exercise)? Within this research area, future studies
could specifically assess the role of basal ganglia within the fronto-striatal network as a form of
language control during early global control processes and cortico-cortical connections within the prefrontal cortices as a later local control processes to develop a hybrid neurobiological framework, combining executive control and language processing models (e.g., Abutalebi & Green, 2007; Abutalebi et al., 2012; Munakata et al., 2011; Stocco et al., 2014).

With current medical developments, a larger proportion of the population lives longer and/or more enriching lives. Research investigating the interactions between cognitive and language processing systems in healthy aging could provide important theoretical contributions because bilingual healthy older adults will have many years of actively using their L2 when they begin to face cognitive decline (Salthouse, 1996). Future research can also focus on the link between executive control and bilingual language processing in populations for whom executive control has been shown to be compromised, such as individuals with schizophrenia, or not yet fully developed, such as young children. Potential research avenues with these populations could specifically focus on language-switching in production and comprehension, using non-invasive eye movement monitoring methods.

Conclusion

The results presented in this thesis make significant contributions to research linking bilingualism and cognition. This significant contribution derives from linking individual differences in executive control and bilingual language production and comprehension. Across experiments in three chapters presented in this thesis, we demonstrated this link when bilinguals
engaged in spontaneous speech production in monologues and dialogues, produced fully
grammatical sentences and to a lesser extent when producing single words. More importantly, we
extended the link to bilingual sentence reading. Throughout these studies, we used newly
developed methods, such as assessing unfolding production in conversations and eye movement
methodology to study bilingual production and comprehension. In sum, the work presented here
provides significant and novel contributions to the already established research foundation linking
bilingualism and cognition (e.g., Bialystok, 2009; Bialystok, Craik, Green, et al., 2009; Cook,
1997; Hamers & Blanc, 1989; Kroll & Gollan, 2014; Lambert, 1981; Macnamara et al., 1968; Peal
& Lambert, 1962), all while opening new doors for further directions to explore interactions
between general cognitive and language-processing systems.
References


Goldrick, M. (2014). Phonological processing: The retrieval and encoding of word form information in speech production. In M. Goldrick, V. Ferreira & M. Miozzo (Eds.), *The
Oxford handbook of language production (pp. 228-244). New York, NY: Oxford University Press.


Retrieved from http://www.r-project.org


In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic

Shallice, T., & Burgess, P. (1996). The domain of supervisory processes and temporal organization

Shea, N. (2014). Distinguishing top-down from bottom-up effects. In S. Biggs, M. Matthen & D.
Stokes (Eds.), *Perception and its modalities*. Oxford: Oxford University Press.

*Topics in Cognitive Science, 1*, 305-319.


Shuai, L., & Gong, Ta. (2014). Temporal relation between top-down and bottom-up processing in

A. (2010). The time course of activity in dorsolateral prefrontal cortex and anterior


Appendices
Appendix A. Example stimulus for Chapter 2
Appendix B. A transcribed example of produced conversational speech from Chapter 2

**Participant**

*Participant:* On va maintenant à des rapides sur ma photo mais ça s'appelle encore vue 4, donc peut-être que c'est la même chose. Enfin... On les contemple, c'est un peu en haut pour demander la place vers la gauche de la rivière.

*Conférencier:* Vous la guidez de la carte... OK.

*Participant:* On la contemple, oui. On la demande... De quoi vous parlez donc quelle page. A gauche, deux ou deux demi pages un peu, pour en savoir dans le document dans de la carte, évidemment... En même temps ça joue un peu.

*Conférencier:* À la verticale... OK.

**Translation**

*Participant:* We go towards a river a little. It's in a straight line... on... horizontally. After, I have white water. It's a hill.

*Conférencier:* I touch the river, it's that OK? I have the river here.

*Participant:* Yes. Well, it looks like rapids in my picture but it's called "white water" no maybe it's the same thing. Uh... so we go around them. We pass above to go down to the left of the map.

*Conférencier:* At 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. It's still go down a little. To get to the front third of the area, let's say... we go down vertically.

*Conférencier:* Vertically... OK.

**Participant**

*Participant:* And then do you have tree area?

*Conférencier:* Um... Yes, at the bottom of the page.

*Participant:* Um... No. There is another one.

*Conférencier:* OK.

*Participant:* Um... It's not like the river but it is... It's basically, blue, at the center of the page. Where the map are blue line, towards the center. So you... You go under the tree area...

*Conférencier:* OK.

*Participant:* After the river, and then there is white water...

*Conférencier:* OK. So... It's probably the same thing. And so you go over it and then...

**Participant**

*Participant:* After that, we are going to go in like, a diagonal line, going towards the stones area. Like, at that point, it's going to be on your right.

*Conférencier:* OK... well... My stone area is like, really at the bottom of the page.

*Participant:* Yes. OK. OK area no... Um... like, the stone area on my page. It's like... It's in the center, behind the map. Except that it's like... They are separated by tree or something.

*Conférencier:* OK. OK area I have rapids that are really a bit down. Below the area. It's like... Some stone area 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...

**Participant**

*Participant:* You go towards the left of the sheet.

*Conférencier:* Also OK. So, I just go in a straight line?

*Participant:* In a straight line between the stone area and the coast.

*Conférencier:* Stone area, I only have a stone area, like, at the bottom of the page. But not...

*Participant:* OK, will you go in a diagonal line, at the left of the coast.

*Conférencier:* At the left of the map... Like, how many kilometers can I move from the natural?

*Participant:* Um... 1.

*Conférencier:* 17 OK. So I go diagonal. Like 45 degrees?

*Participant:* Um... Yeah.

*Conférencier:* And where do I stop? At the rapids?

*Participant:* Um... not yet.

*Conférencier:* Not yet. OK.

*Participant:* When you go down in a diagonal line for maybe about 5 km.

*Conférencier:* 1, 3, 4, 5, OK.

*Participant:* And then you have to go at the top of the planets of the rapids.
Appendix C. A graphic depiction of a typical acoustic output produced in Chapter 2
Appendix D. Graphic representation of the anti-saccade task from Chapters 2 and 4

Eye-Link 1000 tower mounted system will be used to monitor and record fixation durations of the right eye to peripherally presented black targets, while participants complete randomly intermixed 24 pro-saccade and 24 anti-saccade trials. The green central fixation square will cue the onset of a pro-saccade trial, and the red will cue the onset of an anti-saccade trial. The central fixation square will remain on the screen for 1000, 1250 or 1500 ms. Anti-saccade cost variable will be computed on correct trials only by subtracting the average reaction time of pro-saccade trials from the anti-saccade trials.

Pro-saccade trials

- Small black center fixation circle to cue trial onset
- Green central fixation square cues pro-saccade trial
- 12° visual angle to left or right of fixation

Anti-saccade trials

- Small black center fixation circle to cue trial onset
- Red central fixation square cues pro-saccade trial
- 12° visual angle to left or right of fixation
Appendix E. Graphic representation of the non-linguistic Simon and Stroop tasks from Chapters 2-4

The Simon task: Participants press the left button when the arrow points up and the right button when the arrow points down in 40 congruent and 40 incongruent trials. The Stroop task: Participants press the left button when the arrow points left and the right button when the arrows point right in 40 congruent and incongruent trials. Simon and Stroop cost variables will be computed on correct trials only by subtracting the average reaction time of congruent trials from the incongruent trials separately for each task.

Simon task

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="up" alt="Up" /></td>
<td><img src="plus" alt="Plus" /></td>
</tr>
<tr>
<td><img src="plus" alt="Plus" /></td>
<td><img src="down" alt="Down" /></td>
</tr>
<tr>
<td><img src="plus" alt="Plus" /></td>
<td><img src="up" alt="Up" /></td>
</tr>
<tr>
<td><img src="down" alt="Down" /></td>
<td><img src="plus" alt="Plus" /></td>
</tr>
</tbody>
</table>

Stroop task

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="left" alt="Left" /></td>
<td><img src="plus" alt="Plus" /></td>
</tr>
<tr>
<td><img src="plus" alt="Plus" /></td>
<td><img src="right" alt="Right" /></td>
</tr>
<tr>
<td><img src="plus" alt="Plus" /></td>
<td><img src="left" alt="Left" /></td>
</tr>
<tr>
<td><img src="right" alt="Right" /></td>
<td><img src="plus" alt="Plus" /></td>
</tr>
</tbody>
</table>

Responses

<table>
<thead>
<tr>
<th><img src="filled" alt="Filled" /></th>
<th><img src="unfilled" alt="Unfilled" /></th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="filled" alt="Filled" /></td>
<td><img src="unfilled" alt="Unfilled" /></td>
</tr>
<tr>
<td><img src="unfilled" alt="Unfilled" /></td>
<td><img src="filled" alt="Filled" /></td>
</tr>
<tr>
<td><img src="filled" alt="Filled" /></td>
<td><img src="unfilled" alt="Unfilled" /></td>
</tr>
</tbody>
</table>

Congruent trials

Incongruent trials
Appendix F. Graphic representation of the Number Stroop task from Chapters 2 and 4

Participants will use their dominant hand to one of three response buttons that will correspond to the number of presented digits on 40 congruent and 40 incongruent trials. Number Stroop cost variable will be computed on correct trials only by subtracting the average reaction time of congruent trials from the incongruent.
Appendix G. Example stimuli for Chapter 3

High name agreement condition - English

Low name agreement condition - French

Filler trial