Title
The Acquisition and Mechanisms of Lexical Regulation in Multilinguals

Author
Tomoschuk, Brendan

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

The Acquisition and Mechanisms of Lexical Regulation in Multilinguals

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Experimental Psychology

by

Brendan Tomoschuk

Committee in charge:

Professor Victor Ferreira, Co-Chair
Professor Tamar Gollan, Co-Chair
Professor Sarah Creel
Professor Marta Kutas
Professor Harold Pashler

2019

The dissertation of Brendan Tomoschuk is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____

____

____
Co-Chair

____
Co-Chair

University of California San Diego

2019

DEDICATION

I would like to dedicate this dissertation to my supportive husband, Ross, as well as my family and friends that provided me with the opportunities and support that helped me reach this part of my career.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Vic Ferreira and Tamar Gollan for their support as the co-chairs of my committee, as well as my advisors for the last five years. They supported my research questions and ambitions, and were invaluable to me become the researcher and writer I am today.

I would also like to acknowledge the Language Production Lab, whose feedback, guidance and support more generally has shaped my research and my career.

I would like to acknowledge Mayra Murillo and all of the research assistants who helped gather and analyze the data on which this work is built.

Chapter 1, in full, is a reprint of the material submitted to the Journal of Experimental Psychology: Learning, Memory and Cognition. Tomoschuk, Brendan; Duyck, Wouter; Hartsuiker, Robert J.; Ferreira, Victor. S; Gollan, Tamar H. The dissertation author was the primary investigator and author of this paper.

Chapter 2, in full, is a reprint of the material to be submitted to the Journal of Memory and Language. Tomoschuk, Brendan; Gollan, Tamar H.; Ferreira, Victor S. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of the material to be submitted to Language, Cognition and Neuroscience. Tomoschuk, Brendan; Gollan, Tamar H.; Ferreira, Victor S. The dissertation author was the primary investigator and author of this paper.

VITA

Education

2014 – 2019 **Ph.D. Experimental Psychology**, *University of California, San Diego*, San Diego, CA

2010– 2014 **B. Phil. Neurolinguistics**, *Pennsylvania State University*, University Park, PA
B.S. Science, *Pennsylvania State University,* University Park, PA

- Minors: Italian and International Studies

Research Experience

2016 – 2017 **Ghent University Faculty of Psychology and Educational Sciences,** Ghent University

Supervisors: Dr. Robert Hartsuiker and Dr. Wouter Duyck

2014 – 2019 **Language Production Lab**, Department of Psychology, University of California, San Diego

Directors: Dr. Victor Ferreira and Dr. Tamar Gollan

- Independent experimental design, implementation (**Python**), and data analysis (**R, JMP, Audacity**)

2011 – 2014 **Bilingualism and Language Development Lab (BiLD)**, Center for Language Science, The Pennsylvania State University

Director: Dr. Janet van Hell

- Conducted independent ERP research (**Neuroscan, E-prime)**
- Code and analyze various ERP and behavioral tasks **(R, Excel, MATLAB)**

2013 **BULET Lab**, Department of Psychology, Bangor University, Wales

Director: Dr. Guillaume Thierry

- Gather data for ERP study studying Greek-English bilinguals

Select Honors and Awards

- **Fulbright Fellowship,** 2016 - 2017
- **National Defense Science and Engineering (NDSEG) Fellowship,** 2015-2019
- **Competitive Edge Fellowship**, University of California, San Diego, 2015
- **Outstanding Thesis Prize**, Center for Global Studies at the Pennsylvania State University's Center for Global Studies, 2014

Publications and Presentations

2019 Tomoschuk, B., Ferreira, V.S., Gollan, T.H. Translation Distractors Tell You More Than What Not to Say: Language Switching in a Picture Word Interference Task. *Submitted to Language, Cognition and Neuroscience.*

Tomoschuk, B., Gollan, T.H., Ferreira, V.S. The Effect of Language of Instruction on Cognitive Control is Language Wide and Persists with Third Language Practice. *Submitted to the Journal of Memory and Language*

Tomoschuk, B., Duyck, W., Hartsuiker, R.J., Ferreira, V.F., Gollan, T.H. Language of Instruction Affects Language Control in the Third Language. *Submitted to Bilingualism: Language and Cognition.*

2018 Tomoschuk, B., Lovelett, J. A Memory-Sensitive Classification Model of Errors in Early Second Language Learning. In *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications (pp:231-239)*

Tomoschuk, B., Ferreira, V., Gollan, T. A seven is not a seven: Bilingual Self-Ratings Differ Between and Within Language Populations, *Bilingualism: Language and Cognition,* 22(3), 516-536

2016 Tomoschuk, B., Ferreira, V., Gollan, T. Input diversity in Second Language Word Learning. Poster presented at *Conference on Multilingualism,* Ghent, Belgium

2015 Tomoschuk, B., Gollan, T., Ferreira, V. Diversity of Input in Second Language Word Learning. Poster presented at *Architectures and Mechanisms of Language Processing,* Valetta, Malta

2014 Tomoschuk, B., van Hell, J., Thierry, G., Wu, Y. Does Language Modulate Color Perception in Greek-English and Russian-English Bilinguals? Poster presented at *Annual Meeting of the Cognitive Neuroscience Society,* Boston, MA, USA

Tomoschuk, B., van Hell, J., Thierry, G., Wu, Y. An ERP measure of Language Modulated Color Perception. Presentation given at *the Center for Language Science, Pennsylvania State University,* University Park, PA, USA

ABSTRACT OF THE DISSERTATION

The Acquisition and Mechanisms of Lexical Regulation in Multilinguals

by

Brendan Tomoschuk

Doctor of Philosophy in Experimental Psychology

University of California, San Diego, 2019

Professor Victor Ferreira, Co-Chair
Professor Tamar Gollan, Co-Chair

Three sets of studies explore lexical regulation in bi- and trilinguals. Chapter 1 examines the foreign language effect (disproportionate interference between non-native languages) by conducting two experiments in which Dutch-English-French trilinguals monitor phonemes in picture names. Results show evidence of a foreign language effect in this task, and further posit that the possibility that such a phenomenon is driven by language of instruction (the language from which a bilingual learns a third language). Chapter 2 explores this theory with two experiments where Spanish-English bilinguals learned Hebrew from one of their two languages before performing a language switching task between these languages. Results suggest the presence of a language of instruction effect in this task and further explore the mechanics that drive it. Finally, Chapter 3 explores lexical regulation among known languages in a picture word interference task. Spanish-English bilinguals named pictures that had distractor words superimposed. These two experiments show that task strategies are inadequate in explaining translation facilitation effects of this nature, and reveal the translation facilitation effects to be highly robust. Taken together, these three sets of studies establish a new explanation for the acquisition of lexical regulation mechanisms (language of instruction), and explore the nature of bilingual control mechanisms in current theories of bilingual lexical access.

INTRODUCTION

Learning a language is difficult, involving the acquisition of both an entirely new vocabulary as well as the rules and structures that govern the combination of the units in that language. Over and above this, an adult learner of a second (or further) language must integrate this knowledge with their known language, regulating facilitation and interference among these languages on various levels of processing to successfully understand and speak each language separately.

We know that languages can interfere with one another during comprehension (e.g., Dijkstra & Van Heuven, 2002; Thierry & Wu, 2007, see Kroll & De Groot, 2009 for review) and production (see de Bot, 2000 or Kroll & Gollan, 2013 for a review, though see Costa et al., 2017, for a competing account). As such, a bilingual must regulate this interference to successfully use one of their languages. In this dissertation, I explore this regulatory process, posing questions about how it is developed in an adult learner, and the processes by which a fluent bilingual performs this regulation.

In Chapter 1, we investigate interlanguage interference in trilinguals, asking whether and how the three languages interact within the same speaker. Applied linguistic work shows that non-native languages tend to interfere more with each other than they do with the native language (Williams & Hammarberg 1998), despite the evidence in the bilingual literature that the native language typically interferes more with the non-native language than vice versa (except in language switching tasks where the opposite is true). In the first experiment we tested these claims by having Dutch-English-French trilinguals perform a phoneme monitoring task in which phonemes from one of their other languages may interfere with their target language. We found that while working in the third language, French, the non-dominant English interfered more than the native Dutch. In the second experiment, we pose and test a potential explanation for this interference among trilinguals, that the language used in instruction of the third language forms "tighter" lexical connections between that known language and the new language. These links allow for enhanced facilitation and easier inhibition between these languages than between the other known language and the new third language. Figure 0.1 shows a visual depiction of this model, based graphically on the revised hierarchical model (Kroll & Stewart, 1994), with solid lines representing relatively stronger lexical links and dotted lines representing relatively weaker lexical links. In early stages of bilingualism, a speaker gains experience regulating interference from the known language while trying to speak in the new language, which creates tighter links between the known language and the new language, more so than between the unknown language and the concept (or semantic) level, causing speakers to rely on this translation. We posit that with the introduction of a third language, the known language used in instruction creates greater interference in learning, and therefore allows for more practice regulating that interference (and therefore tighter lexical links), as compared to between the other known language and the new third language, or between the new third language and the concept (or semantic) level. To test this, in Experiment 2, Dutch-English bilinguals learned words in an artificial language from either Dutch or English before performing the same phoneme monitoring task as in Experiment 1. We found that language of instruction does modulate language control, causing bilinguals to make more errors monitoring phonemes from the language they did not learn from.

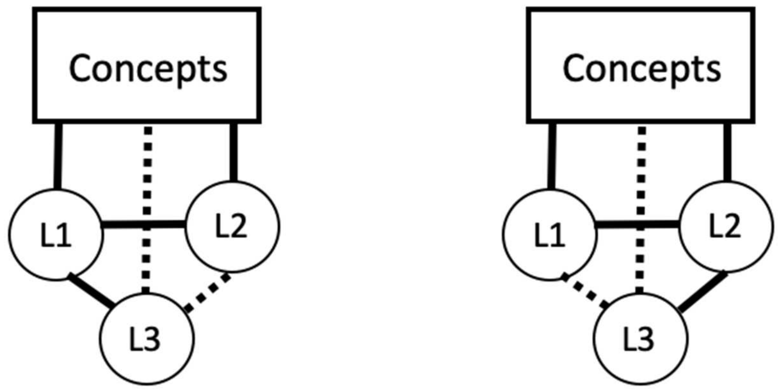

Figure 0.1. *Visual representation of language of instruction effect on third language learning, based on the Revised Hierarchical Model (Kroll & Stewart, 1994).*

In Chapter 2, two additional studies were conducted to probe the specific mechanisms of language of instruction effects. In Experiment 1, Spanish-English bilinguals learned Hebrew via either Spanish or English translations before performing a language switching task. Critically, though the learning paradigm is similar to that of Chapter 1, the subjects speak and learn considerably different languages, and performed a language switching task rather than a phoneme monitoring task. This difference, along with more robust statistical changes (including a within subjects design as opposed to Chapter 1's between subject design), allow us to move from a more exploratory approach to a probing examination of the mechanisms of such an effect using multiple experimental techniques to approach the same question. In Experiment 2, we test the effect at an increased post-learning lag. We replicate the language of instruction effect (found in Experiment 1), and dive deeper into how the effect interacts with language dominance and other factors. These experiments lead us to raise questions about the nature of the language of instruction mechanism, specifically at which level of processing is a learner inhibiting non-target language information.

Finally, in Chapter 3, we further explore the level of processing at which a speaker of two or more languages inhibits their non-target language. In two experiments, Spanish-English bilinguals performed a picture-word interference task in which a distractor word was superimposed on a to-be-named picture. We explore the well-studied phenomenon of translation facilitation in which superimposing the translation of an item onto a picture hastens its production for bilinguals. This effect, though relatively well explored in the bilingual lexical processing literature, has rarely been explored in a language switching paradigm (as in Chapter 2), allowing us to explore the robustness of the effect and the mechanisms that underlie it. We found that regardless of experimental manipulation, the translation facilitation effect does not disappear or become interference. In combining these experimental methods, we question theories about levels of processing raised in Chapter 2 (and in the literature).

Together, these chapters represent both a novel language of instruction explanation of well-established language interference effects, as well as a careful re-examination of existing theories in the literature. By combining experimental paradigms like phoneme monitoring, language switching, and picture-word interference we generate a more complete understanding of the acquisition and mechanisms of lexical regulation in multilinguals.

References

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it?. *Cognitive science*, *41*(6), 1629-1644.

De Bot, K. (2000). A bilingual production model: Levelt's "speaking" model adapted. *The bilingualism reader*, 420-442.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(3), 175-197.

Kroll, J. F., & De Groot, A. M. (Eds.). (2009). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.

Kroll, J. F., & Gollan, T. H. (2013). Speech planning in two languages. In *The Oxford Handbook of Language Production*.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of memory and language*, *33*(2), 149-174.

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

Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied linguistics*, *19*(3), 295-333.

CHAPTER ONE

Language of Instruction Affects Language Control in the Third Language

Brendan Tomoschuk[a]

Wouter Duyck[b]

Robert J. Hartsuiker[b]

Victor S. Ferreira[a]

Tamar H. Gollan[a]

[a] University of California, San Diego,
9500 Gilman Drive, La Jolla, CA, 92093-0109

[b] Ghent University
St. Pietersnieuustraat 33
9000 Gent, Belgium

Abstract

Applied linguistic theories claim that the pattern of interference among multilinguals' non-native languages is based on similarities in cognitive factors like proficiency or age of acquisition (e.g. Bardel & Falk, 2012). In two experiments, we investigated how trilinguals regulate language control of native- and non-native-language lexical representations. In Experiment 1, 46 Dutch-English-French trilinguals completed a phoneme monitoring task in which they decided if phonemes were present in target-language names; phonemes from non-target languages resulted in longer response times and more false alarms compared to phonemes not present in any translation (as in Colomé, 2001). When Dutch and English were target languages, the more dominant of the other two languages interfered more. However, when subjects monitored in their least proficient language, French, the second language (English) interfered more. In Experiment 2, to explore the possibility that the language from which a bilingual learns a language (i.e. the language of instruction) provides practice inhibiting that language, 95 Dutch-English bilinguals learned items in an artificial third language from translation pairs in either their first or second language before performing the same phoneme monitoring task in the artificial language. Subjects controlled potential interference relatively better when distractors were from the language of instruction, suggesting that language of instruction may account for interference effects previously attributed to other cognitive factors.

Keywords: bilingualism, foreign language effect, L2 Status Factor, phoneme monitoring, language learning

Language of Instruction Impacts Language Control in the Third Language

Choosing among multiple labels to express the same meaning is an often challenging but necessary part of knowing more than one language. When the lexicon includes two or more labels to express the same meaning, cognitive mechanisms must regulate and manage the selection of those labels automatically and quickly during speech production. Considerable research has explored the nature of these cognitive mechanisms in bilinguals, showing how co-activation of both languages interplays with inhibitory mechanisms, allowing for successful production of the right words in the right language (see Runnqvist, Strijkers & Costa, 2014 for review). Much less research, however, has explored how the dynamics of the system interact with the trajectory of how bilinguals learned the languages they speak, and more specifically with the methods by which the new language is learned.

Importantly, we know that both of a bilingual speaker's languages are constantly active during comprehension (e.g., Dijkstra & Van Heuven, 2002; Thierry & Wu, 2007, see Kroll & De Groot, 2009 for review) and production (see de Bot, 2000 or Kroll & Gollan, 2013 for a review, though see Costa et al., 2017, for a competing account). One study revealed such dual-language competition in a production task that did not explicitly or obviously present words from both languages (Colomé, 2001). Catalan-Spanish bilinguals saw pictures and were asked to respond via key press "yes" if a specific phoneme was in the Catalan name for the picture and "no" if not. For example, when viewing a picture of a table, the bilingual should respond "yes" to the letter *t* because the Catalan name for table is *taula*. However, they should respond "no" when the letter is *m*, which is in the Spanish translation *mesa*. Response times were slower and false alarms more likely when responding to letters like *m* than *f*, as the *m* sound is present in the bilingual's other, non-target language (Spanish, in this case) while *f* is not present in either translation.

Though this task involves monitoring and not overt production, it suggests that even when formulating a word in Catalan, the Spanish translation of that word from the other language is active and competing for production.

If all of the languages that a bilingual knows compete in production, what is the nature of the cognitive system that regulates such competition? Green (1998) proposes that the attentional system uses cognitive control mechanisms to inhibit one language in anticipation of the production of another. This inhibitory control model suggests that bilinguals actively suppress the non-target language, thereby resolving the competition, but how does such a mechanism develop in language learners? How does it interact with proficiency or use patterns of any given language? A potentially powerful, and relatively underutilized, way to answer these questions is to determine the patterns of interference among the three languages spoken by trilinguals. In particular, any differences that might arise between interference patterns among each of the pairs of trilinguals' three languages could reveal the principles that lead to that interference, in turn revealing how language activation is controlled more generally.

While there is considerable evidence demonstrating the role of dominance in language interference and control (Runnqvist et al. 2014), there is little to no experimental evidence exploring whether and how much non-native languages might interfere with one another. There is, however, some evidence that non-native languages are more closely linked than either are to the native language, regardless of language similarity. In particular, trilinguals' third languages (L3) have more interaction with their second language (L2) than their first language (L1), even though trilinguals are usually more proficient in their L1 than L2. This phenomenon was first referred to as the "foreign language effect" in Meisel (1983), and later the "L2 Status Factor" in Williams and Hammarberg (1998). The latter researchers sought to determine the patterns of interaction among trilinguals' three languages during connected speech. They studied language production in the L3 (Swedish) of English-German-Swedish trilinguals. They found that when speakers switched out of the L3 without any clear pragmatic purpose, they almost exclusively switched into the L2. They proposed that the L2 and L3 are activated in parallel during L3 production while the L1 is more inhibited. If two languages are activated in parallel during production, seemingly unmotivated switches between those two languages should occur more often than either would with the inhibited L1.

One possible explanation for this non-native language interaction is explored in Bardel and Falk (2012). They argued that for trilinguals, the L2 and L3 are often more "cognitively similar" (i.e., are learned in more similar circumstances) within a speaker, and it is this similarity that leads to transfer between L2 and L3 during early stages of L3 learning. That is, the L2 and L3 often have similar ages of acquisition, learning contexts, and other environmental factors that cause the cognitive system to treat them more similarly, leading to transfer and perhaps even interference that may be attributed to non-native language status. Jiang and Forster (2001) likewise suggest that non-native languages differ from native languages in that they are stored in episodic, rather than lexical memory. Falk and Bardel (2011) explored this cognitive similarity idea by testing French-English and English-French learners of German on object pronoun placement. In French, an object pronoun is placed before the verb (*Je* le *vois* – *I* him see), and in English it is placed after the verb (*I see* him). Interestingly, in German, object placement varies based on whether it is in a main (*Ich sehe* ihn – *I see* him) or subordinate clause (*Du weisst dass ich* ihn *sehe* – *You know that I* him *see*). Falk and Bardel found that when rating German sentences, both groups rated sentences as more acceptable when pronoun placement was similar to their respective L2, regardless of what that L2 was. They claim that because the L2 and L3 were both non-native and similar in terms of relevant cognitive factors, transfer between those two arises more than transfer between either non-native language and the L1.

A variable that has not yet been explored, however, is that bilingual learners in studies like these learned their L3 in an L1 environment, possibly even an L1 classroom, and almost none learned their L3 in an L2 environment. Because they learned their L3 in an L1 environment, they have considerably more experience managing dual language activation between their native L1 and their L3 and much less experience managing dual-language activation between their L2 and their L3. In what follows, we term the possible control benefit that might accrue between a newly learned language and the language used to learn that language an effect of *language of instruction*.

There has been very little work investigating whether the language of instruction used to learn an L3 impacts the outcome of learning. One study (discussed in Bjork & Kroll, 2015) investigated whether bilinguals were better at learning an L3 from one of their already known languages. They trained Spanish-English bilinguals, English-Spanish bilinguals, and Chinese-English bilinguals on Dutch vocabulary via English instruction. They found that bilingual learners performed better on a lexical decision task in the L3 Dutch when they had learned via their L1 English (for English-Spanish bilinguals), rather than their L2 English (in the case of Spanish-English and Chinese-English bilinguals). Though this effect was demonstrated between language populations, and there are likely many reasons why the native Chinese-speaking group would have more difficulty learning Dutch than native English and Spanish speakers, the authors suggest that learning a new language through the L1 allows bilinguals to benefit from practice inhibiting their more dominant L1 during acquisition of L3. These results suggest that the hypothesized effect that the language of instruction has on regulating language activation can impact performance in word learning.

In two experiments, we investigated foreign language and language of instruction effects in relatively low L3-proficiency Dutch-English-French trilinguals. Formal age of acquisition was similar for English and French in these subjects, but proficiency in English was much higher than in French. First, in Experiment 1, we asked whether there is a foreign language effect in trilingual language interference. We recruited Dutch-English-French trilinguals and tested them on a lexical interference task adapted from Colomé (2001). Trilinguals did a block of phoneme monitoring in each language, determining whether or not phonemes were present in the name of the picture for the language assigned to that block. Critically, some of the to-be-monitored phonemes came from one of the trilinguals' other languages. For example, trilinguals saw a picture of a girl and were prompted to determine whether the *m, g,* or *f* sounds were present in the name (for *m,* yes in the L1 Dutch *meisje*, but not in the L2 English *girl,* and the L3 French *fille*; similarly for *g* in the L2 English *girl*, and *f* for L3 French *fille*). We expect that in general, subjects should be more likely to false alarm to phonemes of a more dominant language than phonemes of a less dominant language. Critically though, if the foreign language effect affects performance in the non-native language blocks of this task, subjects working in L2 English should more often false alarm to the (L3 French) *f* sound than the (L1 Dutch) *m* sound. Likewise, more false alarms should occur while working in the L3 French for the (L2 English) *g* sound than for the (L1 Dutch) *m* sound. This would suggest that the non-native languages interfere more with one another, and that the lexicon is subject to the same types of foreign language effect as other levels of language processing (see the word order effects of Bardel & Falk, 2012, discussed above).

In Experiment 2, we explored whether the foreign language effect can at least in part be explained as a language of instruction effect. Dutch-English bilinguals were trained in a novel L3 vocabulary via retrieval practice, with the learning prompt coming either from their L1 Dutch or their L2 English. For example, in a particular trial they might see either the word *meisje* (L1) or *girl* (L2), and were then asked to produce *karante,* the novel L3 translation that was phonologically different from both the L1 and L2, before receiving feedback (in the form of the correct answer) in the L3. After many trials like this, they performed the same monitoring task as in Experiment 1 in their novel L3. If language of instruction gives learners experience inhibiting the language of instruction, thus reducing interference from that language, subjects monitoring the item *karante* should more often false alarm to *g* (present in the L2 English *girl*) than *m* (present in the L1 Dutch *meisje*) if they learned *karante* through the L1 Dutch, *meisje*. But if the new language was learned through the L2 English *girl*, phoneme interference effects should be greater between L1 Dutch and the novel L3 than between the L2 English and the novel L3. Put another way, when a bilingual learns a third language, they may be improving their ability to inhibit all non-target languages more generally, in which case the language of instruction should not affect interference patterns while monitoring the novel L3. Alternatively, they may be improving their ability to inhibit the language specifically used in the acquisition of this L3, in which case they will be better able to mitigate interference from phonemes of the language of instruction than from phonemes in the other language while monitoring in the L3.

Experiment 1

Method

Subjects. Dutch-English-French trilingual students (N = 46) at Ghent University, Belgium participated for credit. All trilinguals spoke Dutch dominantly, followed by English, learned in the classroom and reinforced via media (onset age of exposure *M* = 8.63, *SD* = 3.17), and lastly French, learned in the classroom from about the age of 6 or 7. Their average age was 18.59 (1.75) and 78% were female. Full subject characteristics are shown in Table 1.1.

Table 1.1. *Subject characteristics of Dutch-English-French trilinguals of Experiment 1.*

	Dutch			English			French		
	M	SD	Range	M	SD	Range	M	SD	Range
% Pictures Named	81.27	14.77	34-58	43.30	18.53	7-48	10.45	6.18	2-16
Self-rated Listening	6.71	0.82	4-7	5.24	0.95	3-7	4.50	1.00	4-7
Self-rated Speaking	6.65	0.99	4-7	4.80	0.94	2-7	4.07	1.12	4-6
Self-rated Reading	6.57	1.00	3-7	5.00	0.90	2-7	4.35	1.04	3-6
Self-rated Writing	6.40	0.91	3-7	4.61	1.00	1-6	3.85	1.14	3-6

Materials. A list of five hundred concrete nouns were used to generate the stimuli. This list was reduced to 21 items chosen on the following criteria: there were no cognates among the Dutch, English, or French translations of the item, (2) the phonological forms of all of the translations started with a consonant and (3) the initial consonant phoneme of each word was not present in the other-language translations. For example, *meisje-girl-fille* satisfies these criteria because the three translations aren't cognates, all words started with consonants, and *m* is not present in *girl* or *fille, g* is not present in *meisje* and *fille,* and *f* is not present in *meisje* or *girl*. Phonemes with ambiguous grapheme to phoneme mappings were also not used (e.g. *c* in English). Four items were removed after the experiment was run because they were found to have violated one of the above criteria. For each item in each language, three *yes* phonemes were generated and three *no* phonemes. The three *yes* phonemes were consonants in the item. For words with insufficient unique *yes* phonemes, trials were repeated (e.g., *yes* trials for the Dutch word *jas* were *j*, *s* and *s*). Two of the *no* phonemes were the initial phonemes from the other-language translations and one was a *yes* phoneme from another word within the same language that was not present in any of the three translations of the target word (referred to as the no-language condition). This baseline condition ensured that the frequency of different phonemes was relatively consistent across yes and no responses. The full list of items is presented in Appendix 1.A.

Procedure. Subjects were told that they would be doing a task using all three languages that they spoke. Before being instructed on the particular task, they were given pictures of the items with Dutch, English, and French translations below the words and were asked to briefly familiarize themselves with the specific items they would be using in the experiment (Colomé 2001, also presented subjects with the words that they were to be monitoring in Catalan). Once they were ready, they were instructed on the details of the phoneme monitoring task and given seven practice trials (with experimenter supervision). These practice trials deliberately included *yes* trials in which the grapheme was incorrect, but the phoneme was correct (e.g. *cow* and the *k* sound in English) and *no* trials in which the grapheme was correct, but the phoneme was incorrect (e.g. *shovel* and *s*) to ensure that subjects knew to monitor phonemes, and not graphemes. Each trial consisted of a fixation cross appearing for 350 ms followed by a 150 ms blank screen, then a picture for 400 ms followed immediately by a letter for 600 ms. The subject had 2000 ms from the onset of the letter to respond yes or no on a button box (rightmost button was yes, leftmost no) as to whether the sound, and not necessarily the letter, appeared in the name of the picture, using the grapheme to phoneme mapping of the target language (e.g. *w* represented the /w/ sound in the English block but represented the /ʋ/ in Dutch). After the practice, they were corrected on any mistakes and were told which language they were to use first (order fully counterbalanced between subjects). Each block lasted about 6 minutes and subjects could take a break for as long as they wanted before going onto the next block in the next language. After the three language blocks were completed, subjects completed a picture naming task based on the Multilingual Naming Task (MINT; Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012); words were chosen to represent a wide range of lexical frequencies and there were no cognates between Dutch and English. Subjects named the pictures in a set order and were not prompted for alternative terms. Because the experimenter was a native speaker of English and not Dutch or French, subjects completed this picture naming task first in English, then Dutch, then French. This ensured that any mistakes subjects made interpreting the line drawing were resolved in English. Results are shown in Tables 1.1 and 1.2. This task was used to assess language proficiency (Tomoschuk, Ferreira & Gollan, 2018). Finally, they completed a language history questionnaire estimating their self-assessments of each language and estimated language use, and were debriefed on the study.

Analysis. Data were analyzed in R (R Core Team, 2013). All responses with a response time (RT) less than 100 ms or greater than 2000 ms (i.e. responses erroneously measured after the trial ended and before the next trial began) were removed. When analyzing response time as a dependent variable, all incorrect trials were additionally removed from the analysis.

Adjusted residual rates were calculated based on Hughes, Linck, Bowles, Koeth, and Bunting (2013). An accuracy score was calculated for each condition and participant, along with the average RT for that condition. The RT was then converted to minutes and the accuracy was divided by this value to generate an adjusted residual rate score. This metric reflects the average number of correct responses per minute per condition. For example, a participant who scored an average of 90.5% of correct responses in Dutch when the distractor was in English, and did so with an average RT of 1001 ms has an adjusted residual rate of 54.2 correct responses per minute

(.905 / (1001 ms / 60,000 ms/min)) for that condition. This method helps to account for individual differences in the strategies different subjects may take in that those who respond quicker may show effects as error differences, whereas those who try to avoid errors may show effects as response time differences. This method was chosen over other methods of combining response times and accuracies (e.g. inverse efficiency scores, see Bruyer & Brysbaert, 2011) because it is considered to be robust to higher error rates, which are common in learning experiments. Though residual rates are the focus of our discussion, we first report false alarms (responding that a phoneme was present in a word when it was not) and response times as these were the a priori dependent variables.

First, critical trials from all three language blocks were analyzed together to understand whether errors and response times differed between languages. Then within each block, models were built with Helmert contrasts (Wendorf, 2004). Across both experiments, there were two critical contrasts. In Contrast 1, the control condition (letters that did not appear in the non-target languages) was compared to the combination of the two critical conditions (letters appearing in the translations in the two non-target languages). In Contrast 2, phonemes from one non-target language were compared to phonemes from the other non-target language. These data were entered into a linear mixed effect model which was built with maximal random effect structures; when a model failed to converge, correlations were removed from the model, followed by the slope that accounted for the least variance (Barr, Levy, Scheepers & Tily, 2013).

Results and Discussion

Subjects responded correctly (to both *No* and *Yes* trials) in the Dutch language block on 84.3% (*SD* = 7.02%) of trials, 76.9% (*SD* = 10.6%) in the English language block, and 69.9% (*SD* = 9.96%) of the time in the French language block. Figure 1.1 shows the error rates organized by language block for critical (*no*) trials. There was a main effect of target language block such that error rates were highest in the L3 (French, M = 18.5%, SD = 11.6%), followed by L2 (English, M = 14.8%, SD = 12.0%) and L1 (Dutch, M = 10.3%, SD = 7.9%). This difference was significant across all three languages(χ^2 = 10.39, p = .001), marginally significant between just the L1 and L2 (χ^2 = 3.79, p = .051), significant between the L1 and L3 (χ^2 = 9.94, p = .002) and significant between the L2 and L3 (χ^2 = 5.91, p = .015). This pattern confirms our language proficiency assumptions. While monitoring in the L1, phonemes from the L2 and the L3 were significantly more likely to false alarm than letters from the no-language condition (Contrast 1 of the Helmert contrasts, χ^2 = 38.84, p < .001). L2 and L3 were not differentially likely to false alarm (Contrast 2 of the Helmert contrast, χ^2 = 0.473, p =.492). Likewise, in L2, L1 and L3 phonemes led to significantly more false alarms relative to the no-language condition (Contrast 1, χ^2 = 7.14, p = .008), but not differentially (Contrast 2, χ^2 = 0.363, p = .547). In the L3, however, L1 and L2 phonemes were, together, only marginally likely to induce false alarms relative to the no-language condition (Contrast 1, χ^2 = 2.91, p = .088), and there was a significant difference such that L2 phonemes induced more false alarms than L1 (Contrast 2, χ^2 = 4.17, p = .041).

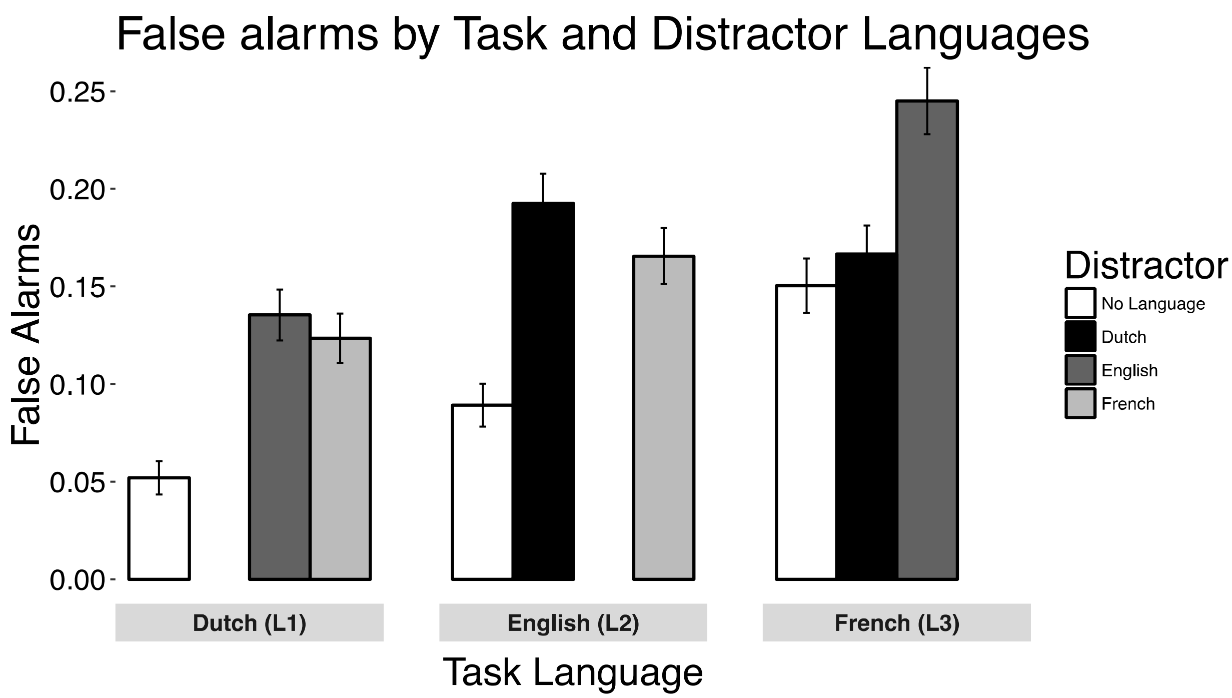

Figure 1.1. *False alarms grouped by whether a phoneme from the distractor language is present in the task-language name of a picture. Error bars represent standard error.*

Figure 1.2 shows response times, also organized by language block. Correct response time trials were log-transformed and analyzed by the same methods. As with the errors, response times first showed a main effect of language (χ^2 = 15.72, p < .001), such that response times were fastest in the L1 block, slower in the L2 block, and slowest in the L3 block. Within the L1 block, response times were slower when monitoring for the phonemes present in the L2 or L3 translations relative to the no-language condition (Contrast 1, χ^2 = 6.79, p = .01), but there was no difference when monitoring for L2 versus L3 phonemes (Contrast 2, χ^2 = 1.02, p = .312). There were no within-language-block differences when monitoring L2 English or L3 French (ps > .11).

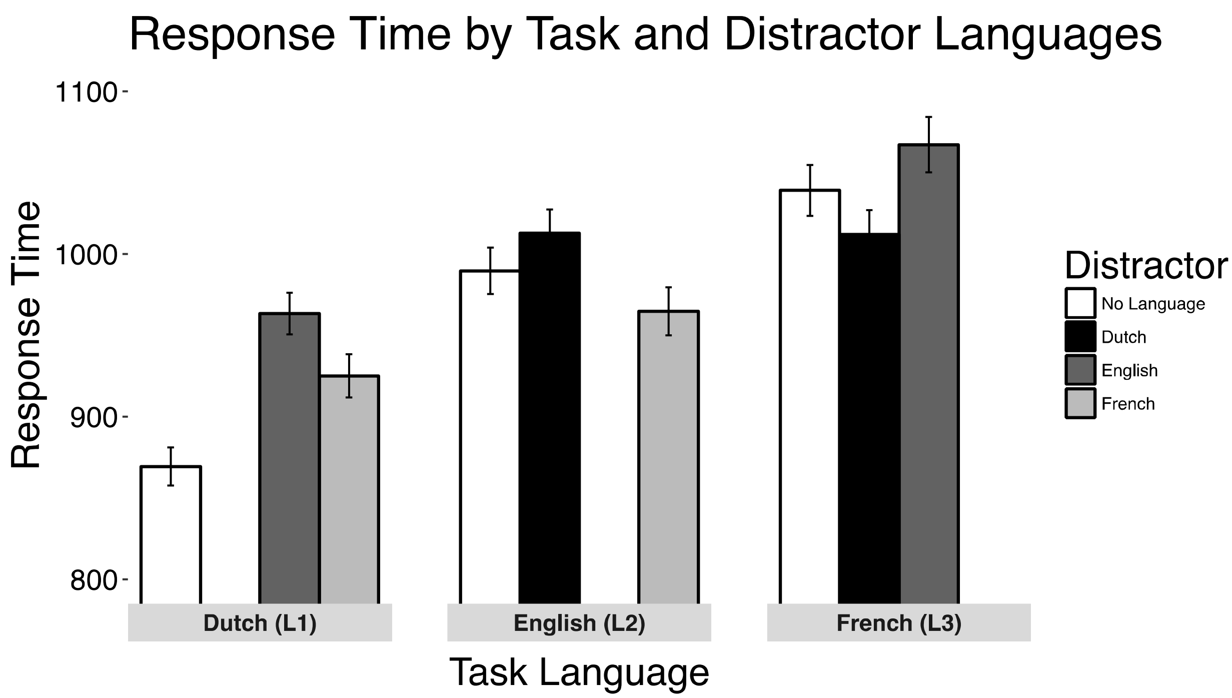

Figure 1.2. *Response times indicating grouped by whether a phoneme from the distractor language is present in the task-language name of a picture. Error bars represent standard error.*

To quantify any potential individual differences in strategies taken in this task, we considered speed-accuracy tradeoffs. On any given trial, a participant could decide to respond more quickly and risk making an error, or could take more time to increase the likelihood of a correct answer. We find evidence that suggests this may be occurring in that in a logistic regression, log response times significantly predict correct responses and interact with both language and condition (*p*s < .05). To capture these tradeoffs, we look at a combined measure of accuracy and response time, residual rates. Figure 1.3 shows residual rates (formula described above). Note that higher residual rates reflect overall better performance in the task (as opposed to higher false alarms which indicate overall worse performance). First, there was an overall main effect of target language block ($\chi^2 = 105.96$, $p < .001$) such that rates were overall highest in the L1 block, lower in the L2 block, and lowest in the L3 block. Residual rates in the L1 were significantly worse when target phonemes appeared in the L2 and L3 translation compared to the no-language condition (Contrast 1, $\chi^2 = 69.29$, $p < .001$). Additionally, phonemes from the L2 had significantly worse residual rates relative to phonemes from the L3 (Contrast 2, $\chi^2 = 6.27$, $p = .012$). Thus, the relatively more dominant L2 interfered more than the less dominant L3 in this block. Likewise, during phoneme monitoring in the L2, there were significantly worse rates when target phonemes were from the L1 and the L3, relative to the no-language condition (Contrast 1, $\chi^2 = 17.68$, $p < .001$), and significantly better when target phonemes were from the L1 than when target phonemes were from the L3 (Contrast 2, $\chi^2 = 10.33$, $p = .001$). Finally, during phoneme monitoring in the L3, residual rates were significantly worse when target phonemes were from the L1 or the L2 relative to the no-language condition (Contrast 1, $\chi^2 = 4.71$, $p = .030$), and they were significantly worse when target phonemes were from the less dominant L2 than when target phonemes were from the L1 (Contrast 2, $\chi^2 = 19.42$, $p < .001$), in contrast to the patterns of the other two language blocks. These effects did not differ significantly between subjects with different block orders (*p*s > .12 for interactions between block and condition), suggesting no carry-over effects between languages.

Additionally, we combined both language and condition in one model in order to examine interactions between language and condition. In this case, the interfering phonemes were refactored into more and less dominant languages. For example, in the L1 Dutch language block, phonemes from the L2 English were factored as the more dominant language and the L3 French was refactored as the less dominant. In the L3 French language block, however, the L1 Dutch was factored as the more dominant language and the L2 English as the less dominant language. When analyzing the residual rates from Experiment 1 in this way, we found that the more dominant language interfered less in the L3 French and more in the L1 English and L2 English, With only 400 ms on the phoneme, preceded by 600 ms on the picture and no blank screen in between may an interaction between language (L1 Dutch and L3 French) and condition (more vs less dominant language). This interaction was significant (t = 2.98, p = .003).

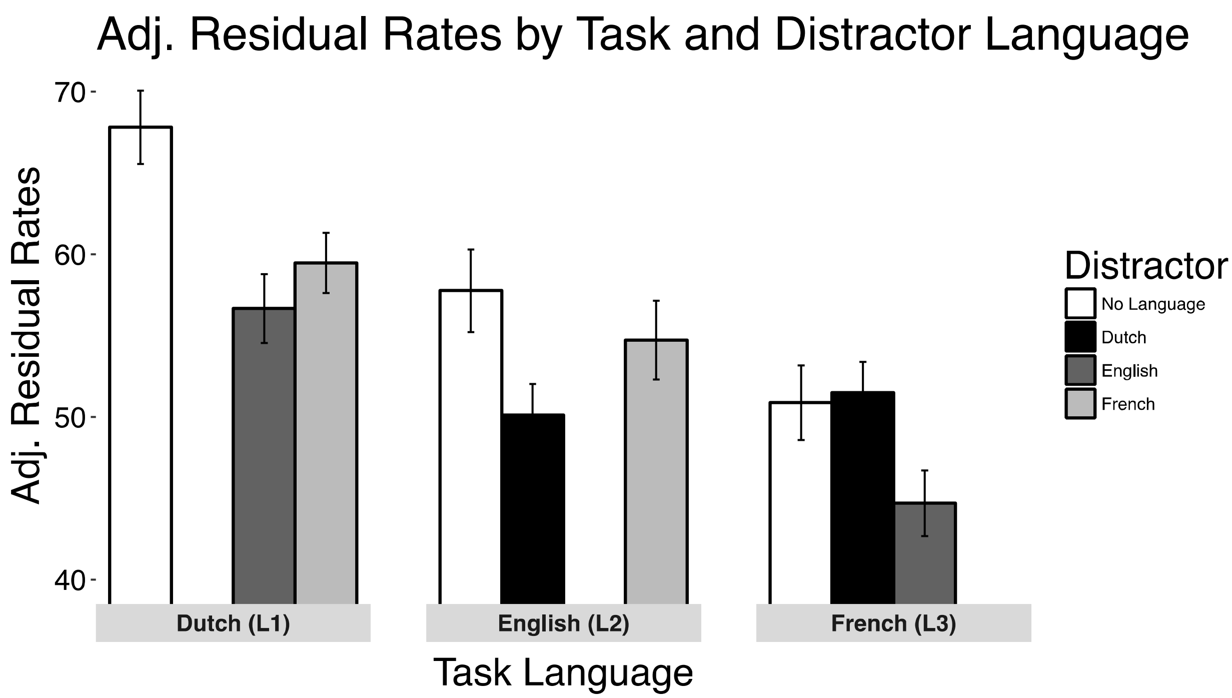

Figure 1.3. *Residual rates grouped by whether a phoneme from the distractor language is present in the task-language name of a picture. Higher rates represent fewer errors and faster responses. Error bars represent standard error.*

These results demonstrate a foreign language effect in a language interference paradigm. In the residual rates (the primary focus of this analysis), we saw that while working in the L3, target phonemes from the non-target L2 reduced rates (i.e. reduced performance) more than target phonemes from the more dominant L1. Additionally, during phoneme monitoring in the L1, target phonemes from the more dominant L2 reduced rates more than target phonemes from the less dominant L3. During phoneme monitoring in L2, target phonemes from the more dominant L1 also reduced rates more than target phonemes from the L3. While working in a higher proficiency language (L1 or L2), the more dominant language tends to interfere more, but while working in a lower proficiency L3, this pattern reverses such that the less dominant L2 tends to interfere more. This pattern suggests that while working in a lower proficiency L3, the cognitive system engages different inhibitory mechanisms to fully suppress the L1 than while working in a higher proficiency L1 or L2. The results of this full analysis pattern identically to the aforementioned Helmert contrasts.

A similar pattern appeared in false alarms. While working in a low proficiency L3, target phonemes from the dominant L1 caused significantly fewer false alarms than target phonemes in the L2. To avoid making errors from their very dominant L1, speakers appear to have inhibited translations from their L1 especially effectively. There was not, however, an effect in response times. This may be due to the nature of the task. In the original experiments in Colomé (2001), effects were seen in response time but not (usually) in error rates. In Colomé's first experiment, the phoneme appeared for 1000 ms, followed by a blank screen for 1000 ms, and the picture for another 2000 ms. With only 400 ms on the phoneme, preceded by 600 ms on the picture and no blank screen in between, subjects may make false alarms than slow responses.

Interestingly, the complementary foreign language effect that one might anticipate seeing in the L2 (such that L3 interfered more than L1) was not found. If the foreign language effect was really about both languages being similar in terms of cognitive profile, we might expect L3 to affect L2 more so than the cognitively dissimilar L1. One possible explanation for this asymmetry is consistent with a weaker foreign language effect explanation, whereby the non-native L3 is especially vulnerable to interference from the non-native L2 because these trilinguals are especially less proficient in their L3, and therefore engage in different, perhaps more top-down, control mechanisms to suppress the more dominant L1 (This rationale is further explored in the General Discussion). It might likewise be suggested that the complementary foreign language effect might not appear when working in the L2 simply because the relative weakness of L3 would not lead to any interference in a stronger language. Critically, though, the L3 did still interfere when monitoring in the L2 and the L1, even in blocks where subjects had not yet performed the monitoring task in the L3. The L3 consistently interferes with other languages at similar levels as the other competing language, suggesting that while the L3 is relatively weak in these subjects, it is strong enough to consistently interfere in other languages during this task.

A different explanation that can more fully account for these data is that these trilinguals learned their L3 in an L1 classroom and surrounding environment – a language-of-instruction effect, as described in the introduction. Because trilinguals were not taught L3 through their L2, they have relatively little experience inhibiting L2 while working in L3. Indeed, foreign language effects in general could be explained by the participant's usage of their two known languages while acquiring the third. As such, in Experiment 2, we manipulate language of instruction and observe how it impacts lexical regulation in the same phoneme monitoring task.

Experiment 2

In Experiment 1, we saw disproportionate interference from the L2 while monitoring phonemes in the L3. This pattern may support a foreign language effect explanation, whereby L3 suffers more interference from L2 than L1, because L3 and L2 are more similar in cognitive profile. Alternatively, it may be due to the fact that these trilinguals' L3 was learned via L1, allowing them to better learn to inhibit L1 than L2 when formulating and monitoring in L3. To explore this language-of-instruction explanation, Dutch-English bilinguals learned new L3 items, either via their L1 or L2. If language of instruction impacts lexical interference, the language of instruction should lead to less interference than the alternative language. In other words, subjects who learn L3 via L1 should show more interference in L3 from L2 (as in Experiment 1), but subjects who learned L3 via L2 should show more interference in L3 from L1. This latter pattern would be the reverse of what should be observed due to a foreign language effect explanation. If language of instruction cannot explain this effect, and instead the disproportionate interference in Experiment 1 is the result of similarity in the mutual cognitive profile of the non-native languages, L2 should interfere with L3 more regardless of the language of instruction. Incidentally, an alternative, simple associative account might predict the opposite pattern: that if the L3 is learned via the L1, then the L3 and L1 words will become associated and so will activate each other. This predicts that if L3 is learned via L1, L3 monitoring should be more difficult for an L1 phoneme, the opposite of the language of instruction predicts that, in this case, monitoring should be more difficult for an L2 phoneme.

Method

Subjects. Dutch-English bilinguals (N = 95) recruited from Ghent University participated for course credit or payment. All subjects were Dutch dominant. Subjects were not recruited based on the knowledge (or lack thereof) of a third language. Five subjects were removed from the analysis for incorrect performance on the task, and a further sixteen were removed for performing statistically at chance, with 60% errors on the phoneme monitoring task (the significance threshold in a binomial test of the same number of trials), leaving 74 subjects. This rate of people performing at or below chance is likely due to the difficult nature of both the learning and monitoring tasks. Further subjects were therefore added until there were 95 that performed above chance, 46 who learned via Dutch, and 49 who learned via English.[1] Subjects in the final Dutch LOI sample were 91.1% female, an average age of 18.8 (1.8) and were first

[1] Note that applying the same filtering procedure to Experiment 1 does not alter the outcomes in any way.

exposed to English at 8.2 (4.0). Subjects in the final English LOI sample were 75.5% female, an average of 20.1 (5.8) years old and first exposed to English at 8.5 (4.0). Full participant characteristics are shown in Table 1.2.

Table 1.2. *Subject characteristics from Dutch-English bilinguals of Experiment 1.*

	Dutch LOI						English LOI					
	Dutch			English			Dutch			English		
	M	SD	Range	M	SD	Range	M	SD	Range	M	SD	Range
% Words produced	87.68	7.54	36-62	59.7	19.46	6-51	87.21	8.51	26-54	55.96	19.09	6-52
Self-rated Listening	6.98	0.15	6-7	5.64	0.64	4-7	6.88	0.44	6-7	5.33	0.8	4-7
Self-rated Speaking	6.91	0.47	4-7	5.04	0.93	3-7	6.72	0.53	4-7	4.76	0.98	3-7
Self-rated Reading	6.93	0.25	6-7	5.51	0.79	4-7	6.93	0.25	5-7	5.64	0.74	3-7
Self-rated Writing	6.88	0.32	6-7	4.95	0.71	3-7	6.86	0.55	5-7	4.91	0.9	2-7

Procedure. Subjects were told that they would do a task that involved 20 Dutch and English words, and they were allowed to familiarize themselves with the Dutch and English names before continuing. Subjects were then told they would learn these words in a new language, called Ibararpa. Artificial words were taken by using Italian pseudowords generated in Wuggy, a pseudoword generator, to ensure naturalistic items (Keuleers & Brysbaert, 2010) that were about equally similar to Dutch and English. They were not told that the language was artificial until the end of the experiment. In each of four learning blocks, they had a brief exposure to each word in Ibararpa, by viewing the picture of the item from the familiarization phase and hearing Ibararpa audio of the word. They had as long as they liked before continuing between words. After being exposed to each word once, they began the training phase. In this phase, half of the subjects saw a Dutch word appear on the left side of the screen, and heard the Dutch audio of the word and had 4 seconds to speak the Ibararpa word. After the four second delay period, the Ibararpa word appeared on the right, also accompanied by Ibararpa audio of the word. Based on pilot data collected to determine the optimal learning structure, a between-subject design was chosen for this study in which subjects learned words in groups of 5, that they practiced retrieving 8 times in a block. Between blocks, they had a break that could last as long as they wanted. At the end, they had one final block that tested their knowledge. They saw each word they had learned presented in the same method (with Dutch prompts and Ibararpa feedback). They only saw each word once. The other half of the subjects completed this same task but used English-based prompts in learning rather than Dutch. This design was made between subject for two reasons: first, a within-subject experiment in which a subject learned some translations from Dutch and others from English would not accurately represent how subjects in Experiment 1 had learned French. Second, a within-subject design would require subjects to learn twice as many translations and therefore, based on pilot data, require exponentially more training time.

At the end of the training, subjects performed one block of phoneme monitoring in Ibararpa. The procedure for this task was identical to the monitoring blocks of Experiment 1. After this, they completed the MINT in Dutch and English and completed a shortened Language History Questionnaire.

Analysis. The analysis was similar to Experiment 1. Helmert contrasts were used to first assess whether there was any effect due to phonemes from the no-language condition, then to assess whether there were differential effects of the two phoneme languages. Language of instruction was also added as a factor. Additionally, Wald Z tests were used to assess model significance.

Results and discussion

Words were considered to be learned if in the final instructional block they were produced within one phoneme of the target word (so an utterance that is only one phoneme different from the target, like *karanta,* was considered a correct production of the target *karante* but an utterance that differed by two or more phonemes like *kamanta* was considered incorrect). Overall, subjects who learned via the L1 Dutch scored slightly higher on the final block of learning (M = 63.6%, SD = 14.9%) than those who learned via the L2 English (M = 62.8%, SD = 16.2%), but this difference was not statistically significant (t = 0.25, p = .80). Additionally, those who learned via Dutch made slightly fewer errors in the monitoring task (M = 25.9%, SD = 7.62%) across all trials than those who learned via English (M = 27.8%, SD = 7.54%), though again the difference was not significant (t = 1.24, p = .221). The response time difference across all trials between the Dutch learners (M = 945, SD = 168) and the English learners (M = 969, SD = 170) was not significant (t = -0.70, p = .484).

Figure 1.4 shows false alarms for Experiment 2. Because the monitoring task was conducted in only the artificial L3, the x-axes in these graphs show the language from which subjects learned the new language. Here we see a significant main effect of interfering language, such that phonemes from Dutch and English induced more errors than phonemes from the no-language condition (Contrast 1, z = -2.20, p = .028), across language of instruction. Subjects who learned via Dutch made slightly more errors with target phonemes from English while subjects who learned via English did not seem to have more errors with the target phoneme from Dutch or English, though this interaction between language of instruction and Contrast 2 was not significant (z = 0.93, p = .35). When analyzing each language of instruction group individually (i.e., when analyzing the data from just the Dutch or English language of instruction groups), there was a significant effect when the language of instruction was English, such that phonemes from Dutch and English were more likely to false alarm than those from neither translation (Contrast 1, z = -2.22, p = 0.026). There were no significant effects when looking at the Dutch

LOI group individually (*p*s > 0.27). In the response times (Figure 1.5), there was a marginal effect when the language of instruction was Dutch such that phonemes from the L2 English were responded to slower than phonemes from the L1 Dutch (Contrast 2, $t = 1.88$, $p = .077$).

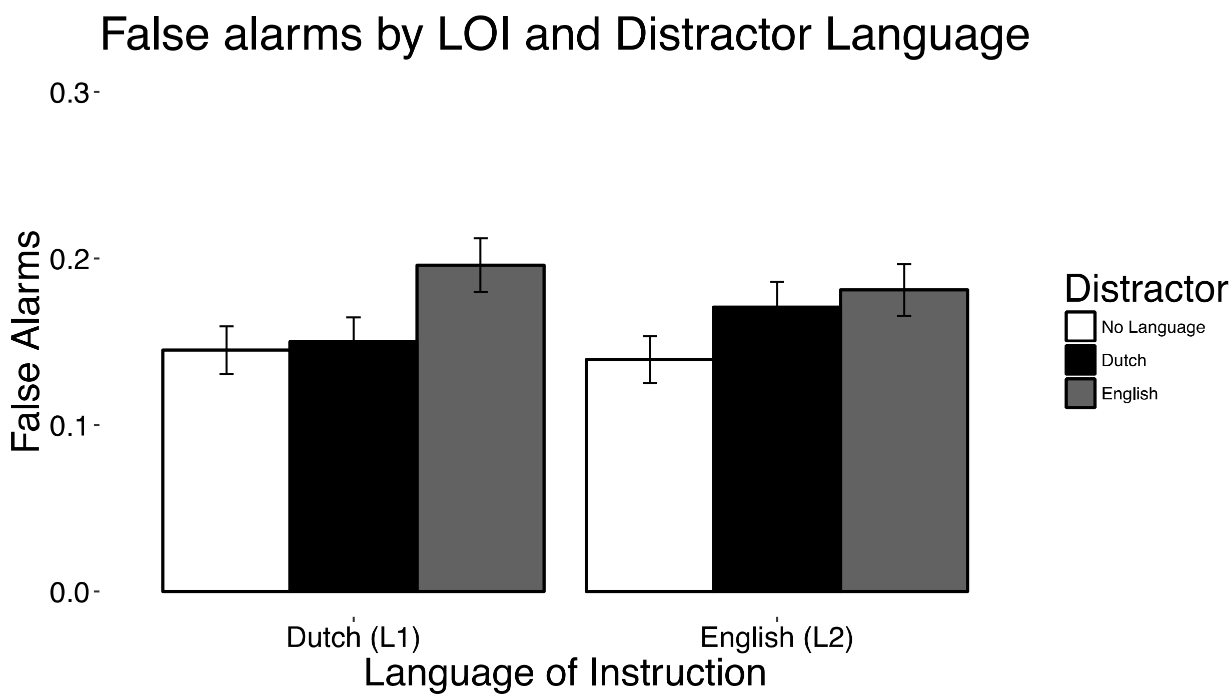

Figure 1.4. *False alarms grouped by whether a phoneme from the distractor language is present in the artificial language, and by language of instruction. Error bars represent standard error.*

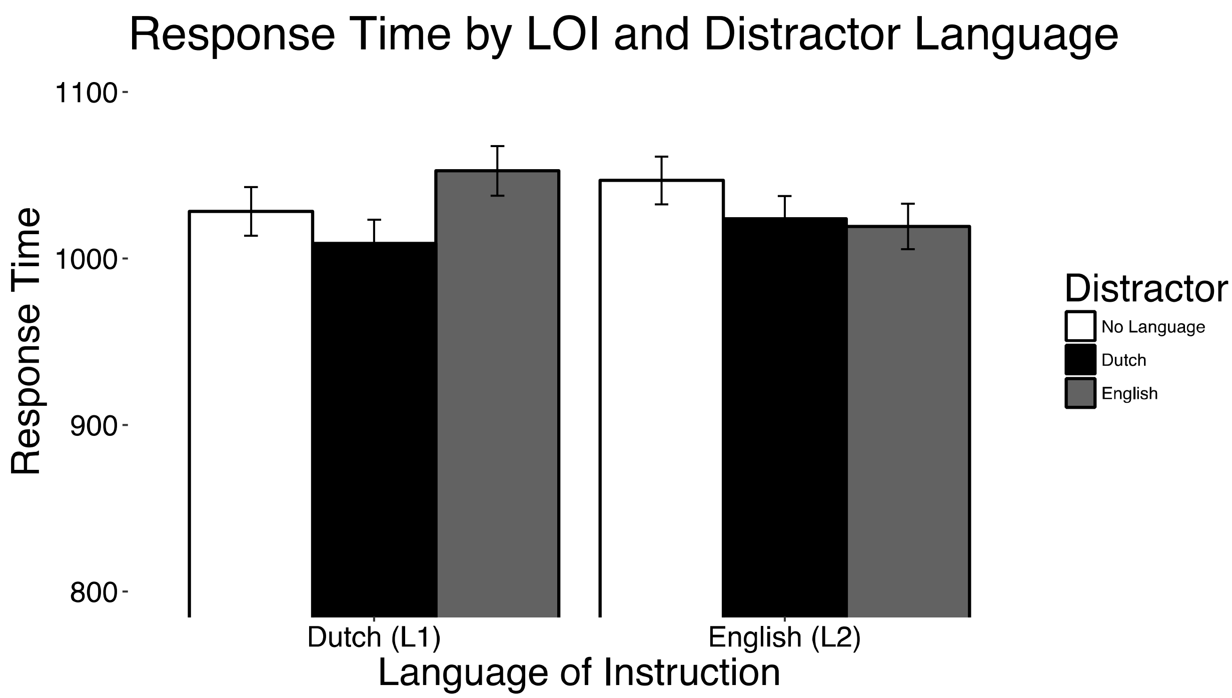

Figure 1.5. *Response times grouped by whether a phoneme from the distractor language is present in the artificial language, and by language of instruction. Error bars represent standard error.*

Figure 1.6 shows residual rates for the monitoring task of Experiment 2. As in Experiment 1, we analyzed speed-accuracy tradeoffs and found that logged response times were a significant predictor of correct responses, and that they significantly interact with both language of instruction (*p*s < .001). As such, we analyzed residual rates. There was a marginal effect, such that rates were lower (i.e. performance was worse) with target phonemes from Dutch and English compared to target phonemes from no language (Contrast 1, $t = -1.83$, $p = .069$). There was an effect in the second contrast such that rates were higher with target phonemes from Dutch, relative to target phonemes from English (Contrast 2, $t = -4.01$, $p < .001$). Finally, there was a significant interaction between language of instruction and the second Helmert contrast such that subjects who learned from Dutch had higher rates with target phonemes from Dutch relative to target phonemes from English, relative to those who learned from English (Contrast 2,

$t = 2.67$, $p = .008$). The latter subjects did not show differential patterns based on which language the target phoneme was from. This difference also appeared when analyzing each language of instruction group individually. When learning occurred via Dutch, the first Helmert contrast was marginally significant ($t = -1.74$, $p = .085$) and the second contrast was significant ($t = -3.82$, $p < .001$). When learning occurred via English, there were no effects in either contrast ($t = -1.63$, $p = .107$ for the first contrast and $t = 0.322$, $p = .748$ for the second).

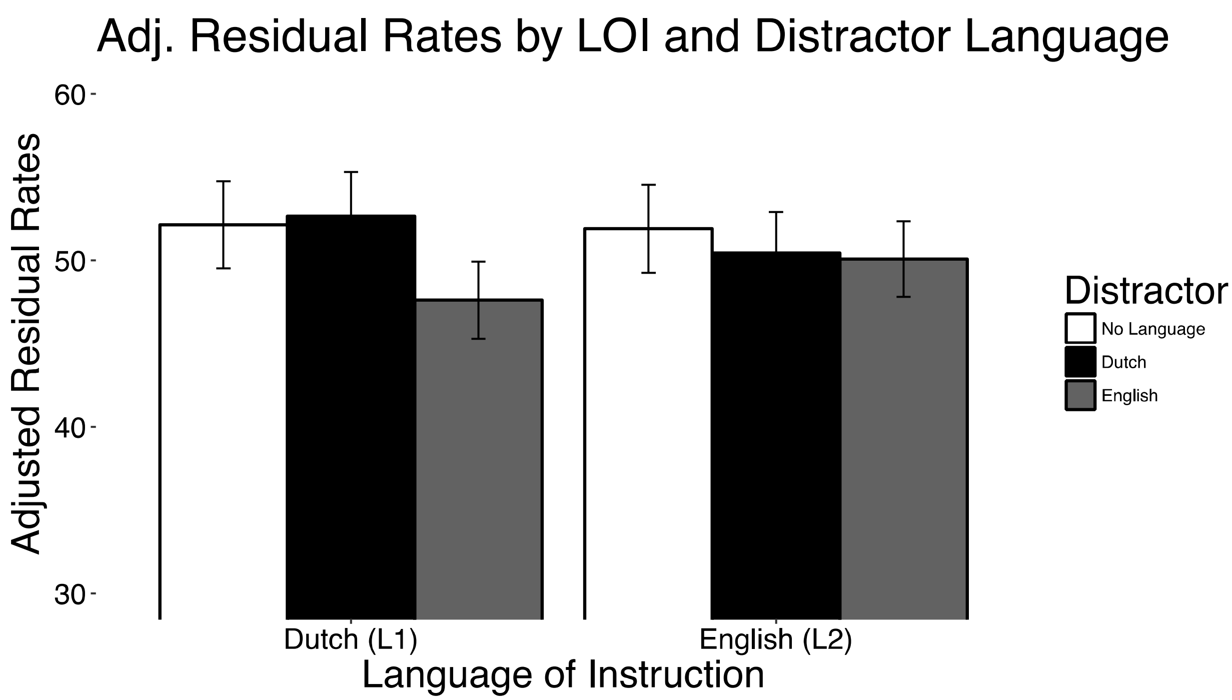

Figure 1.6. *Residual rates grouped by whether a phoneme from the distractor language is present in the artificial language, and by language of instruction. Error bars represent standard error.*

In these results, we expected that while learning, subjects would learn to inhibit the language they learned from, leaving the other language able to interfere while monitoring. While we found this pattern when subjects learned via their L1 (their L2 interfered in monitoring the new L3), we did not see the reverse pattern when subjects learned from their L2, such that their

L1 interfered more in L3 production than L2. Instead, for these learners, there is virtually no interference when working in the L3 and monitoring the L1 or L2. This pattern of data suggests that language of instruction does affect interference patterns in phoneme monitoring, and may explain some effects previously attributed to the foreign language effect.

General Discussion

Two experiments demonstrate that, first, foreign language effects can be shown in language interference tasks and, second, that this effect may be explained in part by language of instruction. In Experiment 1, Dutch-English-French trilinguals performed a phoneme monitoring task, in which they monitored for specific target phonemes in all of their target languages (Colomé, 2001). We observed that phoneme monitoring in L3 (e.g., picture *fille*) was worse when pictures contained target phonemes (e.g., *g*) present in their irrelevant L2 translations (e.g., *girl)*, than when target phonemes appeared in their more proficient L1 (e.g., *m* and *meisje*). In Experiment 2, we tested the possibility that such an effect might actually be driven by a learner's language of instruction. Subjects learned a novel L3 through either their L1 or L2. Afterwards, in the phoneme monitoring task, we again observed that subjects who learned through L1 suffered more interference from L2 than from L1 in the L3 phoneme monitoring task. Interestingly, this effect was not found when subjects learned L3 through L2.

The interaction between the L1/L2 phonemes and language of instruction indicates that language of instruction does impact language interference, at least for languages at low proficiency levels. Though we tested almost 100 multilingual speakers in a difficult language learning experiment, this difference was only significant in the residual rate data, with marginal effects in the false alarms and response times when each language was considered individually. This may be due to a speed-accuracy tradeoff that varies from subject to subject, condition to condition, or trial to trial. Some subjects make quicker responses, and they show effects in errors, while others slow overall response time and make fewer errors, and therefore show effects in response time. This is perhaps not surprising given the outcomes of Colomé (2001). Across the Colomé (2001) experiments, the interference effect was found in response times. In the third experiment, in which the time spent on each trial was shortened, however, the effect was also found in false alarms. The paradigm in the experiments reported here mimics the Colomé (2001) experiment in which the significant differences were found in both response times and false alarms, and so it is possible that different subjects take different strategies on different trials, leading to results being most clear when both dependent variables are considered jointly.

That non-native languages share cognitive resources based on similarity in cognitive profile is a sensible explanation to a commonly reported effect in learners of a third (or later) language. Non-native languages, especially those learned as an adult, have a particular cognitive status, and it is reasonable to theorize that this makes them more likely to interfere with each other when acquiring items in a new language. Non-native languages are often acquired at similar ages, they are less dominant than the native language, and they are similar in other cognitive factors known to impact language interference. There is, however, an alternative explanation that can also account for foreign language effects (in addition to or instead of similarity in cognitive profile), namely language of instruction. By testing language learners who all learned their L2 and L3 in an L1 environment (which is typical), previous studies could not consider language of instruction in the language learning process as a possibly critical factor. Thus, multilinguals tested in those studies had as much experience inhibiting L1 information while working in their non-native language as they had experience in that non-native language. The results of Experiment 2 suggest that this practice, and not exclusively non-native language status per se, can have significant consequences for interference in production of a new language. This language of instruction effect is likely at least a partial explanation of foreign language effects seen in other published studies.

Though we manipulated language of instruction, this effect could instead represent something more like language of the general environment. The trilinguals in Experiment 1 did more than just learn their L3 in an L1 classroom, they were living, working, and studying other topics in their dominant L1. We know that the activation of one of a bilingual's two languages can be boosted, and that this affects their language processing more generally. Elston-Güttler, Gunter, and Kotz (2005) showed German-English bilinguals a twenty-minute-long video subtitled in either German or English before performing a semantic priming task entirely in their L2 English. Subjects performed a lexical decision task after reading a sentence with an interlingual homograph (e.g., *gift* is German for poison). The authors found in both behavioral and neurocognitive measures that semantic priming effects in the first block of the experiment were mediated by the language in which the video was subtitled, despite the identical test materials. This suggests that global language activation can be altered based on a more local environment. Indeed, in our experiment we had subjects working in one of their two languages throughout the majority of the experiment. Our effect, then, may be in part be due to global language activation while learning an L3, rather than or in addition to specific regulation between translation pairs of two languages. Further studies on language of environment may also help us understand the impact of immersion on language control.

Language of environment may also explain why subjects who learned via their L2 did not show L1 interference in Experiment 2. While we manipulated language of instruction, we were not able to manipulate the entire language environment: All subjects were immersed, and dominant, in their L1, likely making it more important to inhibit their very active L1, even when they did not learn from those translation pairs. These data suggest that a bilingual is better able to inhibit their dominant language while monitoring in a less proficient third language, possibly especially when immersed in that dominant language. This, in combination with a language of instruction effect, would explain the pattern of results seen in Experiment 2. Thus, future work can explore the role of the environment and language dominance on language of instruction, and more generally, the inhibitory mechanism used during early stages of learning.

The results of Experiment 1 suggest that the low proficiency L3 is more susceptible to interference from other languages. This idea is also explored in Costa and Santesteban (2004) and Costa, Santesteban, and Ivanova (2006). Across four experiments, Costa et al. (2006) tested highly proficient multilinguals in a variety of picture naming tasks in which speakers were cued to switch between various pairs of their languages. When bilinguals or trilinguals were highly proficient in all their languages, switch costs were the same size in both languages (i.e., were symmetrical), regardless of language similarity or the age of acquisition of the non-native language. However, when proficiency was lower for one of the languages, they found asymmetric switch costs in which switching into the dominant language caused greater slow down, suggesting that the speakers used a different type of control mechanism to manage activation of multiple languages at lower proficiency levels.

In our results, subjects in a relatively low proficiency (Experiment 1) or new (Experiment 2) L3 show a considerable amount of interference from their non-target languages. If these interference patterns were to change with increasing proficiency, as the Costa et al. (2006) study shows, this would suggest that there is some distinct control mechanism used while gaining proficiency in a new language, and, based on Experiment 2, that this mechanism can be impacted by the type, and not just the amount, of experience that a learner receives. More specifically, the asymmetry of the effects seen in Experiment 2 (i.e., that those who learned from Dutch showed interference from English but those that learned from English did not show interference from either English or Dutch) suggests that the control mechanism that trilinguals use to inhibit translations from their L1 when producing in an L3 can be applied to a multilingual's other languages when taught via that language. However, because the L1 did not interfere at all, learners must not need to specifically train to inhibit their dominant L1 while learning. We speculate that the L1 must be dominant enough to warrant inhibition in a low proficiency language regardless of language of instruction.

If, as these studies suggest, low proficiency languages engage a different control mechanism than high proficiency ones, what is the nature of this mechanism? We speculate that the results shown here, in tandem with Costa et al. (2006), suggest that bilinguals need to use strong top-down control mechanisms to prevent dominant languages from interfering in a low proficiency language task. One could argue based on the simple associative account (see Introduction), that the amount of interference seen in this task should be greater from the language of instruction if the connections are simple associations. The experiments here, however, suggest that speakers use top-down control mechanisms to inhibit the most intrusive language, and that language of instruction helps provide experience facilitating that inhibition.

Overall, these studies support the notion that learning lexical items in a new language involves more than just connecting a new word to a known concept; it first involves using learning experience to regulate the connection between those words to allow for successful production in a low proficiency language, and eventually a more direct connection between the new word and the conceptual representation. If a speaker of a second, third, or any new language hopes to reach a level of proficiency in which they do not rely on their language of instruction as a translational scaffold, they must down-regulate the connection between the two lexical items to allow for a stronger link to conceptual representations. Our results suggest that language control mechanisms may differ at earlier stages of acquisition. They also demonstrate that the foreign language effect can be partially explained by this language of instruction, and that this learning experience is critical in first establishing that link between a known word and a new translation.

Chapter 1, in full, is a reprint of the material submitted to Bilingualism: Language and Cognition. Tomoschuk, Brendan; Duyck, Wouter; Hartsuiker, Robert J.; Ferreira, Victor. S; Gollan, Tamar H. The dissertation author was the primary investigator and author of this paper.

References

Bardel, C., & Falk, Y. (2012). The L2 Status Factor and the Declarative/Procedural distinction. *Third language acquisition in adulthood*, *46*, 61-78.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *The American journal of psychology*, *128*(2), 241-252.

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)?. *Psychologica Belgica*, *51*(1), 5-13.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior research methods*, *41*(4), 977-990.

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of memory and language*, *45*(4), 721-736.

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it?. *Cognitive science*, *41*(6), 1629-1644.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of memory and Language*, *50*(4), 491-511.

Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1057-1074.

De Bot, K. (2000). A bilingual production model: Levelt's "speaking" model adapted. *The bilingualism reader*, 420-442.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(3), 175-197.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, *25*(1), 57-70.

Falk, Y., & Bardel, C. (2011). Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research*, *27*(1), 59-82.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(2), 67-81.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *15*(3), 594-615.

Guo, T., Liu, F., Chen, B., & Li, S. (2013). Inhibition of non-target languages in multilingual word production: Evidence from Uighur–Chinese–English trilinguals. *Acta psychologica*, *143*(3), 277-283.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior research methods*, *46*(3), 702-721.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language*, *44*(1), 32-51.

Kroll, J. F., & De Groot, A. M. (Eds.). (2009). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.

Kroll, J. F., & Gollan, T. H. (2013). Speech planning in two languages. In *The Oxford Handbook of Language Production*.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior research methods*, *42*(3), 627-633.

Meisel, J. M. (1983). Transfer as a second-language strategy. *Language & communication*, *3*(1), 11-46.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Runnqvist, E., Strijkers, K., & Costa, A. (2014). 11 Bilingual Word Access. *The Oxford handbook of language production*, 182.

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

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2018). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, 1-21.

Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied linguistics*, *19*(3), 295-333.

Wendorf, C. A. (2004). Primer on multiple regression coding: Common forms and the additional case of repeated contrasts. *Understanding Statistics*, *3*(1), 47-57.

Appendix 1.A. *Items used in Experiments 1 and 2.*

Experiment 1			Experiment2		
English	Dutch	French	English	Dutch	Ibararpa
girl	meisje	fille	girl	meisje	karante
lighthouse	vuurtoren	phare	lighthouse	vuurtoren	fosi
watch	horloge	montre	leg	been	rapo
horse	paard	cheval	horse	paard	borante
smoke	rook	fumée	smoke	rook	bimo
cheese	kaas	fromage	cheese	kaas	gasmirgo
backpack	rugzak	sac à dos	backpack	rugzak	wimu
coat	jas	manteau	coat	jas	pilo
knight	ridder	chevalier	rock	steen	maipa
butterfly	vlinder	papillon	butterfly	vlinder	gotote
feather	veer	plume	feather	veer	kome
fox	vos	renard	pig	varken	siago
trashcan	vuilbak	poubelle	trashcan	vuilbak	zossigi
binoculars	verrekijker	jumelles	binoculars	verrekijker	dototemi
drill	boormachine	percer	drill	boormachine	zoika
deer	hert	cerf	deer	hert	folze
pitcher	kan	lanceur	fox	vos	woddi
			window	raam	tispibi
			tree	boom	poliri
			carrot	wortel	villibu

CHAPTER TWO

The Effect of Language of Instruction on Cognitive Control is Language Wide and Persists with Third Language Practice

Brendan Tomoschuk[a]

Tamar H. Gollan[a]

Victor S. Ferreira[a]

[a] University of California, San Diego,
9500 Gilman Drive, La Jolla, CA, 92093-0109

Abstract

The language from which bilinguals learn a third language can impact their ability to regulate interference between two languages (Tomoschuk et al., submitted). Two experiments probe the mechanism underlying this *language of instruction effect* with an experimental task more relevant to everyday language use. In Experiment 1, Spanish-English bilinguals learned Hebrew vocabulary via either their English or Spanish translations before naming pictures in mixed blocks – one mixed with the language of instruction, one with the other known language – and single language Hebrew blocks. Hebrew naming latencies were faster when mixed with the congruent language – that is, the language used to learn the Hebrew words. In Experiment 2, the language of instruction effect was further tested by having another set of Spanish-English bilinguals perform the same experiment, except with the Hebrew-only picture naming blocks before mixed blocks to increase single language practice before test. The congruency effect persisted despite this single language practice between mixed language production. These results suggest that language of instruction effects exist at the language-wide level and more generally that the circumstances of initial learning are important in forming language-wide control mechanisms.

Keywords: bilingualism, foreign language effect, language control, language switching, language learning

The Effect of Language of Instruction on Cognitive Control is Language Wide and Persists with Third Language Practice

Acquiring a new lexicon is a central and challenging part of attaining proficiency in a new language. Adult learners must map thousands of new labels to existing concepts, often using translation equivalents as an intermediary link (Kroll & Stewart, 1994). In addition to these memory challenges, speakers of two or more languages must regulate activation between languages, choosing the right label for the given context (see Runnqvist, Strijkers & Costa, 2014 for review). While substantial research has explored the mechanisms by which a bilingual regulates language and the processes by which a learner acquires new vocabulary, very little work has explored the development of such regulatory mechanisms during such word learning, and specifically the factors that impact the learning of such regulation.

Importantly, both languages are known to be active during bilingual language comprehension (e.g., Dijkstra & Van Heuven, 2002; Thierry & Wu, 2007, see Kroll & De Groot, 2009 for review) and production (see de Bot, 2000 or Kroll & Gollan, 2013 for a review, though see Costa et al., 2017, for a competing account). Regulating this co-activation is therefore a necessary part of speaking two languages. One relatively well explored question is, at what level do bilinguals administer this control? Do they unilaterally inhibit one language at a time? Or do they suppress lexical units as each competitor is produced? Branzi, Martin, Abutalebi and Costa (2014) posed this question by having bilinguals name pictures in consecutive blocks (all L1 or L2). They found that bilinguals named L1 pictures slower after naming repeated and novel pictures in their L2, suggesting that the entire L1 is inhibited during speech production in the L2. These results overall suggest a language-wide, rather than lexical level, of control (see Green, 1998 for further discussion).

The development of such a lexical control mechanism, however, has been rarely studied. At what point in the learning process does it arise? Is it the same throughout the entire course of adult language learning? Speakers of three or more languages offer a unique opportunity to unpack these factors. For example, there is evidence to suggest that a trilingual's third language (L3) interacts more with the second language (L2) than the first language (L1), over and above any effects caused by language similarity or language dominance. Meisel (1983) coined the term foreign language effect to describe this phenomenon, and it was further explored as the L2 Status factor in Williams and Hammarberg (1998). This work showed that when English-German-Swedish trilinguals switch out of their L3 Swedish without pragmatic purpose, they almost exclusively switched into their L2, and Williams and Hammarberg therefore proposed that the L2 and L3 are activated in parallel to a greater degree than L1 during production.

However, there was neither psycholinguistic evidence for this phenomenon, nor an explanation for its development. Tomoschuk, Duyck, Hartsuiker, Ferreira and Gollan (submitted) sought to provide this evidence and explore a possible explanation for the language-of-instruction effect. Following Colomé (2001), they had Dutch-English-French trilinguals perform a phoneme monitoring task in which some distractor phonemes came from one of the trilinguals' non-target languages. They found that trilinguals performed worse on the task when distractor phonemes came from the trilinguals' other languages, as compared to phonemes from none of the three languages (replicating Colomé, 2001). Critically, they also found that when working in the L3 French, the L2 English induced worse performance than the L1 Dutch, despite the fact that when the target was the L1 Dutch or L2 English the more dominant of the two interfering languages induced worse performance. The authors suggest that that while working in a low proficiency language, a more top-down control mechanism prevents the dominant L1 from interfering. That is, when working in the lower-proficiency L3 French, a top-down control mechanism effectively inhibits the L1 Dutch leaving the L2 English to cause greater interference.

A possible explanation that the authors pursued for how this might come about is that the already known language a bilingual uses to learn a new language (referred to as the *language of instruction*) might form tighter lexical connections between translational equivalents, making lexical-level language control between these two languages easier. To explore this, they taught Dutch-English speakers vocabulary in an artificial language via either Dutch or English (in a between-subject design), and had them perform the same phoneme monitoring task in this artificial L3, to examine whether interference from the L1 Dutch or the L2 English varied by whether it was their language of instruction. They found that indeed the difference in interference from the L1 Dutch or the L2 English varied by language of instruction such that performance on the monitoring task was better for trials where the interfering phoneme came from the language of instruction compared to when the interfering phoneme came from the other language.

This study provided at least a partial explanation of the foreign language effect. Specifically, the language used to learn a third-language vocabulary affects lexical regulation in that language. But there are still many facets of this explanation yet to be explored. First, is the language-of-instruction effect tied to a bilingual's learning experience more generally, or is it tied to specific instructional material? Because Tomoschuk et al. (submitted) manipulated language of instruction between subjects, it is unclear if language of instruction effects would be found for different items within a single learner, or if each learner is geared towards one method or another based on experience. Second, if the effect is tied to material, does it exist at the level of the lexical item, whole-language, or both? Based on Branzi et al. (2014), it may be that language of instruction creates tighter lexical links between translation pairs, but it may alternatively be that the connection is made at some higher level in which a representation encompassing the language of instruction as a whole is involved. Third, does the language-of-instruction effect persist past immediate test? It may be that language of instruction effects only modulate performance immediately after the learning phase, diminishing or fading entirely with increasing L3 proficiency. Fourth, how does language dominance interact with language of instruction effects? The previous study showed that the dominant L1 did not cause interference more than the non-dominant L2 even when the L2 was the language of instruction, suggesting that language dominance is a separate but additive factor that affects the development of this regulatory mechanism. In answering these questions we not only replicate a relatively untested effect, we also explore its nature as an acquisition mechanism in adult language learning.

In two experiments, we address these outstanding questions. First, in Experiment 1, we asked whether the language-of-instruction effect will occur in a new population and extend to a new language control task. Spanish-English bilinguals learned L3 Hebrew vocabulary as in Tomoschuk et al. (submitted). They learned eight Hebrew words via their Spanish translations and performed three picture naming blocks, one with Spanish and Hebrew items (language-of-instruction congruent), one with English and Hebrew items (language-of-instruction incongruent) and one with just Hebrew items. Unlike Tomoschuk et al. (submitted), bilinguals then performed the same series of tasks on a new set of eight Hebrew items but learned the other language (order counterbalanced between subjects). We expect that, if language of instruction effects modulate control over selection of words from the language as a whole, and are tied to instructional material and not experience more generally, then Hebrew naming latencies should be faster when mixed with the congruent language of instruction than the incongruent. Furthermore, if the L3

Hebrew is always faster in the presence of the non-dominant language as the language of instruction, it further suggests that bilinguals are particularly good at inhibiting their dominant language, regardless of language of instruction.

To address whether language of instruction effects persist over immediate test, in Experiment 2, we had bilinguals perform the same learning and picture naming task, but with all Hebrew-only picture naming blocks performed before mixed blocks. Here, we not only replicate effects of Experiment 1, but explore the possibility that effects may change with increased Hebrew-only exposure between learning and mixed blocks. If congruency effects in Experiment 1 are found, and are generated by connections formed between language of instruction and the L3 (either at the whole-language or lexical level) rather than a simple result of increased exposure to the L1 or L2, the difference between mixing costs would persist even after further non-mixed L3 practice.

Experiment 1

Method

Subjects. Spanish-English bilingual students (N = 64) at UC San Diego participated for credit. Three subjects were removed prior to analysis for incorrectly following directions, and as such are not reported. Subject characteristics for the remaining subjects are shown in Table 2.1. All bilinguals spoke Spanish and English fluently. Four bilinguals scored higher in Spanish on the MINT than English and were considered Spanish-dominant for this analysis.

Table 2.1. *Subject characteristics for Experiments 1 and 2.*

	Experiment 1			Experiment 2		
	M	SD	Range	M	SD	Range
Age	20.37	2.30	18-28	20.17	2.12	18-28
% Female	69.84	-	-	70.31	-	-
% English use currently	78.55	12.23	33-95	80.12	14.26	40-99
% English use growing up	58.13	12.78	0-85	55.75	15.34	5-90
Age began using English regularly	5.51	3.29	0-15	5.31	2.71	0-15
Dominant MINT Score	60.45	3.52	52-67	61.03	3.43	51-68
Non-dominant MINT Score	47.00	9.91	23-64	45.37	8.53	22-59
Average English Self-rating	6.41	0.74	4-7	6.43	0.83	4-7
Average Spanish Self-rating	6.19	0.98	2-7	5.97	0.96	3-7

Materials. Sixteen Spanish-English-Hebrew translation sets were generated and split into two groups of eight to-be-learned nouns. Two native Hebrew speakers agreed on Roman alphabet orthographies of the Hebrew items. Items were screened so that they did not share initial onsets between any of the translation pairs. Additionally, the number of phonemes in each language and the frequencies of the items were controlled between lists. The full item list is presented in Appendix 2.A.

Procedure. Subjects first completed a language history questionnaire to ensure they had no previous exposure to Hebrew. Then, subjects were instructed (in English or Spanish based on the language of instruction condition) on the learning phase (adapted from Tomoschuk et al., submitted). They were told they were to use 16 English-Spanish words across the experiment and, to be consistent with other studies of this kind, were asked to familiarize themselves with the particular word-to-picture mappings in English and Spanish (e.g., they should use *gorro* and not *sombrero* to describe the picture of a cap in Spanish). Then, they were instructed on the details of the learning task. In each of two learning blocks, they had a brief exposure to each word in Hebrew, by viewing the picture of the item from the familiarization phase and hearing a recording of a native Hebrew speaker pronouncing the word. They had as long as they liked before continuing between words. After being exposed to each word once, they began the training phase. In this phase, subjects saw an English or Spanish word appear on the left side of the screen (language depending on language-of-instruction condition), and heard the corresponding English or Spanish audio of the word and had 4 seconds to produce the Hebrew translation. After the four second delay period, the Hebrew word (spelled phonetically with the Roman alphabet) appeared on the right, also accompanied by Hebrew audio of the word. Subjects learned words in groups of 4, that they practiced retrieving 6 times in a block (based on pilot data for maximizing learning). Between blocks, they had a break that could last as long as they wanted. After this block, they began again with an exposure phase to the other four words in the same language of instruction, before partaking in a practice phase. At the end of the learning phase, they had one final evaluative block that tested their knowledge. They saw each word they had learned presented in the same method (with English or Spanish prompts and Hebrew feedback). They only saw each word once. After this was complete, they participated in three blocks of a picture naming task.

In the both the single and mixed language response blocks of the picture naming task, subjects were first instructed to name each picture as quickly and accurately as possible, in accordance with a flag cue at the top of the screen (the American flag signaled English, the Mexican flag signaled Spanish, and the Israeli flag signaled Hebrew). First, a fixation cross appeared on the screen for 150 ms, followed by a blank screen for 250 ms, and then the picture and cue appeared. Subjects had 3 seconds from the onset of the picture appearing in which to respond before the picture disappeared. There was an inter-stimulus interval of 850 ms before the next picture appeared. At the beginning of the first block, subjects were given five practice trials in their language of instruction to confirm the sensitivity of the microphone. The first block was either mixed English and Hebrew picture naming, or mixed Spanish and Hebrew picture naming, with a 50% switch rate. The second picture naming block was Hebrew mixed with the not yet tested language, with the order of these mixed language blocks counterbalanced between subject and language of instruction condition. Then, every subject completed a block of picture naming in just Hebrew. The same picture was never named twice in a row, within languages or switching between. After this, the subject completed the entire series of tasks again with the other eight vocabulary items, using the other language of instruction. The order of language of instruction was also fully counterbalanced between subjects. After these tasks were completed, subjects completed the Multilingual Naming Task (MINT; Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012) as a metric of language proficiency and dominance (Tomoschuk, Ferreira & Gollan, 2018).

Analysis. Data were analyzed in R (R Core Team, 2013). All responses with a response time (RT) less than 100 ms or greater than 3000 ms were removed. Trials with erroneous microphone triggers were removed from the analysis. Response times were log transformed. Language-of-instruction congruency was also added as a factor; a trial was considered congruent if it was a naming trial in the English-Hebrew mixing condition, and was also in the English language-of-instruction condition, or if it was a naming trial in the Spanish-Hebrew mixing condition and had been learned in the Spanish language-of-instruction condition. Trials that mismatched between mixing block and language-of-instruction condition were considered language-of-instruction incongruent.

Though we report and analyze response times from all three languages, only Hebrew naming trials were considered critical trials. Three linear mixed effect models were applied to the log-transformed response times. Because language of instruction congruency is not a factor that affects single language Hebrew blocks, the first model compared just language mixing comparing single language blocks to mixed language blocks, the second looked within mixed blocks, where language of instruction congruency can be applied as a factor, using switching, congruency and the interaction as factors. We refer to this model throughout as the *congruency model.* In the final model, to understand the impact of language dominance, we separate language of instruction block with naming block, which we refer to as the *dominance model*. All models used items and subjects as random effects. A full random effects structure was used. To achieve model convergence, correlations among random factors were first removed, followed by factors from the random effect structures that accounted for the least amount of variance (Barr, Levy, Scheepers & Tily, 2013). This analysis was repeated for error rates.

Results and Discussion

In the evaluative learning phase, subjects produced 77.5% (*SD* = 21.3%) of Hebrew words correctly when the language of instruction was English, and 74.8% (*SD* = 23.8%) correctly when the language of instruction was Spanish. This difference was not significant ($t < 1$).

Figure 2.1 shows response times and error rates on Hebrew naming trials organized by switching condition and language of instruction congruency. Blocked items were those named in Hebrew blocks, while stay and switch trials were named in blocks with either Spanish or English trials intermixed. Stay trials were those on which the language of the previous trial was Hebrew, while switch trials were those with either English or Spanish trials immediately preceding them. In the response times we see that bilinguals named pictures slower in mixed blocks than Hebrew-only blocks ($\chi^2 = 54.47$, $p < .001$). In the congruency model, and counter to typical switching experiments, bilinguals named switch trials significantly faster than stay trials (χ^2 = 5.55, p = .018). Additionally, trials mixed with the congruent language of instruction were named significantly faster than trials mixed with the incongruent language of instruction (χ^2 = 5.85, p = .015). Finally, the congruency effect was somewhat greater on switch trials than stay trials, a marginal interaction between language switching and language of instruction congruency (χ^2 = 2.96, p = .085).

In the error rates, bilinguals made more errors naming pictures in mixed blocks than Hebrew-only blocks (χ^2 = 22.44, p < .001). In the congruency model, and in contrast to the response times, bilinguals were somewhat more likely to name pictures incorrectly when the language-of-instruction was congruent than incongruent with the mixing language, though this effect was only marginally significant (χ^2 = 3.60, p = .058). There were no other significant effects in the congruency model (ps > .15).

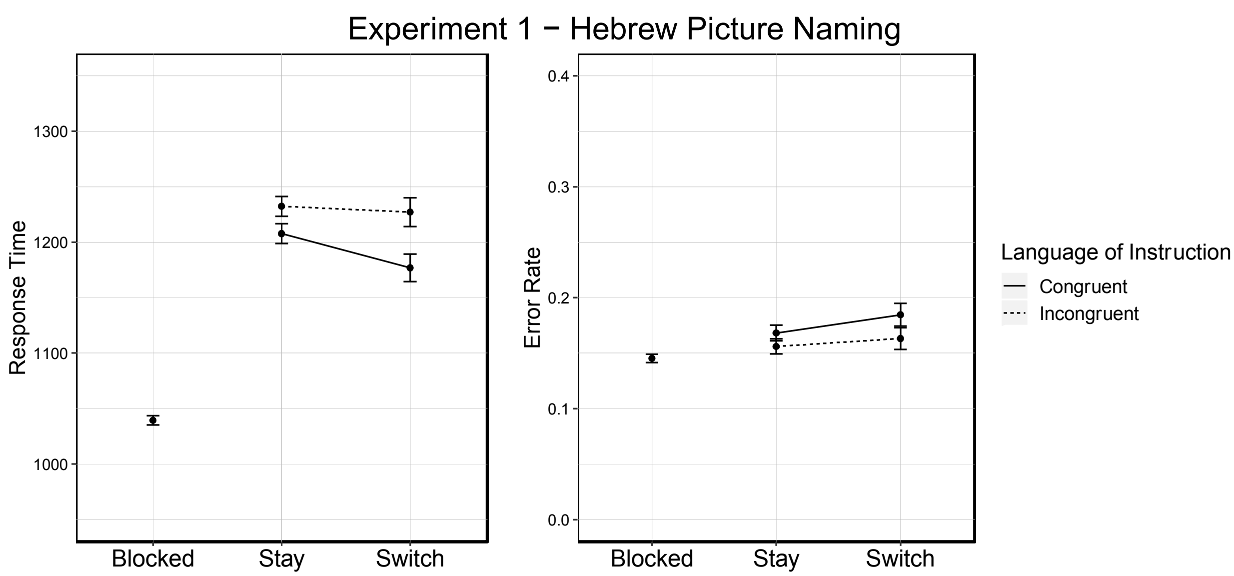

Figure 2.1. *Response times and error rates from Experiment 1 grouped by switching condition and language of instruction. Error bars represent standard errors.*

Figure 2.2 shows the data separated by mixing language (as a function of language dominance) and language of instruction. In the dominance model, considering trial type (stay vs. switch) and language-of-instruction (dominant LOI vs. non-dominant LOI) in this way also shows that switch trials were named significantly faster than stay trials ($\chi^2 = 11.43$, $p < .001$), that mixing Hebrew with the dominant language resulted in faster Hebrew naming latencies than when mixed with the non-dominant language ($\chi^2 = 17.20$, $p < .001$), and pictures were named fastest when learned via the dominant language and mixed with the dominant language, an interaction between mixing language and language of instruction ($\chi^2 = 7.57$, $p = .005$). This effect was slightly stronger on switch trials than stay trials, resulting in a marginal three-way interaction ($\chi^2 = 2.90$, $p = .084$). All other interactions were non-significant (*p*s > .22).

When looking at the error rates in the dominance model, there were three marginally significant effects. First bilinguals made slightly more errors naming Hebrew pictures when the language of instruction was the dominant language ($\chi^2 = 2.72$, $p = .099$). Second, bilinguals made more errors on switch trials as compared to stay trials when mixing with the dominant language, an interaction between language switching and mixing language ($\chi^2 = 3.24$, $p = .073$). Finally, bilinguals made more errors when naming in the dominant language as compared to the non-dominant language when the language of instruction was the dominant language, and vice versa, a marginally significant interaction between language dominance and mixing language ($\chi^2 = 3.22$, $p = .072$). All other effects were non-significant in the error rates (*p*s > .14).

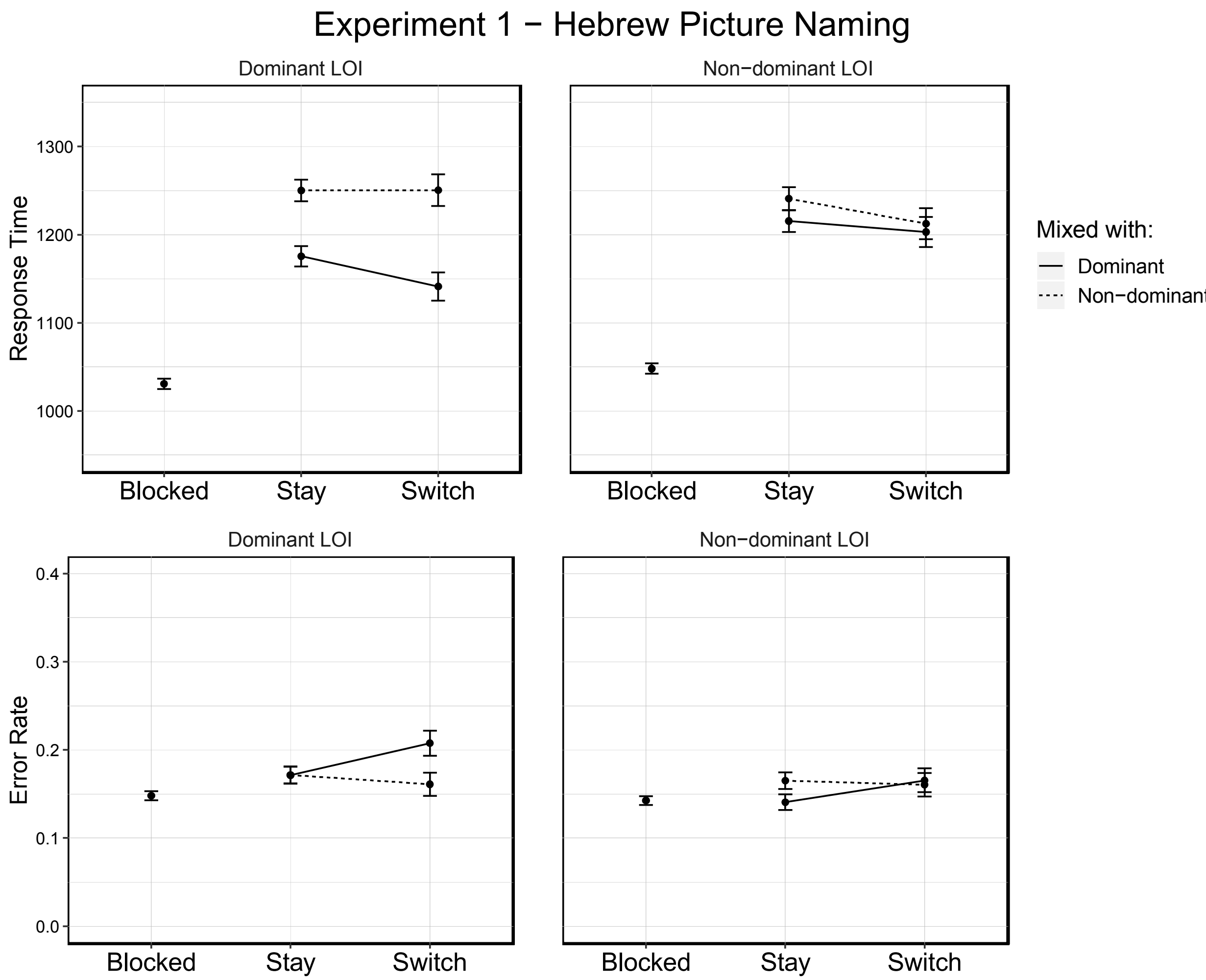

Figure 2.2. *Response times and error rates from Experiment 1 grouped by language, switching condition and mixing language. Error bars represent standard errors.*

Table 2.2 shows the dominant and non-dominant (i.e. Spanish and English) rather than the Hebrew naming times separated by congruency and trial type (stay vs. switch). Though not the focus of these analyses, we report the response time outcomes for these data. First, bilinguals named switch trials slower then stay trials, a significant switch cost (χ^2 = 14.27, $p < .001$). Next, bilinguals named non-dominant language trials slower than dominant language trials (χ^2 = 5.13, $p < .024$). Finally, bilinguals named pictures slower when the language was inconsistent with the language of instruction (χ^2 = 7.97, $p = .005$). None of the interactions were significant (*p*s > .24).

Table 2.2. *Mean and standard deviations of dominant and non-dominant language response times separated by trial type (stay vs. switch) and whether it was the congruent or incongruent mixing language.*

	Dominant		Non-dominant	
	Stay	Switch	Stay	Switch
Congruent	855 (295)	950 (355)	917 (340)	984 (364)
Incongruent	939 (350)	1023 (395)	973 (360)	1039 (401)

These results demonstrate a language-of-instruction effect in a language switching paradigm. In the mixed block, Hebrew response times were faster when the language of instruction was congruent with the language from the mixed language block. Figure 2.2 shows that this result is primarily driven by the mixed block in which the language of instruction was the dominant language, as well as the mixing language. The error rates, however, suggest that bilinguals are taking different strategies in different language-of-instruction conditions. Bilinguals make slightly more errors when the language of instruction was congruent than when it was incongruent, the opposite pattern seen in the response times.

There are two possible explanations when considering this speed-accuracy tradeoff. It may be that the language-of-instruction manipulation causes only this task strategy difference. It may also be that the manipulation causes both a language-of-instruction effect and a speed-accuracy tradeoff. To try to understand this individual contribution of these effects, we conducted a post-hoc analysis. The subjects were median split by overall error rate, and each group was considered separately. The mean error rate for the entire sample was 15.92% (*SD* = 13.96% *, Range* = 1.02% , 59.6%). Thirty-one subjects were considered the high-performing group and had a mean error rate of 5.27% (*SD* = 3.85%, *Range* = 1.02% ,13.10%), and thirty subjects were considered the low-performing group the mean error rate was 27.66% (*SD* =11.36% *, Range* = 13.11% , 59.60%). We performed the analyses again on both groups.

Figure 2.3 shows the data split by dominance for the high-performing group of subjects. In the congruency model, for response times, we found that bilinguals named switch trials significantly faster than stay trials ($\chi^2 = 4.57$, $p = .032$). Additionally, trials mixed with the congruent language of instruction were named significantly faster than trials mixed with the incongruent language of instruction ($\chi^2 = 6.54$, $p = .011$). For errors, neither the interaction, nor any effects in the congruency model were significant (*p*s > .18).

In the dominance model, switch trials were named significantly faster than stay trials ($\chi^2 = 6.15$, $p = .013$), mixing Hebrew with the dominant language resulted in faster Hebrew naming latencies than when mixed with the non-dominant language ($\chi^2 = 10.39$, $p = .001$), and pictures were named fastest when learned via the dominant language and mixed with the dominant language, an interaction between mixing language and language of instruction ($\chi^2 = 9.73$, $p = .002$). All other interactions were non-significant (*p*s > .28). In the error rates of the dominance model, bilinguals made slightly more errors when the language of instruction was dominant, as compared to non-dominant, though this effect was only marginally significant ($\chi^2 = 2.73$, $p = .098$). No other effects or interactions were significant (*p*s > .15).

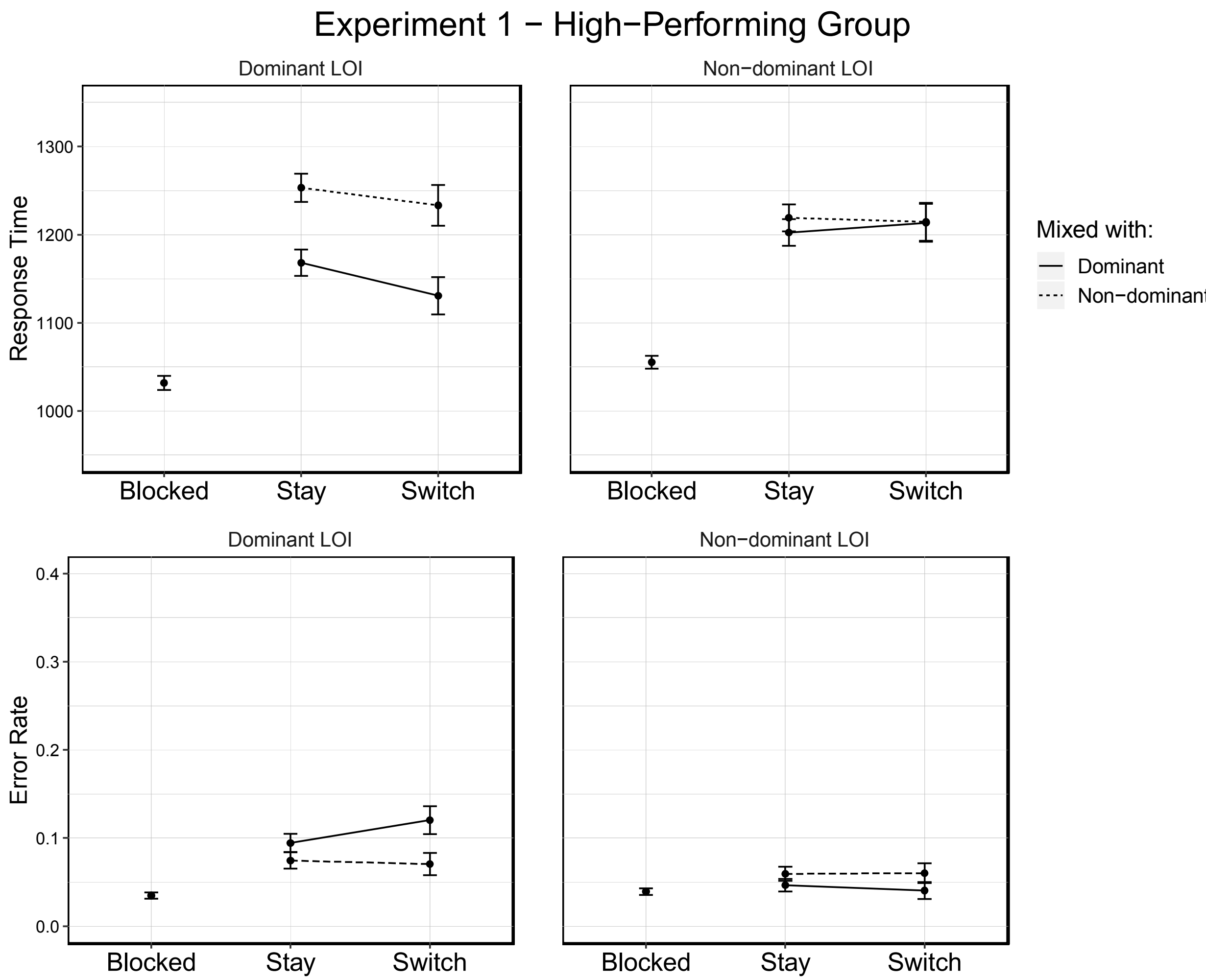

Figure 2.3. *Response times and error rates from the high-performing group in Experiment 1 grouped by language, switching condition and mixing language. Error bars represent standard errors.*

Figure 2.4 shows the data split by dominance for the low-performing group of subjects. For response times, in the congruency model we found only that the congruency effect differed slightly by trial type, a marginal interaction (χ^2 = 3.56, *p* = .059). Neither main effect was significant (*p*s > .15). In the congruency model for the error rates, bilinguals made more errors on switch trials than stay (χ^2 = 3.98, *p* = .046).

In the response times of the dominance model, switch trials were named significantly faster than stay trials (χ^2 = 3.98, *p* = .046) and mixing Hebrew with the dominant language resulted in faster Hebrew naming latencies than when mixed with the non-dominant language (χ^2 = 7.90, p = .005). All other interactions in the response times were non-significant (ps > .18). In the error rates of the dominance model, bilinguals made more errors on switch trials than stay trials (χ^2 = 3.94, p = .047). Additionally, switch costs were slightly greater when mixed with the dominant language than the non-dominant (χ^2 = 4.45, p = .035). No other effects or interactions were significant (ps > .21).

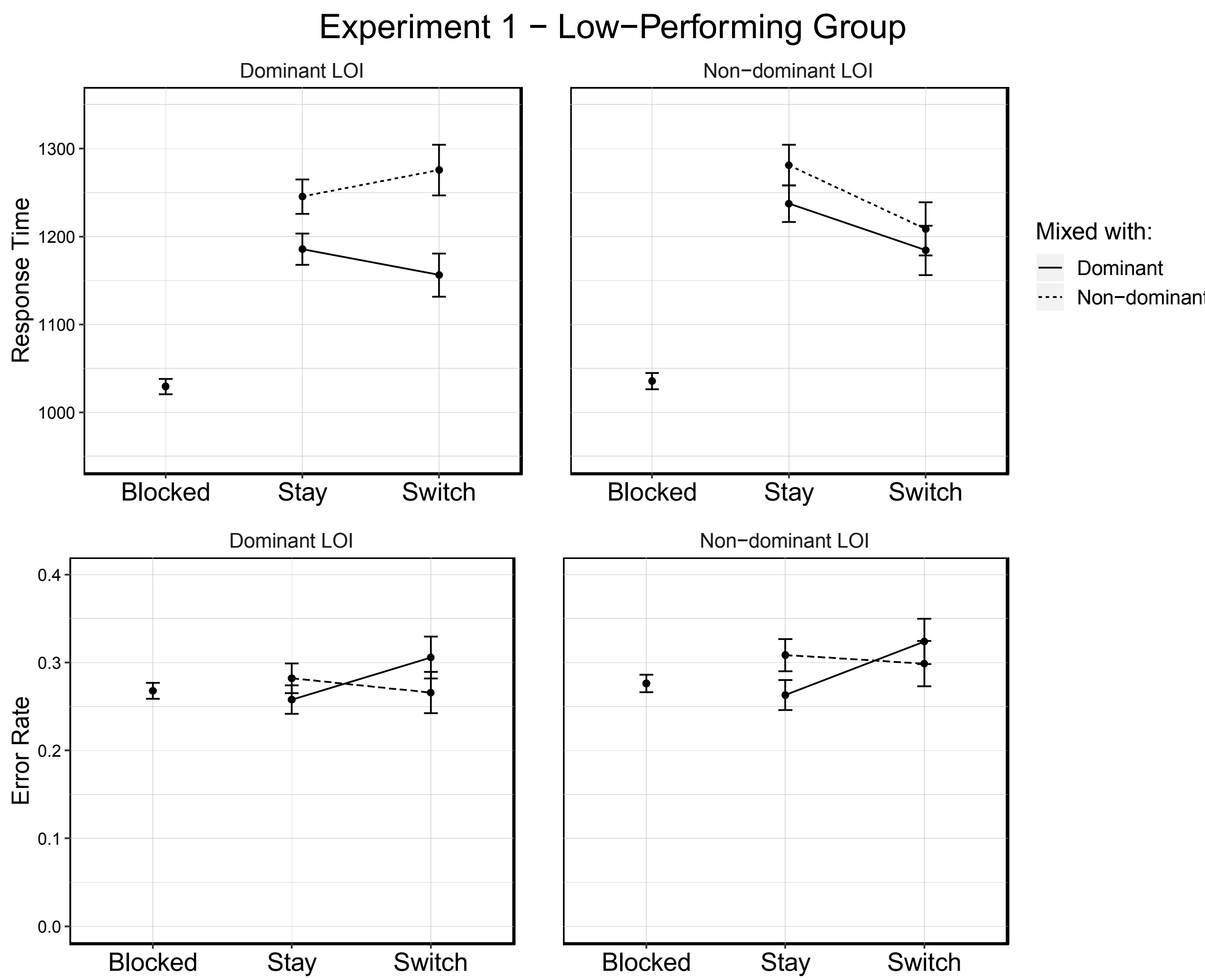

Figure 2.4. *Response times and error rates from the low-performing group in Experiment 1 grouped by language, switching condition and mixing language. Error bars represent standard errors.*

In this experiment, the effects demonstrating language of instruction effects remained robust in the response times of the high-performing group, while effects in the error rates strongly diminished. Because the patterns hold for the subjects in the high-performing subject group in the response times but not error rates, we believe that the speed-accuracy tradeoff is tied to the difficulty of the task, more so than the language-of-instruction effect.

Taken together, these results inform the first two of the questions posed in the introduction. First, the language-of-instruction effect shows that the effect is tied to instructional material, and not simply overall learning experience, as these effects are found within-subject. Second, the presence of the language-of-instruction effect in a language-switching task suggests that language of instruction can indeed impact some level of language-wide control (as switching is never done on the same picture). Interestingly, we saw in both models an overall switch facilitation (here and in Experiment 2), discussed further in the General Discussion.

Interestingly, we see that the language-of-instruction effect is driven primarily by when the language of instruction and the mixing language are both dominant. Picture naming is hastened relatively to nearly every other condition when bilinguals learned from their dominant language and also mixed with that language. Tomoschuk et al. (submitted) noted a different effect in their second study. In that experiment, language-of-instruction effects manifested primarily in the non-dominant language. Bilinguals monitored phonemes in their new third language and were more likely to make errors when the to-be-monitored phoneme was from their dominant language when they learned from their dominant. Put simply, the effects of language of instruction manifested primarily in the dominant language of this experiment, while they manifested primarily in the non-dominant language of Tomoschuk et al. (submitted, see General Discussion for further comments).

Experiment 2

Though Experiment 1 provided considerable evidence furthering our understanding of the language of instruction effect, two outstanding questions still remain. First, is the language-of-instruction effect tied to immediate test, or does it persist over time? There is some evidence that recent exposure to one of a bilingual's languages can facilitate access to that language when subsequently processing the other language (Elston-Güttler, Gunter & Kotz, 2005). It may be that recent exposure to the language of instruction causes faster overall naming latencies in that language, as well as in Hebrew, suggesting that language of instruction effects are due to recent exposure, rather than connections formed between the language of instruction and L3. Second, it is not fully clear whether language-of-instruction manifests as a speed-accuracy tradeoff, or as an effect independent of task strategy.

To examine both of these possibilities, we repeated Experiment 1 with the sole change that Hebrew-only naming blocks came before mixed-language blocks. The rearrangement of these blocks served two purposes. First, it increased the lag between learning and mixed-language naming to reduce facilitated access to the language of instruction, perhaps improving overall accuracy on Hebrew picture naming, and reducing the benefit of a speed-accuracy tradeoff. Second, bilinguals gained further practice with the L3 that is not intermixed with one of their known languages. If language-of-instruction effects persist in this condition, we have further evidence of the language-of-instruction effect being caused by connections formed between languages in acquisition, rather than more transient factors related to ease of access of one of the two known languages.

Method

Subjects. Spanish-English bilingual students (N = 64) taken from the same pool of subjects as Experiment 1 participated for class credit with the additional requirement that they had not participated in Experiment 1. Four subjects were removed before analysis for incorrectly following directions. Three subjects were considered Spanish dominant based on their MINT scores. Subject characteristics are shown in Table 2.1.

Materials, procedure, analysis. The materials, procedure and analysis were identical to Experiment 1 with the exception that during the testing phase subjects completed blocks of Hebrew-only naming before completing blocks mixed with either English or Spanish. The order of these mixed blocks was still counterbalanced between subject and language of instruction condition.

Results and Discussion

In the evaluative learning phase, subjects produced 75.2% (*SD* = 24.0%) of Hebrew words correctly when the language of instruction was English, and 72.5% (*SD* = 25.3%) correctly when the language of instruction was Spanish. This difference was not significant ($t<1$).

Figure 2.5 shows Hebrew naming latencies and response times by block and language congruency. Looking first at response times, in the first model, we see that Hebrew pictures were named slower when mixed with English or Spanish trials ($\chi^2 = 22.33$, $p < .001$). When looking at the congruency model, for response times, switch trials were named faster than stay trials ($\chi^2 = 11.32$, $p < .001$) and pictures were named faster in Hebrew when mixed with the language that was congruent to their language of instruction, as compared to when mixed with their incongruent language of instruction ($\chi^2 = 10.36$, $p = .013$). There was no significant interaction between these two factors ($\chi^2 = 1.35$, $p = .25$).

In the error rates, for the congruency model, bilinguals made more errors naming pictures in mixed blocks than Hebrew-only blocks (χ^2 = 8.48, p = .004). In the second model, bilinguals were more likely to name pictures incorrectly when the language-of-instruction was congruent than incongruent with the mixing language (χ^2 = 6.10, p = .014). There were no other significant effects looking only at mixed blocks in the second model (ps > .19).

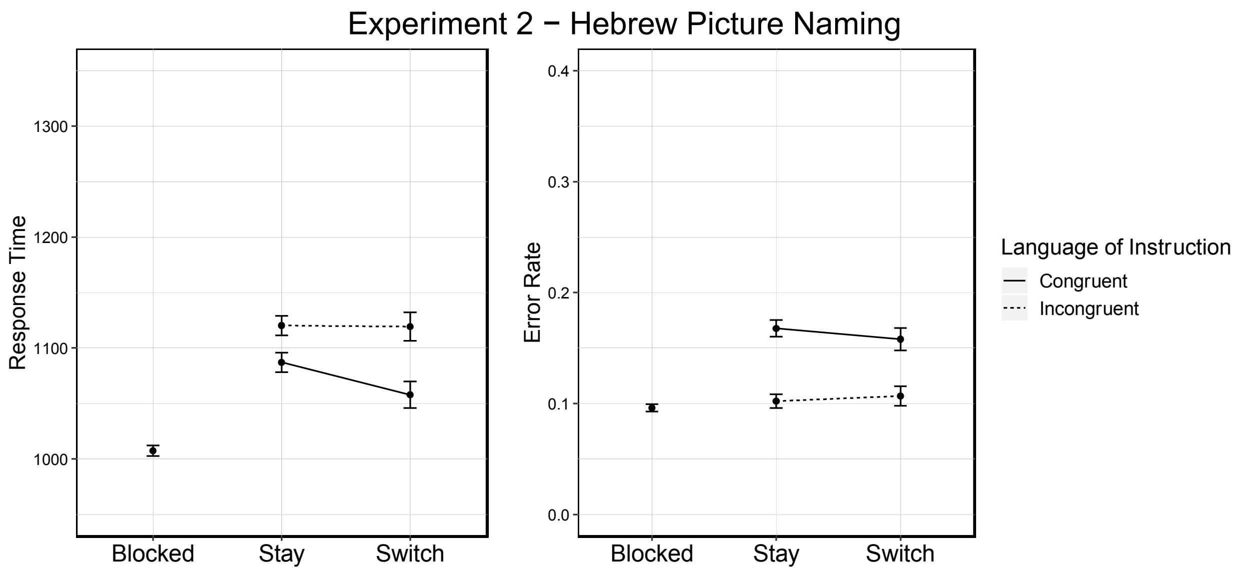

Figure 2.5. *Response times and error rates from Experiment 2 grouped by switching condition and language of instruction. Error bars represent standard errors.*

Figure 2.6 shows the data separated by mixing language and language of instruction. For the dominance model, switch trials were named faster than stay trials (χ^2 = 10.59, p = .001), pictures were named overall faster in Hebrew when mixed with the dominant language than with non-dominant (χ^2 = 27.61, p < .001), and finally the difference between response times between the mixing language conditions was greater when the language of instruction was the dominant language than the non-dominant (χ^2 = 9.64, p = .002). All other interactions were non-significant (ps > .21).

When looking at the error rates in dominance model, when the language of instruction was the dominant language, bilinguals made more errors when Hebrew naming was mixed with the dominant language as compared to the non-dominant language, but when the language of instruction was the non-dominant language, they made more errors when Hebrew naming was mixed with the non-dominant language than the dominant language, an interaction between language dominance and mixing language ($\chi^2 = 6.55$, $p = .01$). All other effects were non-significant (*p*s > .20).

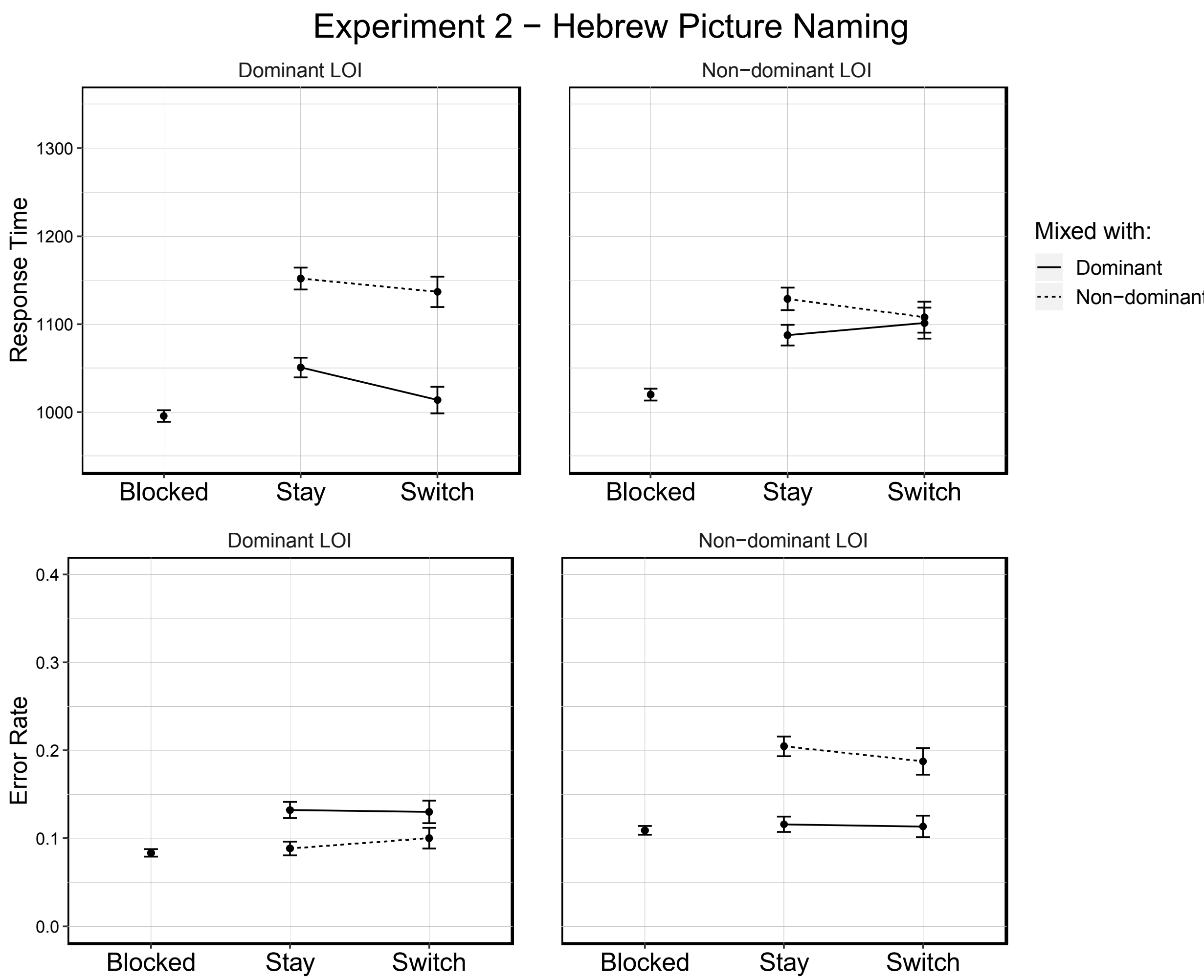

Figure 2.6. *Response times and error rates from Experiment 2 grouped by language, switching condition and mixing language. Error bars represent standard errors.*

Table 2.3 shows the dominant and non-dominant (i.e. English and Spanish) language data separated by congruency and trial type (stay vs. switch). Though not the focus of these analyses, we report the response time outcomes for these data. First, switch trials were named slower then stay trials ($\chi^2 = 22.04$, $p < .001$). Next, non-dominant language trials were named slower than dominant language trials ($\chi^2 = 10.82$, $p = .001$). Pictures were named slower when they were inconsistent with the language of instruction ($\chi^2 = 25.95$, $p < .001$). Switch costs were significantly larger in the dominant language as compared to the non-dominant language ($\chi^2 = 8.23$, $p = .004$). Finally, the congruency effect was significantly greater in the dominant language ($\chi^2 = 20.73$, $p < .001$).

Table 2.3. *Mean and standard deviations of dominant and non-dominant language response times separated by trial type (stay vs. switch) and whether it was the congruent or incongruent mixing language.*

	Dominant		Non-dominant	
	Stay	Switch	Stay	Switch
Congruent	856 (350)	909 (367)	941 (374)	974 (378)
Incongruent	934 (393)	983 (409)	981 (389)	1011 (431)

As with Experiment 1, to assess concerns with possible speed-accuracy trade-offs, we furthermore conducted the same post-hoc analysis as in Experiment 1. The mean error rate for the entire sample was 11.87% (*SD* = 8.96%*, Range* = 1.28%, 37.67%). Thirty subjects were considered the high-performing group and had a mean error rate of 4.67% (*SD* = 3.51%, *Range* = 1.28% ,10.45%), and thirty subjects were considered the low-performing group the mean error rate was 17.17% (*SD* =5.35% *, Range* = 11.05% , 37.67%). We performed the analyses again on both groups.

Though we analyze the data using both the congruency and dominance models, Figure 2.7 shows the data split by dominance for the high-performing group of subjects. The analysis shows for the congruency model, for response times, we found that bilinguals named switch trials significantly faster than stay trials ($\chi^2 = 11.22$, $p < .001$). Additionally, trials mixed with the congruent language of instruction were named significantly faster than trials mixed with the incongruent language of instruction ($\chi^2 = 8.35$, $p = .004$). No effects in the congruency model of the errors were significant (*p*s > .19).

In the dominance model, for response times, switch trials were named significantly faster than stay trials ($\chi^2 = 10.93$, $p < .001$), mixing Hebrew with the dominant language resulted in faster Hebrew naming latencies than mixing with the non-dominant language ($\chi^2 = 11.90$, $p < .001$), and pictures were named fastest when learned via the dominant language and mixed with the dominant language, an interaction between mixing language and language of instruction ($\chi^2 = 7.94$, $p = .004$). All other interactions were non-significant (*p*s > .22). In the error rates of the dominance model, all effects and interactions were not significant (*p*s > .13).

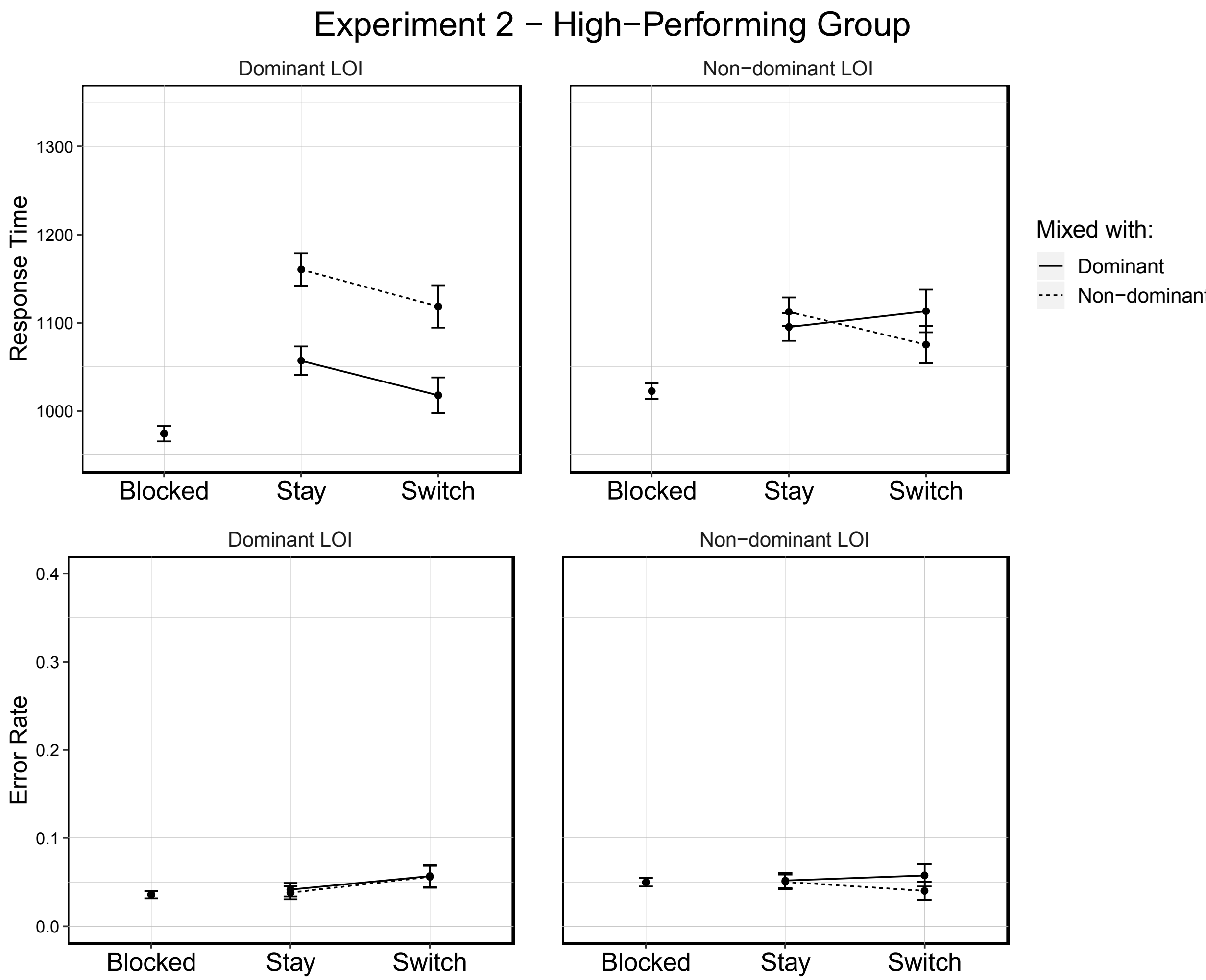

Figure 2.7. *Response times and error rates for the high-performing group from Experiment 2 grouped by language, switching condition and mixing language. Error bars represent standard errors.*

Figure 2.8 shows the data split by dominance for the low-performing group of subjects. For response times, in the congruency model we found no significant effects (*p*s > .13). In the congruency model for the error rates, bilinguals made more errors on when the language of instruction was congruent, as compared to incongruent (χ^2 = 7.85, p = .005).

In the response times of the dominance model, mixing Hebrew with the dominant language resulted in faster Hebrew naming latencies than when mixed with the non-dominant language (χ^2 = 15.44, $p < .001$). All other interactions in the response times were non-significant (*p*s > .17). In the error rates of the dominance model, bilinguals made more errors mixing with the dominant language when the language of instruction was the dominant language and more when mixing with the non-dominant language when the language of instruction was the non-dominant language, an interaction between mixing language and language of instruction ($\chi^2 = 7.97$, $p = .005$). No other effects or interactions were significant (*p*s > .14).

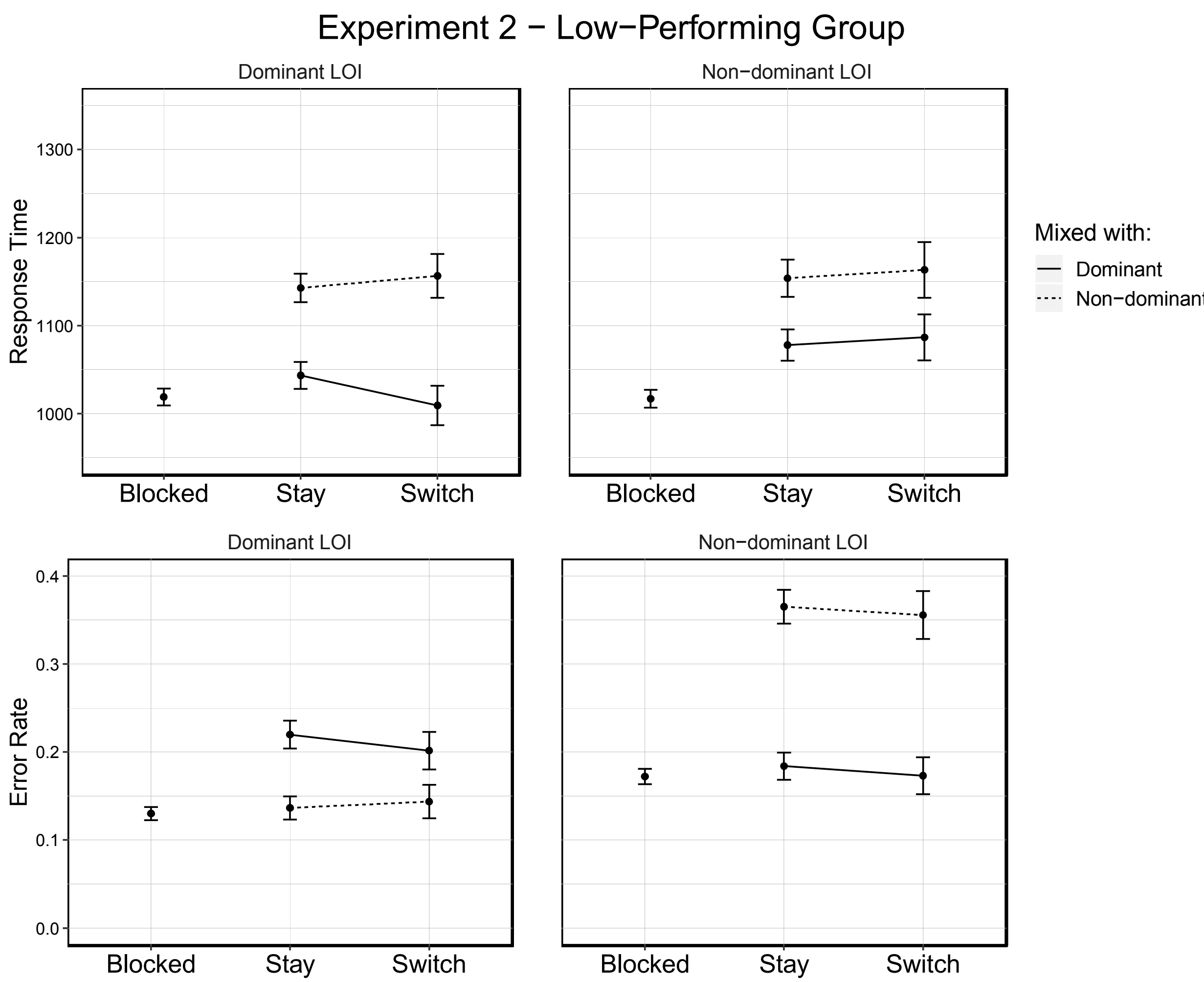

Figure 2.8. *Response times and error rates for the low-performing group from Experiment 2 grouped by language, switching condition and mixing language. Error bars represent standard errors.*

As in Experiment 1, language of instruction effects remained robust when considering only the high-performing subjects, while no differences were seen in the error rates. In the low performing group, error rates were higher for the incongruent language of instruction, suggesting that the speed-accuracy tradeoff seen in the whole sample is driven by low-performing subjects.

In Experiment 2, we expected that if recent exposure to one of the languages of instruction drives these response-time differences, that having bilinguals perform Hebrew-only blocks before the mixed blocks should reduce, if not eliminate language-of-instruction effects. We saw that between experiments the overall response times of the mixed language blocks decreased ($\chi^2 = 29.12$, $p < .001$), as well as the overall error rates ($\chi^2 = 4.74$, $p = .029$), suggesting that Hebrew-only practice did improve performance on the mixed language tasks. Additionally, numerically we do indeed see the congruency difference diminish in stay trials.

General Discussion

In two experiments, we show that first, language of instruction impacts language control when learning a third language in a switching task, and second that the language-of-instruction effect is persistent over a block of L3-only practice. In Experiment 1, Spanish-English bilinguals learned Hebrew words from both of their languages, naming Hebrew pictures in blocks mixed with either English or Spanish naming, followed by Hebrew-only naming. When the language that Hebrew picture naming was mixed with was congruent with the language of instruction, pictures were named faster than when incongruent, though this effect is driven primarily by the dominant language. In Experiment 2, we further tested the presence of this learning effect by having subjects name pictures in Hebrew only before mixed blocks, giving them additional non-mixed experience. Here we found that main effects remain for both switching and congruency.

These results are accompanied by a speed-accuracy trade off. Both experiments showed that mixing with the congruent language of instruction hastens language production, as well as leads to more errors relative to when the language of instruction is incongruent. Though we see the effect in error rates diminish when analyzing the data from only the high-performing group, the tradeoff exists in both experiment groups (before splitting to high- and low- performing groups). One possibility is that, in additional to task difficulty, manipulating language of instruction modulates a bilingual's language control strategy, causing them to favor making more mistakes but name pictures faster. Because subjects with overall fewer errors do not show this tradeoff, future studies may consider paradigms with higher inter-stimulus intervals, or increased L3 training, to boost overall accuracy rates.

Speed-accuracy tradeoffs acknowledged, these studies address four points outlined in the introduction. First, we asked whether the effect was tied to instructional material or a bilingual's learning experience more generally. The replication of language-of-instruction effects found in Tomoschuk et al. (submitted) in a within-subjects design suggest that language of instruction is indeed tied to the instructional material, and that such effects can vary within a bilingual based on the learning experience.

Second, we asked if the effect modulates learning at the lexical or language-wide level. In Tomoschuk et al. (submitted), subjects showed language interference modulated by language of instruction in a phoneme monitoring task. Such a task suggests that the effect could exist at the lexical level, or instead that the effect operates at the whole-language level as may be suggested by Branzi et al (2014). The task in this experiment however is language switching. Critically, in these experiments, subjects never named the same picture twice in a row, switching only between translations. Instead, switch trials necessarily involved switching languages as well as to-be-named pictures. If switching only between translations, it may be argued that language-of-instruction effects modulate language only at the lexical level, as costs could be the result of switching between lexical items, not semantic concepts. Instead, bilinguals switched languages (on some trials) and to-be-named pictures (on all trials). In such a task, response time differences must be modulated at least partially at a whole-language level. That is, the type of language control modulated by language of instruction must exist at least at the whole-language level, if not both levels.

Third, we asked whether the language-of-instruction effect persists past immediate test. For example, it may be that simply having additional exposure to English or Spanish during training facilitates naming overall suggesting that language of instruction effects are explained by the recent exposure to one of the languages of instruction. As such, we gave subjects Hebrew-only practice to increase the interval between the learning trials and the mixed block. After this change, the effect of language of instruction remained, primarily in switch conditions, suggesting that while the effect may be influenced by recent exposure to the language of instruction, there remains some effect whereby tighter connections are formed between the language of instruction and new L3 in acquisition, facilitating switching between these languages.

Fourth, we asked how language dominance plays into language of instruction. In both studies, language of instruction facilitates bilingual lexical access in a dual language task, suggesting its importance as a factor of L3 acquisition. Tomoschuk et al. (submitted), however, showed that the dominant L1 did not interfere more with the non-dominant L2 even when the L2 was the language of instruction, suggesting that, orthogonal to language-of-instruction effects, the dominant language is less prone to interference of any kind in a phoneme monitoring context. Here, in a language switching context, the opposite was found. As shown in Figures 2.2 and 2.4, the new L3 is affected more by language of instruction in the presence of the dominant language. Post-hoc analyses shows that when considering only one language at a time, the language of instruction effect was significant for the dominant language (Experiment 1: $\chi^2 = 8.08$, $p = .004$, Experiment 2: $\chi^2 = 7.15$, $p = .008$) but non-significant for the non-dominant language (Experiment 1: $\chi^2 = 0.35$, $p = .55$, Experiment 2: $\chi^2 = 0.96$, $p = .33$). This suggests that it is difficult to mix with the non-dominant language regardless of language of instruction condition, and that mixing with the dominant language facilitates naming in the L3 when learned from that language. In both tasks, the dominant language is easier to control than the non-dominant language.

What, then, causes language-of-instruction effects to manifest in the dominant language of one task but the non-dominant language of another? Tomoschuk et al. (submitted) claims that because phoneme monitoring is an interference task, it is difficult for bilinguals dominant and immersed in their L1 to show interference patterns while monitoring in that dominant language. They are so good at producing in the dominant language that external linguistic information that is not a part of the current task is less likely to affect production. Here, in a language switching task, bilinguals are not producing in only one language at a time, unilaterally inhibiting all language information that is not pertinent to the task at hand. They must instead actively produce in both task languages. Instead of introducing difficulty into a relatively easy task, we may be introducing a means of facilitation into a relatively difficult task. When bilinguals learn from their dominant language, they can use the ease of their dominant language to hasten production across that entire block – all other blocks and tasks are hard because they need the alignment of both language of instruction and mixing language to facilitate production.

Interestingly, in both experiments, we anticipated language of instruction effects to manifest in switch costs. Instead, we found that subjects produce an L3 Hebrew item faster when the item before it was in the L1 or L2 as compared to when the previous item was the L3 – that is, a switch benefit rather than switch cost – and this seems particularly true for the congruent language-of-instruction condition in the second experiment. One might expect this to be a result of accuracy. Because speakers of this L3 do not know the L3 words as well as the L1 or L2 words, they may be more likely to not produce anything rather than correctly produce an item with a longer latency, thus causing response times to "speed up" on harder trials because there are fewer long responses being factored in. Such a speed-accuracy tradeoff was found in Tomoschuk et al. (submitted). However, such difficulty should affect stay and switch trials alike. Furthermore, when looking at naming trials for pictures that subjects correctly produced by the end of the learning phase, none of the results change. It is still easier to switch from the L1 or L2 into the L3 Hebrew than produce a second L3 item correctly. One possible explanation is that the difficulty of switch trials manifests in error rates as opposed to response times, as suggested by the marginally significant switch cost seen in the error rates of Experiment 2. This is, however, the only place we see higher errors on switch trials compared to stay trials. Another possibility is that this switch facilitation is caused simply by the difficulty of producing L3 items. In Experiment 1, the average response time in mixed blocks for Hebrew (1258 ms) is considerably higher than that of Spanish (978 ms) and English (933 ms). The difficulty of first naming a picture in Hebrew may carry over into the next Hebrew naming trial.

These experiments support the findings of Tomoschuk et al. (submitted) in suggesting that learning experiences modulate language control; learners of a new language scaffold that language onto previous language experience. Such an effect has been demonstrated in two populations and two language control tasks. Additionally, though, the two sets of studies in conjunction suggest that language of instruction effects are tied to instructional material, and not increased experience, that they modulate control mechanisms at a whole-language level, and that they persists even after additional L3 exposure. Learning regulatory control is an important part of the language learning process, and is modulated by the specifics of our learning experiences.

Chapter 2, in full, is a reprint of the material to be submitted to the Journal of Memory and Language. Tomoschuk, Brendan; Gollan, Tamar H.; Ferreira, Victor S. The dissertation author was the primary investigator and author of this paper.

Appendix 2.A *Items used in Experiments 1 and 2.*

English	Spanish	Hebrew
fork	tenedor	mazleg
queen	reina	malka
teeth	dientes	shinayim
bear	oso	dov
glass	vaso	kos
moustache	bigote	safam
hat	gorro	kova
tie	corbata	aniva
spoon	cuchara	kapit
fish	pez	dag
chicken	pollo	ohf
hammer	martillo	patish
straw	popote	kash
broom	escoba	matate
purse	bolsa	teek
eye	ojo	ayin

Appendix 2.B. *Residual rates analysis*

To align the results with Chapter 1, Figure 2.9 shows the data from Experiments 1 and 2 calculated as adjusted residual rates. Adjusted residual rates were calculated based on Hughes, Linck, Bowles, Koeth, and Bunting (2013). An accuracy score was calculated for each condition and participant, along with the average RT for that condition. The RT was then converted to minutes and the accuracy was divided by this value to generate an adjusted residual rate score. This metric reflects the average number of correct responses per minute per condition (see Chapter 1 for example calculation), as such a higher residual rate indicates better performance. In Experiment 1, residual rates were significantly higher when naming was blocked than mixed (χ^2 = 65.84, p < .001), but there were no effects of switching or congruency (ps > .52). Likewise in Experiment 2, residual rates were significantly higher when naming was blocked than mixed (χ^2 = 29.95, p < .001), and no effects of switching or congruency (ps > .19).

That no effects appear in this analysis could represent multiple things. First, it is possible that the speed-accuracy tradeoff accounts for more of the effect seen in response times than we believe. Alternatively, by grouping trials within-subject we reduce the number of critical trials and the variance of the sample shrinks, reducing the effect size that the statistical analysis can reliably detect. In either case, considerably more subjects would be necessary to detect a reliable difference using residual rates.

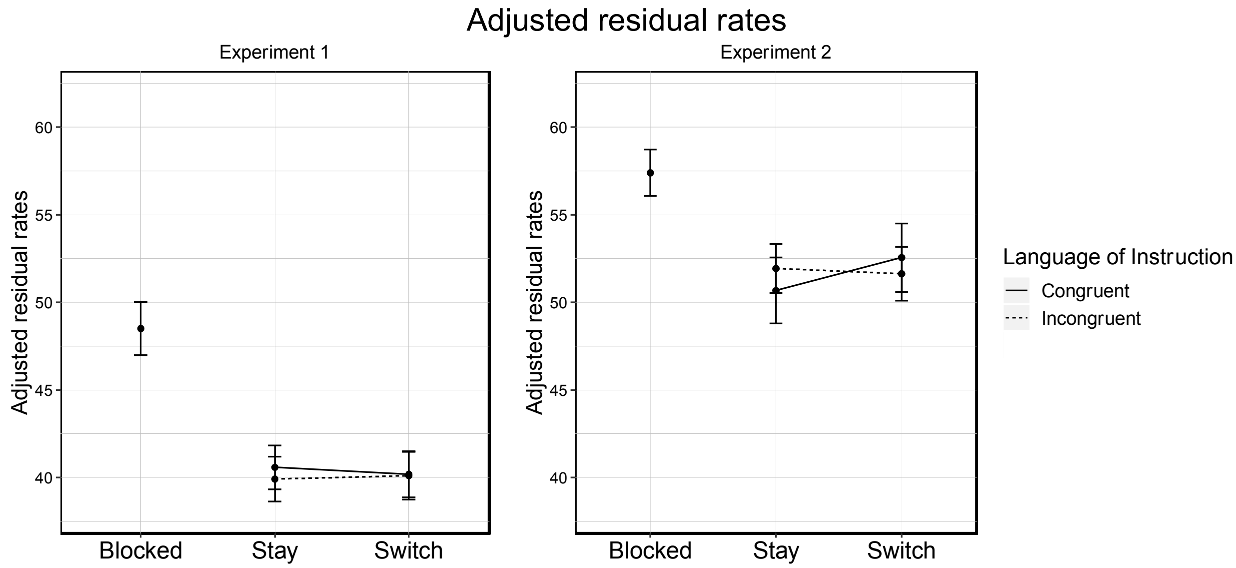

Figure 2.9. *Adjusted residual rates for Experiments 1 and 2. Error bars represent standard errors.*

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Branzi, F. M., Martin, C. D., Abutalebi, J., & Costa, A. (2014). The after-effects of bilingual language production. *Neuropsychologia*, *52*, 102-116.

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it?. *Cognitive science*, *41*(6), 1629-1644.

De Bot, K. (2000). A bilingual production model: Levelt's "speaking" model adapted. *The bilingualism reader*, 420-442.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(3), 175-197.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research*, *25*(1), 57-70.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *15*(3), 594-615.

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior research methods*, *46*(3), 702-721.

Kroll, J. F., & De Groot, A. M. (Eds.). (2009). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.

Kroll, J. F., & Gollan, T. H. (2013). Speech planning in two languages. In *The Oxford Handbook of Language Production*.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of memory and language*, *33*(2), 149-174.

Meisel, J. M. (1983). Transfer as a second-language strategy. *Language & communication*, *3*(1), 11-46.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Runnqvist, E., Strijkers, K., & Costa, A. (2014). 11 Bilingual Word Access. *The Oxford handbook of language production*, 182.

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

Tomoschuk, B., Duyck, W., Hartsuiker, R., Ferreira, V.S., Gollan, T.H. (submitted). Language of Instruction Impacts Language Control in the Third Language. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2018). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, 1-21.

Williams, S., & Hammarberg, B. (1998). Language switches in L3 production: Implications for a polyglot speaking model. *Applied linguistics*, *19*(3), 295-333.

CHAPTER THREE

Translation Distractors Do Much More Than Just Tell You "What Not to Say": Language Switching in a Picture Word Interference Task

Brendan Tomoschuk[a]

Victor S. Ferreira[a]

Tamar H. Gollan[a]

[a] University of California, San Diego,
9500 Gilman Drive, La Jolla, CA, 92093-0109

Abstract

Evidence from picture-word interference tasks (in which a speaker names a picture with a superimposed distractor word) suggests that bilingual lexical regulation is language-specific, meaning that a bilingual can restrict selection to a single target language (Costa, Miozzo & Caramazza, 1999). Evidence from language switching literature, on the other hand, suggests that bilinguals must inhibit the dominant language to produce the non-dominant language (Meuter & Allport, 1999). The conflict between these literatures may arise, at least in part, from the difference in the nature of the tasks. Picture-word interference paradigms, on the one hand, most often restricted responses to one language (subjects were asked to produce names only in a single target language, see Hall, 2011, for review) while in language switching paradigms bilinguals responded in both languages. In two experiments, we resolve this difference by asking bilinguals to perform a picture-word interference task in both single and mixed language blocks. In Experiment 1, Spanish-English bilinguals named pictures with superimposed distractors while cued to name pictures only in Spanish, only in English, or in both languages (based on a cue). In Experiment 2, a set of Spanish-English bilinguals taken from the same population performed the same task, with the only difference being that all distractors (not simply the translation condition) were presented in the non-target language. Results across experiments showed that translation facilitation effects are highly robust in a language-switching task, regardless of distractor language, and likely not an artifact of task strategy but instead reflect facilitation at the semantic level offsetting competition at the lexical level.

Keywords: picture-word interference, language switching, bilingualism, language control

Translation Distractors Do Much More Than Just Tell You "What Not to Say": Language Switching in a Picture Word Interference Task

Speaking two languages is effortless for many bilinguals despite some of the cognitive challenges that come with it. There is substantial evidence that both languages are active when a bilingual is comprehending (e.g., Dijkstra & Van Heuven, 2002; Thierry & Wu, 2007, see Kroll & De Groot, 2009 for review) and producing (see de Bot, 2000 or Kroll & Gollan, 2013 for a review, though see Costa Pannunzi, Deco, & Pickering, 2017, for a competing account) just one of their two languages. Thus, to speak one of these languages without interference from the other, a bilingual must inhibit non-target language information. While considerable work has been invested towards understanding language selection in bilinguals, many questions remain about the nature of these mechanisms.

The picture-word interference task allows exploration of how lexical items affect production as they become active during speech production in real time. In this task, a speaker names a picture while attempting to ignore distractor words written over the picture (for review see Hall, 2011; Roelofs, 1992; Roelofs, 1998). For example, when a semantically related word appears on the to-be-named picture (e.g., the word *cat* is superimposed on the picture of a dog), naming latencies are slowed relative to an unrelated word (e.g., when the word *table* appears on the picture of a dog). This is referred to as the *semantic interference effect*. Interestingly though, other types of distractor words facilitate production. For example, when a phonologically related word appears on the to-be-named picture (e.g., the word *doll* appears on the picture of a dog), naming latencies are hastened relative to unrelated words, known as the *phonological facilitation effect*.

Bilingual distractor words have provided critical evidence shaping alternative models of lexical access in bilingual speech production. A widely cited paper by Costa, Miozzo and Caramazza (1999) asked if lexical access in bilinguals is language specific or non-specific. In particular, they investigated whether bilinguals can select entirely within an intended target language, and are thereby completely unaffected by the activity of representations in the non-target language. If bilinguals can select in one language entirely independent of activity in the other, there should be no effect of distractors from another language. Catalan-Spanish bilinguals named pictures in Catalan with both target language (Catalan) and non-target language (Spanish) distractors. In addition to finding the same within language interference effects discussed above (semantic interference and phonological facilitation), they also found similar between-language effects when the bilinguals named Catalan pictures with Spanish distractors. For example, naming a picture of a dog in Catalan was slower when the word imposed was semantically related (e.g. *gato,* the Spanish translation of cat) and hastened when the superimposed word was phonetically related (e.g. *dama,* the Spanish translation of lady) relative to an unrelated Spanish word.

However, of great interest, Costa et al. (1999) found *facilitation effects* from translation distractors. For example, bilinguals named a picture of a dog more quickly when *perro* (its translation) was the distractor relative to the unrelated Spanish word. This was a surprising result because if there is competition between languages at the lexical level, then translations should have elicited unusually strong interference effects (Hermans, Bongaerts, De Bot & Schreuder, 1998). Costa et al. (1999) theorized that translation facilitation effects support a language specific model of control in which the non-target language is set aside as a whole; if lexical access were not language specific, translation distractors should have elicited effects similar to semantically related distractors. But in a language specific model, the translation distractors can only boost activation of the intended target at a semantic level, and cannot compete for selection at the lexical level, thereby speeding (rather than interfering with) production times.

This effect was further explored in Roelofs, Piai, Rodriguez and Chwilla (2016). In this study Dutch-English bilinguals completed a picture word interference task in their second language (L2) English in which some distractor items were presented in their first language (L1) Dutch, while EEG was recorded. In addition to replicating the commonly reported semantic interference and phonological facilitation effects on response times, they found a reduction in N400 amplitude (a measure typically interpreted as difficulty in semantic integration, see Kutas and Hillyard, 1980) for both semantically related and translation distractors. The authors claim that a reduction in N400 amplitude suggests that both semantic distractors and translation distractors facilitate access to the name of the picture. Because the facilitation of semantic processing manifests as interference for within-language semantic distractors and facilitation for between-language translation distractors, the data support a language specific model of lexical access where only items from the same language compete for production (see Piai, Roelofs, Jensen, Schoffelen & Bonnefond, 2014 for additional discussion of neurocognitive results).

In most (but not all, see below) picture-word interference studies, however, bilinguals completed the task with responses in just one language. In Roelofs et al. (2016) bilinguals completed the entire experiment in their non-native Dutch. Frequently, however, when languages are mixed in naturalistic settings bilinguals both speak and comprehend the two languages. That is, the controlled situation elicited in the picture-word interference task in which the response set is restricted to just one language while the other language is continuously intermixed in just one modality might not exist outside the laboratory, and could reflect a mode of processing that only arises in such unusual circumstances rather than what bilinguals typically do. In circumstances that involve switching back and forth between languages there is more consistent evidence for competition between languages. Specifically, when bilinguals must switch between languages, they name pictures more slowly than when they speak the same language they used on the previous response (Meuter & Allport, 1999; for review see Declerck & Philipp, 2015). Furthermore, this cost is often asymmetric, that is, switch costs are larger for switching into the dominant than the nondominant languages. To explain the switch cost asymmetry, inhibition of the dominant language in the service of non-target language production is often invoked; switch costs are larger when switching back to the dominant language because it was necessary to have inhibited it more strongly to be able to use the non-dominant language (Green, 1998; Abutalebi & Green, 2007). The notion of inhibitory control presupposes competition between languages, which might be greater when bilinguals switch back and forth between languages, but also functional when bilinguals speak in just one language (i.e., inconsistent with the views developed by Costa et al., 1999 and Roelofs et al., 2016).

Additional evidence for competition between translation equivalents comes from a related effect also from the picture-word interference task. Hermans et al. (1998) investigated cross-language activation in bilingual production by having Dutch-English bilinguals name pictures in their L2 English (e.g., *mountain*) while auditory distractors played in either their L1 Dutch or L2 English. They found that distractors in either language that were phonologically related to the L1 translation of the target word (e.g., *bench* in English or *berm* in Dutch for *berg* the Dutch translation of mountain) caused slowdowns in naming latencies relative to an unrelated distractor. The authors referred to these effects as phono-translation effects, and interpreted these effects as evidence of a lexical level of language control as interference occurred in word production even up to the phonological level. The phono-translation effect has been replicated and shown to increase with proficiency (Boukadi, Davies & Wilson, 2015; Costa, Colomé, Gómez & Sebastián-Galles, 2003; see Klaus, Lemhöfer & Schriefers, 2018, for discussion). Theoretical interpretations of the phono-translation effect on the one hand, and translation facilitation effect on the other in picture-word interference paradigms are completely incompatible. Why should translations facilitate production when presented directly in written form, but inhibit it when auditory distractors that are phonologically similar to the translations are presented?

Lexical access must, then, be modulated by the presentation of the translation itself and other factors including the nature of the task, the experience of the bilingual, and so forth (see Kroll, Bobb & Wodniecka, 2006 for discussion). Some argue that translation facilitation effects are in fact an artifact of a task-specific strategy. The distractor, being the only information provided in the non-target language, provides a "what not to say" cue that eliminates the translation itself as a possible competing item, hastening production of the target word in the correct target language (in Kroll et al., 2006, a what is "not to be spoken" strategy). Similarly, Roelofs et al. (2016) suggested that translation facilitation effects might be found only when the response set contains only one of the two languages. That is, in all previous studies discussed above, bilinguals knew in advance they would never have to produce any words in the non-target language (i.e., they only named pictures in one language while distractor words were sometimes in the same language, and sometimes in the nontarget language). On this view, if bilinguals were tested with the same picture-word interference task but were also cued to switch languages on some trials, translation distractors might then elicit interference rather than facilitation (because both languages would be in the response set).

Only one experiment to date has begun to explore this possibility. In Experiment 2 of Costa et al. (1999), Catalan-Spanish bilinguals named pictures in a mixed language response block. In this block, subjects named 30% of the trials in their non-dominant Spanish. The authors treated these trials as filler, causing several notable differences between this study and a typical language switching task, making it difficult to determine whether language switching might change the nature of translation facilitation. First, the pictures were different from the critical Catalan naming pictures and were only presented four times as compared to the critical trials' six presentations. Second, the authors did not report on any of the effects of the non-dominant language. Third, the Spanish translation that appeared on the critical Catalan naming trials was always a cognate (e.g. the Spanish word *limón* superimposed on a picture of a lemon in a Spanish-English task). Though not a full exploration of translation facilitation in language mixing, the authors show that a robust effect of Spanish translation identity remains in Catalan naming even in the presence of Spanish filler trials.

Here we explored the nature of the translation facilitation effect in a language mixing task. In Experiment 1, Spanish-English bilinguals completed six blocks of picture naming, two blocks in which they produced names only in Spanish, two in which they produced picture names only in English (*single language response blocks*) and two intermixed blocks in which they were cued to name pictures in English on half the trials, and in Spanish on the other half, or *mixed language response blocks*. In each block words were superimposed over the picture in one of four conditions, mimicking the experimental conditions of Roelofs et al. (2016): control (a row of XXXXs), same-language semantically related (the word *nose* superimposed on a picture of an eye), same-language semantically unrelated (the word *napkin* superimposed on a picture of an eye), and translation equivalent (the word *ojo* superimposed on the picture of an eye). Several key differences exist between this and Costa et al (1999). First, all distractors were non-cognate translations. In order to test language-wide inhibition we test translation facilitation between non-cognate items, as cognates facilitate naming (for example see de Groot & Nas, 1991). Second, both Spanish and English picture naming are considered critical trials. This allows us to understand the impact of mixing on translation facilitation in the non-dominant language. Third, the switching rate was increased to 50%. Finally, language mixing was manipulated within-subject to ensure high statistical power in detecting translation facilitation differences between block type.

If translation facilitation in single-language response blocks is caused by a “what not to say” strategy, translation facilitation effects should become interference in mixed-language response blocks. In other words, the presence of two languages in the response set should make it harder for bilinguals to use any task specific strategy to exclude alternative responses in one of the languages and hasten production relative to unrelated distractors. If, however, translations continue to facilitate production in mixed-language response blocks, this would demonstrate the robustness of translation facilitation effects, reducing the likelihood that they could be attributed to strategies that are possible only in limited experimentally controlled and artificial conditions. By extension, this would provide additional support for bilingual processing models in which lexical access not language specific (in which activation spreads between translations without causing interference, see Klaus, Lemhöfer & Schriefers, 2018). Experiment 2 further explored translation facilitation in language mixing by presenting all distractors in the non-target language.

Experiment 1

Method

Subjects. Spanish-English bilingual undergraduates (N = 48) at the University of California San Diego (UCSD) participated for credit. All bilinguals spoke Spanish and English fluently and completed the Multilingual Naming Task (MINT; Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012) as a metric of language proficiency and dominance (Tomoschuk, Ferreira & Gollan, 2018). The MINT score determined their dominance classification. Forty-two bilinguals scored higher on the MINT in English than Spanish, and six scored higher in Spanish than in English. Subject characteristics are shown in Table 3.1.

Table 3.1. *Subject characteristics for Experiments 1 and 2.*

	Experiment 1			Experiment 2		
	M	SD	Range	M	SD	Range
Age	20.00	2.13	18-28	19.88	2.05	18-28
% Female	67.80	-	-	72.92	-	-
% English use currently	82.93	14.19	40-100	80.23	14.53	40-98
% English use growing up	47.83	19.31	5-80	51.40	17.40	5-87
Age began using English regularly	5.66	4.23	0-16	5.63	2.85	0-14
Dominant MINT Score	60.50	2.58	56-65	61.41	2.76	55-68
Non-dominant MINT Score	48.33	7.99	22-62	48.25	9.40	28-61
Average English Self-rating	6.52	0.63	5-7	6.57	0.56	5-7
Average Spanish Self-rating	6.16	0.79	3-7	5.87	0.74	3-7

Materials. Thirty-two Spanish-English translations selected from eight semantic categories were selected to resemble those used in Roelofs et al. (2016). The number of characters in the English translations ($M = 5.16$, $SD = 1.83$) and the number of characters in the Spanish translations ($M = 5.69$, $SD = 1.42$) did not differ significantly in length. Materials are listed in Appendix 3.A. Example materials are shown in Table 3.2.

Table 3.2. *Example materials for Experiment 1, showing response language, picture name, and superimposed distractor.*

		Control	Unrelated	Related	Translation
English	Picture name	eye	eye	eye	eye
	Distractor	XXXXX	napkin	nose	ojo
Spanish	Picture name	ojo	ojo	ojo	ojo
	Distractor	XXXXX	servilleta	nariz	nose

Procedure. Bilinguals first completed a language history questionnaire and then were shown the pictures along with the names to be used in the task and were asked to familiarize themselves with the names. Bilinguals then completed the picture naming task. In this task, bilinguals were first instructed to name the picture as quickly and accurately as possible, based on the flag cue shown at the top of the screen (the American flag cued English responses, and the Mexican flag cued Spanish responses). First, a fixation cross appeared on the screen for 350 ms, followed by a blank screen for 150 ms, the flag cue for 250 ms, and last the picture appeared. Bilinguals had 3 seconds from the onset of the picture in which to respond before the picture disappeared. Distractors appeared simultaneously with the picture, and remained until the picture disappeared, and voice onset triggered both to disappear. The distractors were five Xs (control condition), a translation of the picture in the non-target language (translation condition), a semantically related word in the target language from the same semantic list (related condition), and a semantically unrelated word also in the target language taken from another semantic list (unrelated condition). For example, when a bilingual was to name a picture of a nose, the Spanish translation *nariz* appeared on the picture in the translation condition, the English word *eye* appeared in the semantically related condition, and the English word *glass* appeared in the semantically unrelated condition. There was an inter-stimulus interval of 850 milliseconds before the next picture appeared. At the beginning of the first block, bilinguals were given three practice trials in English to assess the sensitivity of the microphone. Bilinguals then completed six blocks of this task: two in Spanish, two in English, and two mixed language response blocks. The order was counterbalanced across bilinguals using a sandwich design such that half of subjects performed single language response blocks at the beginning and end of the experiment with mixed language response blocks in the middle and half performed mixed language response blocks at the beginning and end of the experiment with single language response blocks in the middle. After these tasks were completed, bilinguals completed the MINT. Results are shown in Table 3.1.

Analysis. Data were analyzed in R (R Core Team, 2013). All responses with a response time (RT) less than 100 ms or greater than 3000 ms were removed. Trials with erroneous microphone triggers or incorrect naming were removed from the analysis. Response times were log transformed and then entered into by-subject (F_1) and by-item (F_2) repeated measure analyses of variance (ANOVA). This was chosen over mixed-effect models, to match analyses reported by Roelofs et al., (2016; and because design complexity led to convergence issues in the mixed-effects models, Barr, Levy, Scheepers & Tily, 2013). (Note also that Barr et al, 2013, showed that for continuous measures such as response times, F1xF2 analyses are at least as suitable as mixed-effect models.) We first report entire model outcomes, with language dominance, language mixing and all four distractors as conditions, leaving out trial type (stay versus switch) to simplify interpretation. Then we report the critical planned comparisons looking at only trials with semantically unrelated and translation distractors as these index translation distractor effects as reported in Costa et al., (1999; the control and semantically related conditions were included to replicate the procedure in Roelofs et al., 2016). We also report logistic mixed-effect regression models of only the planned comparisons of the errors looking at unrelated and translation conditions, in order to rule out the possibility of error rates driving the effects. Finally, we report by-subject and by-item analyses of the planned comparison within mixed-language-response blocks reintroducing trial type as a factor.

Results

Figure 3.1 shows the response times for Experiment 1 organized by picture-word interference condition and language mixing, shaded by language dominance. Among these conditions, bilinguals were slower to name pictures in mixed language response conditions compared to single language response conditions, $F_1(1,35) = 173.48$, $MSE = 996{,}638$, $p < .001$, $F_2(1,20) = 606.32$, $MSE = 285{,}607$, $p < .001$. Bilinguals responded fastest with control and translation distractors and slowest with semantic distractors, a main effect of distractor type, $F_1(3,126) = 138.17$, $MSE = 173.962$, $p < .001$, $F_2(3,81) = 52.74$, $MSE = 456{,}709$, $p < .001$. Bilinguals named pictures equally quickly in their two languages, a nonsignificant effect of language dominance (Fs < 1). Distractor type effects were generally bigger in the non-dominant than in the dominant language, an interaction between dominance and distractor type, $F_1(3,134) = 8.85$, $p < .001$, $MSE = 78{,}351$, $F_2(3,89) = 4.55$, $MSE = 147{,}408$, $p = .005$. Additionally, the non-dominant language typically elicited slower responses in the single language response block, but in the mixed language response blocks bilinguals exhibited reversed language dominance effects (i.e., they named pictures more quickly in the language that is usually nondominant), an interaction between dominance and block type, $F_1(1,43) = 13.45$, $MSE = 157{,}629$, $p < .001$, $F_2(1,28) = 8.68$, $MSE = 226{,}498$, $p = .006$. Distractor type effects were generally bigger in the single language response blocks, an interaction between distractor type and block type $F_1(3,131) = 12.60$, $MSE = 67{,}638$, $p < .001$, $F_2(3,86) = 2.78$, $MSE = 318{,}561$, $p = .046$. Finally, distractor type effects were bigger in the non-dominant language in the single language response blocks, but bigger in the dominant language and the mixed language response blocks, a 3-way interaction between language dominance, block type, and distractor type that was a significant by subject but not by item (but see below), $F_1(3,138) = 4.02$, $MSE = 82.783$, $p = .008$, $F_2(3,93) = 1.89$, $MSE = 163,050$, $p = .14$.

In a planned comparison between only trials with unrelated and translation distractors, we found that bilinguals were slower to name pictures in mixed language response conditions compared to single language response conditions, $F_1(1,41) = 172.46$, $MSE = 478,601$, $p < .001$, $F_2(1,26) = 269.87$, $MSE = 307,284$, $p < .001$. Bilinguals named pictures faster when the distractor was the translation as compared to the unrelated word, $F_1(1,40) = 78.63$, $MSE = 199,232$, $p < .001$, $F_2(1,25) = 29.77$, $MSE = 517,755$, $p < .001$. Bilinguals named pictures equally quickly in their two languages, a non-significant effect of language dominance (*F*s <1). The translation facilitation effect was greater when bilinguals named pictures in the non-dominant language (but see the below interaction with block type), $F_1(1,44) = 21.63$, $MSE = 89,818$, $p < .001$, $F_2(1,29) = 13.38$, $MSE = 141,246$, $p = .001$. Bilinguals also named pictures more slowly in the non-dominant than the dominant language in the single language response block, but exhibited reversed dominance effects in the mixed language response blocks, $F_1(1,45) = 12.36$, $MSE = 105,220$, $p = .001$, $F_2(1,30) = 7.05$, $MSE = 182,590$, $p = .013$. Of great interest, translation facilitation effects were larger in the single language response blocks than in the mixed language response block in the by-subject analysis, $F_1(1,43) = 13.31$, $MSE = 54,070$, $p < .001$, $F_2(1,28) = 3.53$, $MSE = 218,340$, $p =.071$, though, critically, in a post-hoc analysis translation facilitation effects were significant in the mixed language response block alone in the by-subject analysis ($F_1(1,44) = 21.59$, $MSE = 53,392$, $p < .001$, $F_2(1,29) = 8.14$, $MSE = 78,502$, $p =.071$). Finally, with unrelated distractors, bilinguals named pictures faster in the dominant language, but with translation distractors language dominance effects reversed, such that bilinguals named pictures faster in the non-dominant language, a significant 3-way interaction language dominance, block type, and distractor type, $F_1(1,46) = 6.87$, *MSE* = 118,489, $p < .001$, $F_2(1,31) = 4.40$, *MSE* =169,995, $p = .044$.

Figure 3.2 shows the naming error data, which were analyzed using logistic mixed-effect regression models to examine the unrelated and translation conditions. Bilinguals produced more errors in mixed language response blocks than in single language response blocks ($\chi^2 = 32.4$, $p < .001$). Bilinguals also produced more errors when naming pictures in the non-dominant language than the dominant ($\chi^2 = 4.28$, $p = .039$). Finally, bilinguals produced more errors in the non-dominant than in the dominant language with an unrelated distractor as compared to the translation distractor, specifically in the mixed language response blocks as compared to the single language response block, causing a three-way interaction ($\chi^2 = 4.67$, $p = .030$). All other effects were non-significant (*p*s > .14).

Though not the focus of these analyses, we additionally analyzed switching effects within the mixed language response blocks, shown in Table 3.3. In a by-subject and by-item analysis of trial type (stay vs. switch), distractor type (focusing only on the planned comparison, unrelated vs. translation), and language dominance, we found that bilinguals produced switch trials slower than stay trials $F_1(1,41) = 79.46$, *MSE* = 115,875, $p < .001$, $F_2(1,21) = 11.41$, *MSE* =338,643, $p < .01$. They also produced the non-dominant language faster than the dominant, $F_1(1,42) = 13.39$, *MSE* = 170,378, $p < .001$, $F_2(1,27) =11.51$, *MSE* = 169,452, $p < .01$, and produced pictures with translation distractors faster than unrelated distractors $F_1(1,40) = 22.46$, *MSE* = 104,567, $p < .001$, $F_2(1,25) = 9.21$, *MSE* = 246,990, $p < .01$. No interactions were significant (*p*s > .13).

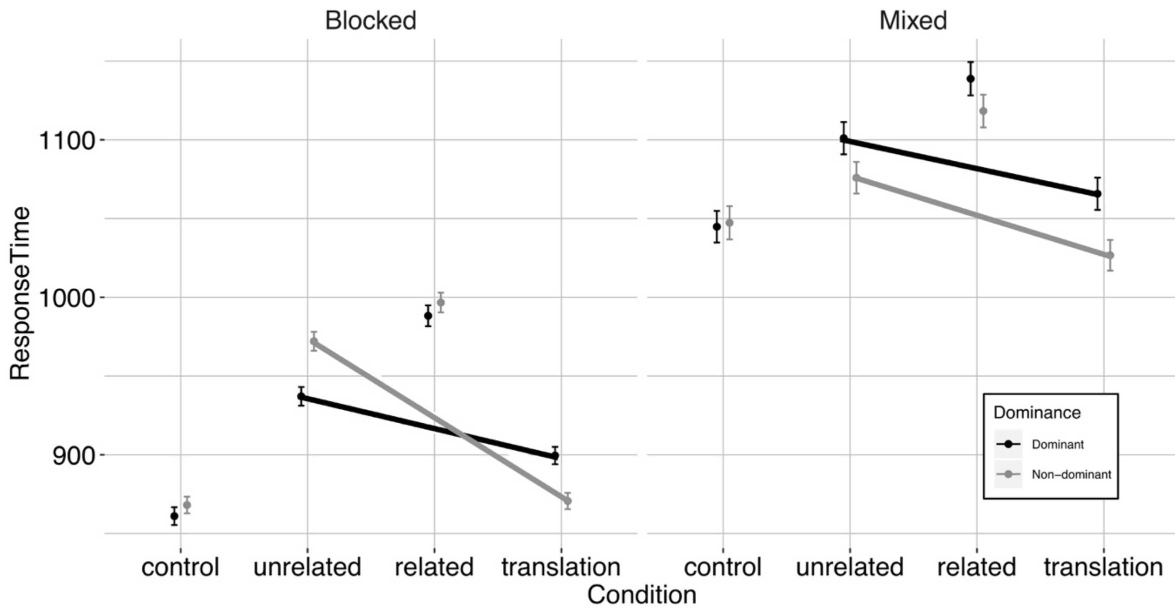

Figure 3.1. *Response times for Experiment 1 grouped by distractor condition and block type, shaded by language dominance. Error bars represent standard error. Lines connect conditions in planned comparison.*

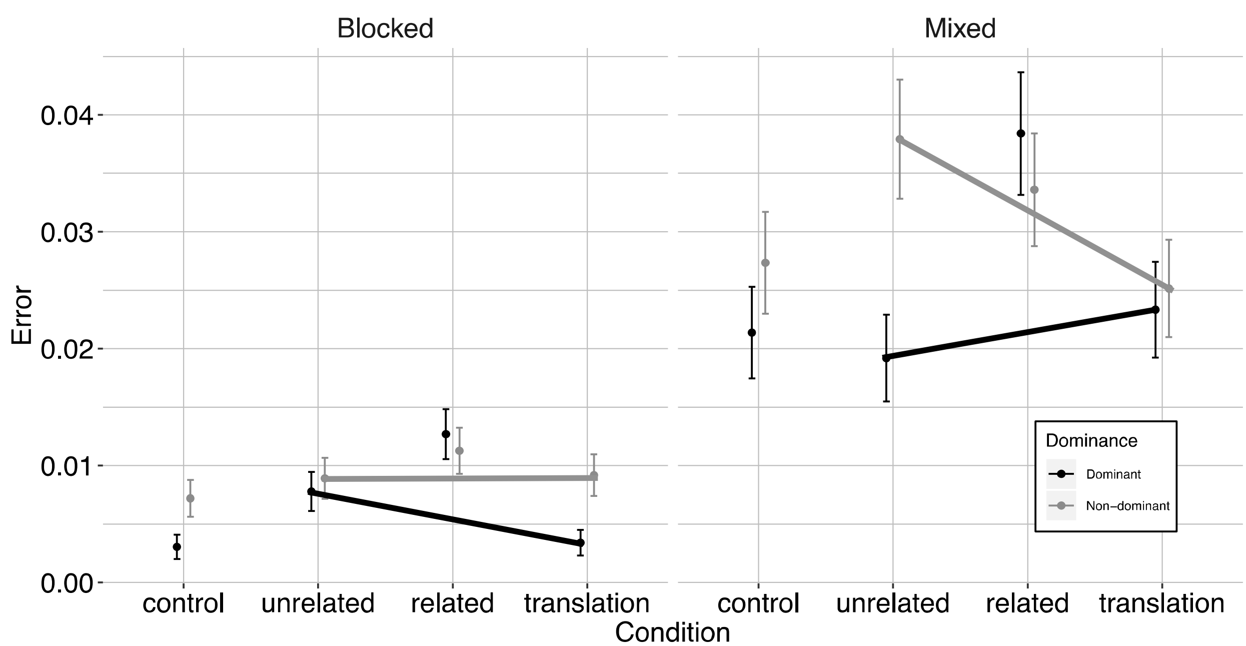

Figure 3.2. *Error rates for Experiment 1 grouped by distractor condition and block type, shaded by language dominance. Error bars represent standard error. Lines connect conditions in planned comparison.*

Table 3.3. *Response times means and standard deviations for Experiments 1 and 2 by trial type (stay versus switch), language dominance, and distractor type.*

		Control		Unrelated		Related		Translation	
		Stay	Switch	Stay	Switch	Stay	Switch	Stay	Switch
Experiment 1	Dominant	995 (363)	1084 (384)	1074 (385)	1131 (390)	1100 (379)	1170 (405)	1021 (345)	1114 (423)
	Non-dominant	996 (400)	1068 (381)	1015 (363)	1121 (382)	1086 (386)	1135 (376)	986 (333)	1062 (403)
Experiment 2	Dominant	971 (347)	1054 (378)	1037 (369)	1113 (392)	1044 (374)	1120 (401)	1000 (367)	1044 (358)
	Non-dominant	973 (348)	1042 (374)	1036 (363)	1109 (403)	1098 (401)	1106 (395)	975 (353)	1037 (369)

Discussion

In Experiment 1, we tested the possibility that facilitation from translation distractors comes from a "what not to say" strategy. If so, mixing the language of response (i.e., so that bilinguals needed to name pictures in one language half the time, and in the other language the other half of the time) should reduce the extent to which such a strategy can be used, thereby eliminating translation facilitation effects. Such an effect may even cause translation distractors to slow picture naming, compared to unrelated distractors, revealing competition between translation equivalents in an experimental design that better matches circumstances bilinguals face in naturalistic language use (in which languages are mixed in speech production and comprehension alike, not just in one modality). Instead we found that translation facilitation effects are maintained in mixed-language-response blocks for the dominant language, and though translation facilitation effects were significantly reduced for the non-dominant language, here too translation distractors facilitated responses.

In the error rates, we see almost no effects in the single language response blocks. Bilinguals were, however, more likely to make an error in their non-dominant language in the unrelated condition of the mixed language response blocks relative to the translation condition (i.e., an effect of translation facilitation for the non-dominant language in the mixed-language-response block, see General Discussion).

In single language response blocks, bilinguals responded more slowly in the non-dominant language, significantly so in the unrelated condition (and numerically in the same direction in both the control and related conditions). By contrast, in mixed language response conditions, bilinguals tended to exhibit reversed language dominance effects in all but the control condition. The reversal of language dominance on only some trials matches a pattern reported by Meuter and Allport (1999) in which dominance was reversed only on switch trials, and could imply that inhibition is applied only on trials in which the language of the task (name pictures in language X) and the language of the distractor (presented in language Y) conflict. Such inhibition might be stronger at the block-wide level when bilinguals know in advance that they will need to produce picture names in both languages. It is not clear, however, why bilinguals would inhibit a language that does not compete for selection. That is, the pattern of results that would have provided more unequivocal support for competition between languages (the non-selective model) would have been translation distractors delaying naming. Instead, translation distractors facilitated responses, significantly so in both languages, and in both single-language response and mixed-language response blocks (see General Discussion for discussion of three-way interaction).

Experiment 2

Experiment 1 manipulated the language of response by introducing mixed language response to a typically single language response paradigm. It showed that translation facilitation is maintained in a mixed language context, suggesting that a "what not to say" strategy is an inadequate account of translation facilitation in that language. Experiment 2 seeks primarily to reproduce the results of Experiment 1 and manipulate the distractor language. In Experiment 2, all distractors were presented in the non-target language, as in Costa et al. (1999). For example, when naming pictures in English, the semantically related distractor, the semantically unrelated distractor, and the translation distractors were all presented in Spanish in Experiment 2, but just the translation distractors were in Spanish in Experiment 1. If, in this scenario, translation facilitation continues to not differ between single- and mixed-language response blocks, this would provide stronger evidence for a language specific control model. In other words, presenting distractors entirely in the non-target language should further reduce the possibility of a "what not to say" strategy.

Separately, it may be that in single language response blocks, having all distractors in the non-target language may speed up production as the non-target language will be easier to ignore as a whole. Alternatively it may be that because responses are already all given in one language, having a language-specific cue will not help, as subjects can already represent which language to use to respond. In the mixed language response block, however, it may be that when the distractor is always in the non-target language, bilinguals get an early cue (in addition to the flag) as to which language they should not speak on this trial. If so, responses might be faster in all trials of Experiment 2 relative to Experiment 1, given that the distractor could serve to cue which language is not to be spoken on every trial.

Method

Subjects. Spanish-English bilingual students (N = 48) from the same pool of subjects but that had not participated in Experiment 1 were recruited and completed the study for course credit. All bilinguals spoke Spanish and English fluently and MINT score determined their dominance classification. Forty-five subjects scored higher on the English MINT than the Spanish MINT and three subjects scored higher on the Spanish MINT than the English MINT. Full participant characteristics are listed in Table 3.1.

Materials, Procedure, Analysis. The materials, procedure and analyses were identical between Experiments 1 and 2, with one exception: distractor words in the semantically related and unrelated conditions appeared in the non-target language. For example, when the picture of a nose was to be named in English, the semantically related distractor appearing on the picture was no longer the English word *eye* but its Spanish translation equivalent *ojo*. Example materials shown in Table 3.4.

Table 3.4. *Example materials for Experiment 2, showing response language, picture name, and superimposed distractor.*

		Control	Unrelated	Related	Translation
English	Picture name	eye	eye	eye	eye
	Distractor	XXXXX	servilleta	nariz	ojo
Spanish	Picture name	ojo	ojo	ojo	ojo
	Distractor	XXXXX	napkin	nose	nose

Results

Figure 3.3 shows the response times for Experiment 2 organized by picture-word interference condition and language mixing, shaded by language dominance. When considering all distractor types in a model, bilinguals were slower to name pictures in mixed-language response conditions compared to single-language response conditions, $F_1(1,36) = 123.51$, $MSE = 1,191,655$, $p < .001$, $F_2(1,20) = 545.66$, $MSE = 272,749$, $p < .001$. Responses were fastest with control and translation conditions and slowest with semantic distractors, a main effect of distractor type, $F_1(3,129) = 230.65$, $MSE = 99,224$ $p < .001$, $F_2(3,81) = 47.54$, $MSE = 479,383$, $p < .001$. Bilinguals named pictures equally quickly in their two languages, a nonsignificant effect of language dominance (*F*s < 1). Additionally, the non-dominant language typically elicited slower responses in the single language response block, but in the mixed language response blocks bilinguals exhibited reversed language dominance effects (i.e., they named pictures more quickly in the language that is usually nondominant), an interaction between dominance and block type in the by-subject analysis, $F_1(3,137) = 3.94$, $MSE = 101,744$, $p = .010$, $F_2(3,89) = 2.45$, $MSE = 171,196$, $p = .070$. Unlike in Experiment 1, neither the language mixing by dominance, nor three way interaction were significant (*p*s > .49).

In the planned comparison, bilinguals were slower to name pictures in the mixed language response condition relative to single language response blocks $F_1(1,42) = 104.88$, $MSE = 708,853$, $p < .001$, $F_2(1,26) = 253.53$, $MSE = 296,738$, $p < .001$. They named pictures more quickly with translation than with unrelated distractors, $F_1(1,41) = 156.72$, $MSE = 109,228$, $p < .001$, $F_2(1,25) = 57.92$, $MSE = 296,805$, $p < .001$, and named pictures equally quickly in their two languages, a non-significant effect of language dominance, $F_1(1,43) = 1.14$, $MSE = 409,090$, $p = .29$, $F_2(1.27) = 0.99$, $MSE = 419,935$, $p = .33$. The translation facilitation effect was also greater in the single language response block than the mixed language response block in the by-subject analysis, $F_1(1,45) = 6.36$, $MSE = 96,383$, $p = .015$, $F_2(1,29) = 2.80$, $MSE = 215,551$, $p = .11$. None of the remaining two-way, nor three-way interactions were significant (*p*s > .47).

Figure 3.4 shows the naming error data, which were analyzed using mixed-effect models to examine the unrelated and translation conditions. Among these conditions, bilinguals produced more errors in mixed language response than single language response blocks ($\chi^2 = 43.9$, $p < .001$), and they produced fewer errors in the translation condition than in the unrelated condition, that is, a significant translation facilitation effect, ($\chi^2 = 5.53$, $p = .019$). All other effects were non-significant (*p*s > .13).

As in Experiment 1, we additionally analyzed switching effects within the mixed language response blocks, shown in Table 3.3. In a by-subject and by-item analysis of trial type (stay vs. switch), distractor type (focusing only on the planned comparison, unrelated vs. translation), and language dominance, we found that bilinguals produced switch trials slower than stay trials $F_1(1,42) = 64.28$, $MSE = 95,955$, $p < .001$, $F_2(1,21) = 6.14$, $MSE = 484,350$, $p = .022$. They also produced pictures with translation distractors faster than unrelated distractors $F_1(1,40) = 45.27$, $MSE = 106,592$, $p < .001$, $F_2(1,25) = 29.31$, $MSE = 158,938$, $p < .001$. Bilinguals did not name pictures faster in either the dominant or non-dominant language, nor were any interactions significant (*p*s > .18).

We additionally asked whether response times were faster across Experiment 2 relative to Experiment 1 as a result of the change of distractor language. To assess this we performed a by-subject and by-item analysis of response times using experiment, language response blocks, and distractor type as factors. Response times were marginally significantly faster in Experiment 2 relative to Experiment 1, $F_1(1,90) = 3.04$, $MSE = 1.55e15$, $p = .085$, $F_2(1,28) = 3.33$, $MSE = 1.46e15$, $p = .08$, and these differences did not differ significantly between single language response and mixed language response blocks or between distractor types (*F*s < 1).

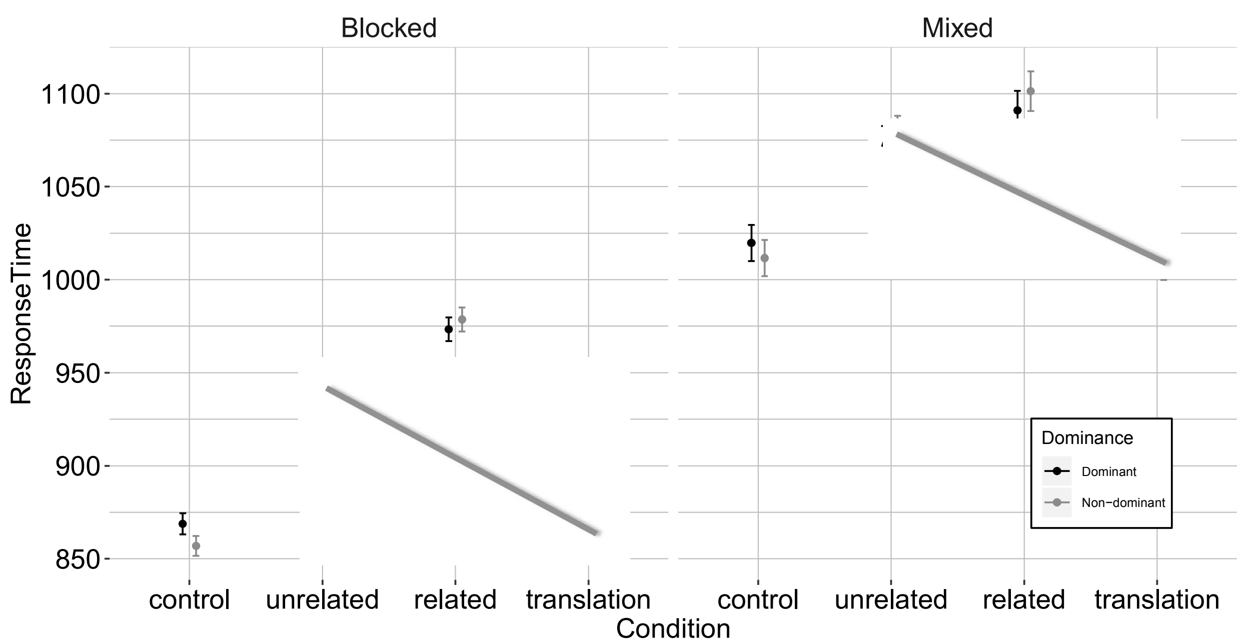

Figure 3.3. *Response times for Experiment 2 grouped by distractor condition and block type, shaded by language dominance. Error bars represent standard error. Lines connect conditions in planned comparison.*

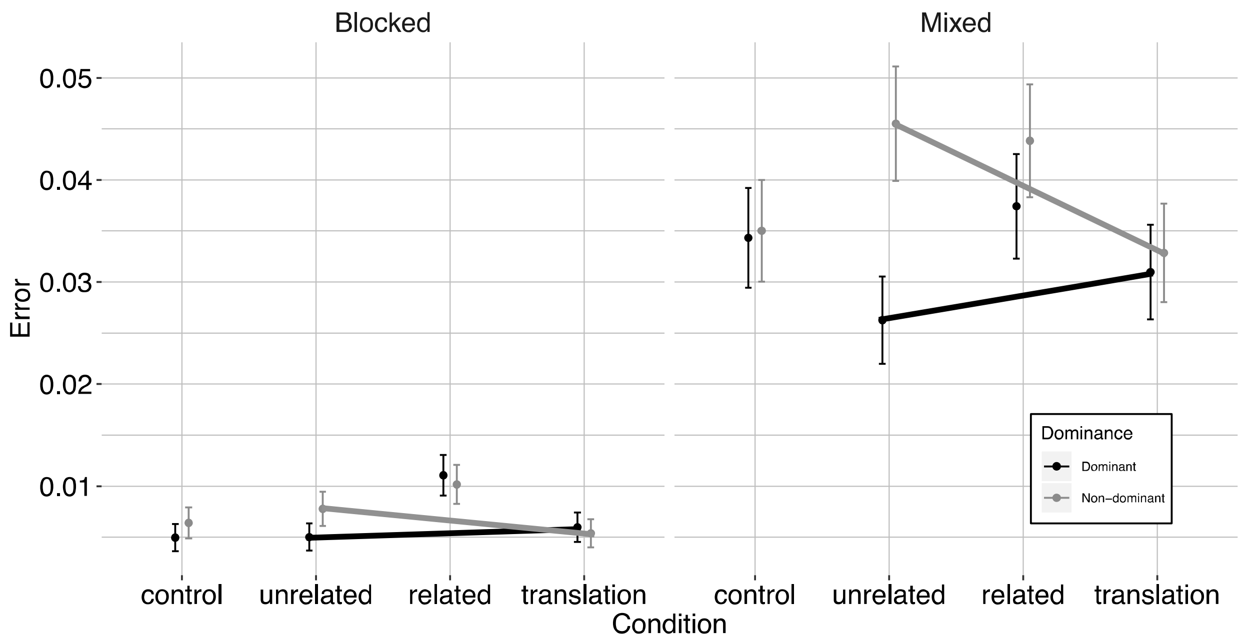

Figure 3.4. *Error rates for Experiment 2 grouped by distractor condition and block type, shaded by language dominance. Error bars represent standard error. Lines connect conditions in planned comparison.*

Discussion

In Experiment 2 we predicted that translation facilitation would again be robust, even in the presence of other distractors of the same language, and indeed we found a significant translation facilitation effect in the response times. The error rates showed the same pattern as in Experiment 1. Bilinguals were more likely to make an error in their non-dominant language in the unrelated condition of the mixed language response blocks, though the three-way interaction was non-significant in Experiment 2.

If presenting all distractors in the non-target language impacted overall production in our experiments, we would have expected that subjects may use the fact that all distractors now appear in the non-target language as a task strategy to hasten production. We saw that response times across Experiment 2 were somewhat faster – about 15 ms – but that this difference was only marginally significant. Critically, though, regardless of distractor language, translation facilitation remains robust in response times across experiment, block type, and language.

General Discussion

Two experiments show that the translation facilitation effect is robust in both single language and mixed language response blocks tasks. In Experiment 1, we showed that there was a consistent translation facilitation effect in both single language response blocks (aligning with typical picture-word interference tasks) and mixed language response blocks, in which bilinguals switched between languages. Additionally, translation facilitation was greater for the non-dominant language in the single language response block. In Experiment 2, we changed only the language of the unrelated and semantically related distractor conditions from the target language to the non-target language (so that all distractors were in the non-target language), and demonstrated that even with all distractors coming from the same language as the translation condition, translation facilitation effects remain in both the single language and mixed language response conditions. The magnitude of this effect did not differ for the non-dominant language in the single language block as in Experiment 1, and translation facilitation effects did not vary across single versus mixed language response blocks (with a possible exception in Experiment 1 for the non-dominant language only in RTs, with a trend in the opposite direction in errors).

According to Roelofs et al., (2016) a language non-specific model of control in bilingual lexical access predicts that translation facilitation effects should become translation interference effects when the response set in a picture word interference experiment involves both languages. Experiment 2 of Costa et al. (1999) began to explore this effect, showing that the "Spanish identity condition" (in this experiment a Spanish cognate distractor was superimposed on a picture to be named in Catalan), produced significant facilitation in a mixed response language block. Here, in two experiments, we extend and bolster that effect, showing that the translation facilitation effect never becomes interference in mixed language response blocks, even with noncognate distractors, whether bilinguals named pictures in their dominant or nondominant languages, with the same subjects tested in both single- and mixed language response blocks, and with materials counterbalanced across languages. A "what not to say" strategy cannot account for this facilitation when a bilingual does not know which language they will need to produce on an upcoming trial. As such, Costa et al.'s (1999) explanation of translation facilitation, in which activation spreads between languages after language-wide inhibition is activated, remains the theory that is most consistent with the range of effects observed.

In Experiment 1, we saw that translation facilitation was greater in the non-dominant language of the single-language-response blocks than in any other conditions, causing a significant three-way interaction. In error rates, however, there was an opposite pattern such that the non-dominant language showed *greater* translation facilitation in the mixed-language-response blocks than the single-language-response blocks. This possible speed-accuracy tradeoff in the magnitude of translation facilitation effects across block types makes it difficult to argue that block type (single- vs. mixed-language response) had any effect on the nature of translation facilitation effects. Another possible interpretation is that error rates in the dominant language were at floor– in mixed-language-response blocks error rates may have been too low for the dominant language in the translation distractor condition to show any effect. By contrast, the non-dominant language in the mixed-language response conditions created the only circumstances in which the task was difficult enough to show effects in the error rates. Assuming the latter interpretation (the floor effect) it is possible the reduction of translation facilitation effects for the non-dominant language (in RTs) suggests that bilinguals are capable of using a task strategy in some particular circumstances; a “what not to say” strategy may be used to aid a weaker language in a block in which the strategy is feasible. Such a strategy, though, may use different mechanisms than the core translation facilitation effect, which remained with equal magnitude across other block types and in Experiment 2. Future studies may consider paradigms that reduce error rates (e.g. increased inter-stimulus intervals), to potentially allow for effects to manifest solely in response times, or paradigms that increased error rates to examine the possibility of floor effects.

Why, given that translation facilitation remains robust in mixed language setting, do effects differ between picture-word interference tasks and phono-translation tasks? The primary difference is likely that phono-translation tasks present near translation distractors whereas picture-word interference tasks present translation equivalents – which provide a strong semantic cue as to the identity of the picture, perhaps overriding any slowing caused by competition for selection (if such competition exists). One possible, though speculative, explanation is that when activation spreads between the distractor and the target a near-translation it must go through an additional step (the translation itself), diverting attentional resources to that additional step, hindering language-wide inhibition. Further study is needed to pinpoint the source of the differences between these effects.

Roelofs et al. (2016) suggests that this difference is explained at another stage of production, stating that competition in phono-translation tasks exists at a phonological encoding layer of production, stating that “Distractor words that are phonologically related to the translation of the picture name will increase the activation of phonologically related syllable motor programs, making them stronger competitors during phonetic encoding.” Though this would explain the pattern seen in phono-translation effects, it does not fit into a cohesive theory of lexical selection. While the studies presented here cannot provide strong evidence of between-language lexical competition, considerable work does (for examples. see Declerck & Philipp, 2015; Green, 1998). Within picture-word interference tasks, it is clear that phono-translation and translation facilitation effects are tied to semantic facilitation. The data presented here show that any models without lexical competition are an inadequate explanation. Translation facilitation persists even when the response set contains both languages. That is, whether bilinguals name pictures exclusively in one language across an entire block of trials, or in the more difficult mixed language response blocks, translation facilitation effects remained significant and robust in both languages. This was true whether the translation distractors were presented in the non-target language only, or if distractors in both languages were presented. A“what not to say” strategy is inadequate for explaining translation facilitation effects.

Chapter 3, in full, is a reprint of the material to be submitted to Language, Cognition and Neuroscience. Tomoschuk, Brendan; Gollan, Tamar H.; Ferreira, Victor S. The dissertation author was the primary investigator and author of this paper.

Appendix 3.A. *Items used in Experiments 1 and 2.*

Category	English	Spanish
body parts	arm	brazo
body parts	tongue	lengua
body parts	nose	nariz
body parts	eye	ojo
animals	cow	vaca
animals	horse	caballo
animals	dog	perro
animals	bird	pájaro
furniture	table	mesa
furniture	mirror	espejo
furniture	bed	cama
furniture	door	puerta
clothing	shoes	zapatos
clothing	skirt	falda
clothing	dress	vestido
clothing	shirt	camisa
kitchenware	glass	vaso
kitchenware	spoon	cuchara
kitchenware	fork	tenedor
kitchenware	napkin	servilleta
foods	cheese	queso
foods	strawberry	fresa
foods	meat	carne
foods	watermelon	sandía
school items	pencil	lápiz
school items	backpack	mochila
school items	ruler	regla
school items	book	libro
tools	hammer	martillo
tools	saw	sierra
tools	shovel	pala
tools	scissors	tijeras

References

Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of neurolinguistics*, *20*(3), 242-275.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255-278.

Boukadi, M., Davies, R. A., & Wilson, M. A. (2015). Bilingual lexical selection as a dynamic process: Evidence from Arabic-French bilinguals. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *69*(4), 297.

Costa, A., Colomé, À., Gómez, O., & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, *6*(3), 167-179.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection?. *Journal of Memory and language*, *41*(3), 365-397.

Costa, A., Pannunzi, M., Deco, G., & Pickering, M. J. (2017). Do bilinguals automatically activate their native language when they are not using it?. *Cognitive Science*, *41*(6), 1629-1644.

De Bot, K. (2000). A bilingual production model: Levelt's "speaking" model adapted. *The bilingualism reader*, 420-442.

De Groot, A. M., & Nas, G. L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of memory and language*, *30*(1), 90-123.

Declerck, M., & Philipp, A. M. (2015). A review of control processes and their locus in language switching. *Psychonomic bulletin & review*, *22*(6), 1630-1645.

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and cognition*, *5*(3), 175-197.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *15*(3), 594-615.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, *1*(2), 67-81.

Hall, M. L. (2011). Bilingual picture–word studies constrain theories of lexical selection. *Frontiers in psychology*, *2*, 381.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language?. *Bilingualism: language and cognition*, *1*(3), 213-229.

Klaus, J., Lemhöfer, K., & Schriefers, H. (2018). The second language interferes with picture naming in the first language: Evidence for L2 activation during L1 production. *Language, Cognition and Neuroscience*, *33*(7), 867-877.

Kroll, J. F., Bobb, S. C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and cognition*, *9*(2), 119-135.

Kroll, J. F., & De Groot, A. M. (Eds.). (2009). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford University Press.

Kroll, J. F., & Gollan, T. H. (2013). Speech planning in two languages. In *The Oxford Handbook of Language Production*.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of memory and language*, *40*(1), 25-40.

Piai, V., Roelofs, A., Jensen, O., Schoffelen, J. M., & Bonnefond, M. (2014). Distinct patterns of brain activity characterise lexical activation and competition in spoken word production. *PloS one*, *9*(2), e88674.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1-3), 107-142.

Roelofs, A. (1998). Lemma selection without inhibition of languages in bilingual speakers. *Bilingualism: Language and Cognition*, *1*(2), 94-95.

Roelofs, A., Piai, V., Rodriguez, G. G., & Chwilla, D. J. (2016). Electrophysiology of cross-language interference and facilitation in picture naming. *cortex*, *76*, 1-16.

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

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2018). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*, 1-21.

CONCLUSION

Three sets of studies explored the acquisition and mechanisms of multilingual lexical regulation. In Chapter 1, we explored non-native language interference in trilinguals, finding that the language from which a bilingual learns a third language might explain effects previously attributed to cognitive similarity between non-native languages. In Chapter 2, we probed the mechanisms of this language of instruction effect, with results suggesting that it affects language control at a language-wide level, and that it persists beyond third language practice. Finally, in Chapter 3, we explore language-wide vs language non-specific cognitive control among bilinguals combining two commonly used experimental paradigms, and found that translation facilitation effects are robust to the difficulty of a two-language context, supporting a language-wide control model of multilingual language control.

Taken together, these studies suggest that during the learning of a new language, a learner uses a top-down, language-wide control mechanism to prevent interference from known languages while producing the newly acquired language. This control mechanism shifts from a top-down method to a more automatic process that fluent bilinguals use, though it appears to remain at the language-wide level.

These studies also speak to the necessity of using a variety of paradigms and populations to study the same phenomenon. In Chapters 1 and 2, we explored language of instruction in two different tasks, with two different sets of bilinguals. While we do see consistent modulation of language control between groups, they are modulated in different ways. In Chapter 1, when bilinguals were monitoring the new L3 for phonemes, having learned the L3 from the non-dominant language affected performance more. In contrast, in Chapter 2, when bilinguals were naming L3 pictures mixed with pictures from their known languages, learning from the dominant language affected performance more, suggesting that task difficulty may determine the nature of language-of-interference effect. This converging evidence allowed us to draw conclusions about the nature of the effect, that it is not tied to particular instructional material, that it persists across language group, and that it must modulate control at the whole-language level. Chapter 3 also explored the relationship between task and phenomenon. We drew from both the picture-word interference and the language mixing literatures to examine translation facilitation, allowing us to adjudicate between contrasting evidence from different paradigms. We ultimately ascribe differences in the literature to the differences in tasks, concluding that a mixture of tasks suggests a language-wide level of control in fluent bilinguals.

In a more applied sense, these studies have various implications for a new language learner. One surface-level takeaway is that a bilingual, when given a choice, should think critically about the language they use during the acquisition of a third. Using the more dominant of the two languages may facilitate learning at first, but could leave the third language more susceptible to interference from the non-dominant known language, whereas learning via the non-dominant language may help mitigate non-native language interference more generally. Another takeaway is that, because language control seems to operate at a whole-language level, any practice mitigating interference from one language helps a bilingual learn to mitigate any interference from that language. Experience inhibiting one known language may ease future control of that and the other language, even in a new context or with new vocabulary.

Language control, like many other aspects of language, is a skill that is tied to the both the amount and the nature of the exposure we receive when learning a new language. By exploring these mechanisms in both fluent and emerging bilinguals, we not only build a more complete picture of language development and acquisition, we lay the foundation for building tools to improve the language-learning process.