Title

Grammatical constraints on language switching: Language control is not just executive control

Authors

Gollan, Tamar H

Goldrick, Matthew

Publication Date

2016-10-01

DOI

10.1016/j.jml.2016.04.002

Peer reviewed

Published in final edited form as:
J Mem Lang. 2016 October ; 90: 177–199. doi:10.1016/j.jml.2016.04.002.

Grammatical Constraints on Language Switching: Language Control is not Just Executive Control

Tamar H. Gollan and
University of California, San Diego

Matthew Goldrick
Northwestern University

Abstract

The current study investigated the roles of grammaticality and executive control on bilingual language selection by examining production speed and failures of language control, or *intrusion errors* (e.g., saying *el* instead of *the*), in young and aging bilinguals. Production of mixed-language connected speech was elicited by asking Spanish-English bilinguals to read aloud paragraphs that had mostly grammatical (conforming to naturally occurring constraints) or mostly ungrammatical (haphazard mixing) language switches, and low or high switching rate. Mixed-language speech was slower and less accurate when switch-rate was high, but especially (for speed) or only (for intrusion errors) if switches were also ungrammatical. Executive function ability (measured with a variety of tasks in young bilinguals in Experiment 1, and aging bilinguals in Experiment 2), slowed production and increased intrusion rate in a generalized fashion, but with little or no interaction with grammaticality. Aging effects appeared to reflect reduced monitoring ability (evidenced by a lower rate of self-corrected intrusions). These results demonstrate robust effects of grammatical encoding on language selection, and imply that executive control influences bilingual language production only after sentence planning and lexical selection.

Keywords

bilingualism; aging; speech errors; intrusion errors; read aloud; switching; inhibition; reversed dominance

Some of the most impressive demonstrations of agility in the domain of spoken language production include the prosodic and expressive speech often exhibited by newscasters or storytellers, unusually rapid production in a single language (e.g., when sportscasters describe team sports action live), and fluid alternation back and forth between languages by

Corresponding Author: Tamar H. Gollan, Department of Psychiatry, University of California, San Diego, 9500 Gilman, Drive, 0948, La Jolla, CA 92093-0948, tgollan@ucsd.edu.

Declaration of Conflicting Interests
The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

bilingual or multilingual speakers. Indeed, language switching has become one of the most broadly studied topics in the field of research on bilingualism—without question, it has played a major role in shaping theories of control over bilingual language selection. With relatively few exceptions the bulk of research on this topic has focused somewhat narrowly on obligatory and intended switches in out of context speech, relatively little on switches in connected speech, and even less on *un*intended language switches, though these can provide a unique and powerful form of evidence on the cognitive mechanisms underlying bilingual language selection.

Bilingual language switching can be considered as a specific example of the more general problem of regulation and control of production processes, which may be driven by two general types of mechanisms. Language-specific knowledge clearly plays a role in controlling speech production. For example, word exchange errors (*I'm writing a mother to my letter*) overwhelmingly respect word class – or *part of speech* (i.e., nouns exchange with nouns, verbs with verbs, etc.; Garrett, 1975), suggesting that knowledge of syntactic properties helps control which lexical items are selected for production (because lexical items in different parts of speech serve distinct syntactic functions; see Dell, Oppenheim, & Kittredge, 2008, for review and discussion). In addition to such domain-specific knowledge, mechanisms that support goal directed behavior in non-linguistic cognitive domains might also support control of speech production. For example, the need to select from among competing response alternatives arises across many situations, and common mechanisms may subserve this function across these cases (e.g., Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). Few studies have considered how these two types of mechanisms might function jointly to control bilingual language selection, and fewer still have considered how language-specific and domain-general factors might modulate the control of both intended and *unintended* language switches in production of connected speech.

Factors Modulating Control of Language Switching in Out-of-Context Speech

A great deal of the work in bilingual language control has examined single word production tasks. These have suggested a number of task properties that seem to facilitate switching between languages. Onset of speech in such studies is typically slower when speakers are cued to switch languages relative to when they are cued to continue speaking the same language they used on the previous trial. This difference is assumed to reflect *switch costs*, i.e., the processing cost associated with switching languages. Switch costs are smaller, but not eliminated entirely, when preceded by longer preparation times (e.g., Costa & Santesteban, 2004; Fink & Goldrick, 2015; Philipp, Gade, & Koch, 2007; Verhoef, Roelofs, & Chwilla, 2009), when they are predictable (Declerck, Koch, & Philipp, 2015), when bilinguals know exactly which words they will produce ahead of time (Declerck, Philipp, & Koch, 2013), and when they are voluntary (Gollan & Ferreira, 2009) rather than forced by an experimentally provided cue (Gollan, Kleinman, & Wierenga, 2015). Switch costs are found even when no overt switch is produced (e.g., when alternating between reading a word in one language and producing a word in the other; Peeters, Runnqvist, Bertrand, & Grainger, 2014). Cost free switches have been reported in just a few cases. In one study, switch costs were found in a task performed in both languages, but not if each task was performed in just one language (e.g., digit naming in one language, and picture naming in the other;

Finkbeiner, Almeida, Janssen, & Caramazza, 2006). In another, cost free switches were found when bilinguals switched languages voluntarily, but with an experimental requirement to use each language about equally often which led switching to become the default behavior (Gollan & Ferreira, 2009, Experiment 2). Finally, cost free switches were also found in voluntary switching when a small set of pictures was presented repeatedly and bilinguals chose (Gollan et al., 2015) or were instructed (Kleinman & Gollan, in press) to use just one language to name each picture and then consistently used only that same language on all subsequent appearances of each picture. Together these studies reveal both the persistence of switch costs in a variety of experimental settings, thereby resembling the literature on non-linguistic task-switching, but also the possibility that language switches are sometimes as efficient as (if not more efficient than) using just one language. Such cost-free switches might be easier to observe when presented with contextual support, a property that might be easier to study in the domain of language.

Grammatical structure—Sentence contexts, and connected speech as it is produced when multiple sentences are strung together, include grammatical as well as semantic structure that may facilitate control of language selection. Studies of spontaneous code switching corpora have observed that grammatical properties constrain the distribution of naturally occurring language switches in connected speech (e.g., Poplack, 1980; Muyksen, 2000; Myers-Scotton, 1997; 2005; 2006). However, experimental studies of language switches in sentence contexts have yielded inconsistent results. In one study, cued switches between sentences in unscripted connected speech (descriptions of actions shown in pictures) were found to be costly (Tar owski, Wodniecka, & Marzecová, 2013). Another study revealed speech to be slower when preceding naturally occurring code-switches relative to single-language speech (while controlling part of speech and utterance length; Fricke, Kroll, & Dussias, 2015). However, naturally occurring code-switches might sometimes be initiated to recover from access difficulties in one language, thereby masking the possibility that some switches are fully intentional and cost-free. In a different study, no switch costs were found but connected speech was not measured; sentence context was read silently, and only a single (highlighted) target word within each sentence was produced long after the language switch actually occurred (Gullifer, Kroll, & Dussias, 2013). Finally, another study had bilinguals (professional translators and highly proficient matched controls) read sentences one word at a time with self-paced button presses and exhibited switch costs only if they later had to repeat the sentence aloud (Ibáñez, Macizo, & Bajo, 2010).

Beyond simply asking participants to switch within sentences, some work has explicitly examined whether switch costs are modulated by the extent to which they match habitual, or henceforth *grammatical*, patterns of switching in natural language use. Some results suggest that the size of switch costs is influenced both by the exact location of the switch within the sentence, and the type of verb participating in the switch (Dussias, 2003; for similar evidence using fMRI see Rossi, Ting, Diaz, Newman, van Hell, & Dussias, submitted). Specifically, faster reading times (measured in gaze duration times) were observed for switches that occurred at a syntactic boundary (i.e., before the auxiliary; *terroristas have injured)* than between the auxiliary and the main verb (*terroristas han injured)*. However, this was found much more for closely bound syntactic elements *(haber + participle)*, than for less closely bound elements (i.e., *estar + participle*). The former is more bound to the participle in that it cannot occur on its own but *estar* can, and there appear to be stronger restrictions on the occurrence of switches that split *haber* than *estar* from the participle. These findings imply grammatical constraints on the magnitude of switch costs in bilingual sentence processing that mirror the distributional properties of naturally occurring switches.

However, a recent investigation of language switches in full sentence production provides more mixed support for the notion that switch costs are influenced by grammaticality. In this study, German-English bilinguals memorized and then repeatedly produced four different sentences with five words in them while switching languages on every other word, and with a forced 1500 ms pacing interval between words (Declerck & Philipp, 2015a). Sentences with shared word order in the two languages exhibited negligible switch costs (just 6 ms), whereas sentences with different word order in each language, or scrambled sentences, exhibited larger switch costs (36 and 40 ms respectively). On this basis the authors suggested that language switch costs can be abolished by sentence structure, but only in this very restricted special case (fully overlapping syntax; consistent with the equivalence constraint proposed by Poplack, 1980). Other syntactic constraints within this same study seemed to have no effect on switch costs. Specifically, switches that occurred on functional elements (e.g., *you, he, this*) were not more costly than switches that occurred on content words (e.g., *boy, flowers, shopping*), even though naturally occurring switches tend not to involve insertion of single functional elements (Muysken, 2000; Myers-Scotton, 1993).

Part of speech, specifically the contrast between functional elements (which signal grammatical relationships between words) and content words (which communicate more meaningful elements), appears to exert powerful but diametrically opposed constraints on intended versus *unintended* language switches (e.g., saying *pero* instead of *but*), or henceforth, cross-language *intrusion* errors. Unlike intentional switches (which, as just noted, rarely involve single functional elements), most spontaneously produced intrusion errors involve single function word targets (and interjections; Poulisse, 1999). Although bilinguals rarely lose control over language selection (Gollan, Sandoval, & Salmon, 2011), intrusions can be induced in large numbers by asking bilinguals to read aloud paragraphs that switch languages haphazardly (Kolers, 1966). As discussed in more detail below, based on data from eye-tracking and errors in this task, Gollan, Schotter, Gomez, Murillo, and Rayner (2014) argued that the majority of such intrusions arise within speech production processes. Gollan et al.'s (2014) results confirmed the vulnerability of function word targets to intrusion errors in production in Spanish-English bilinguals. Function words outnumbered content words as targets of intrusion errors, particularly for English targets in paragraphs with Spanish word order. Weaker effects were observed for English targets in English word order (and across paragraphs, Spanish targets exhibited relatively less prominent part of speech effects). The sensitivity of part of speech effects to word order implies that they are not simply an artifact of other non-syntactic differences between function and content words (e.g., function words might draw less attention because they tend to be shorter than content words; Poulisse & Bongaerts, 1994).

Author Manuscript

Author Manuscript

Author Manuscript

Executive functions—Some have suggested that bilingual language control is achieved, at least in part, by utilizing domain-general *executive control* processes (Abutalebi & Green, 2007; Hernandez, 2009; Kroll, Bobb, Misra, & Guo, 2008), including the same processes that support goal directed behavior outside the domain of linguistic functioning. This proposal has been supported by studies showing bilinguals are sometimes advantaged relative to monolinguals on tests of executive function (Bialystok, Craik, Green, & Gollan, 2009), with such advantages sometimes linked explicitly to language switching ability and frequency (Hartanto & Yang, in press; Prior & Gollan, 2011; Soveri, Rodriguez-Fornells, & Laine, 2011; Verreyt, Woumans, Vandelanotte, Szmalec, & Duyck, 2016), ability to avoid unwanted switches (Festman & Münte, 2012), or proficiency (Tao, Taft, & Gollan, 2015). However, the literature specifically comparing language switching and task switching often suggests that there is relatively little overlap in the underlying processing mechanisms (Calabria, Hernández, Branzi, & Costa, 2012; Gollan, Kleinman, & Wierenga, 2014; Prior & Gollan, 2013), and a number of investigators have been unable to replicate bilingual advantages (for review see Hilchey, Saint-Aubin, & Klein, 2014; Paap, Johnson, & Sawi, 2015; Valian, 2015).

Some of the most often cited support for the role of executive control in bilingual language selection has come from *language dominance* effects in studies of cued language switching. Most bilinguals have one language that is *dominant* i.e., relatively more proficient in many or most respects, than the other, non-dominant, language. Switch costs are often larger in the dominant language than in the non-dominant language, a result that could suggest inhibition of the dominant language when the non-dominant language is used, that must then be overcome to return to the dominant language (Meuter & Allport, 1999; for review see Declerck & Philipp, 2015b). These effects are most often attributed to inhibitory control mechanisms, specifically those that might also support non-linguistic task-switching (Green, 1998; Meuter & Allport, 1999; see also Philipp, et al., 2007; Philipp & Koch, 2009; but see Bobb & Wodniecka, 2013). Consistent with strong inhibition of the dominant language in mixed language contexts, under a limited set of conditions fully *reversed* language dominance effects – in which bilinguals perform less well in their dominant than in the otherwise non-dominant language — have been observed. Gollan et al. (2014) reported that intrusion errors exhibited such reversed dominance effects; words in the dominant language were replaced by words in the non-dominant language more often than vice versa. Reversed dominance effects have also been found in a handful of studies that measured timed picture naming, with both cued (Christoffels, Firk, & Schiller, 2007; Costa & Santesteban, 2004; Verhoef et al., 2009), and voluntary, language switches (Gollan & Ferreira, 2009).

The Current Study

In the current study we investigated the role of these factors in controlling intended and unintended language switches using the read-aloud task. Both previous studies that used the read-aloud task contained paragraphs with haphazard language switches (Gollan et al., 2014; Kolers, 1966). Here, we asked if bilinguals would still produce intrusion errors when switches followed more typically occurring patterns of bilingual language use. Spanish-English bilinguals read paragraphs that manipulated the degree to which switches conformed to naturally occurring patterns of language switching. We hypothesized that grammatical switches might allow bilinguals to rely on automatic control mechanisms that function in conversational speech between bilingual interlocutors, and therefore might elicit more rapid and also more accurate production of mixed language speech. Because previous evidence concerning the role of grammaticality on switching produced mixed results, we also manipulated the rate of switches, crossing grammaticality with switch rate so that paragraphs either had mostly grammatical or mostly ungrammatical switches, and either a low or a high switch rate (with twice as many switches in high as in low). To provide a further index of grammatical influences on switching, we examined part of speech effects. Based on previous work (Gollan et al., 2014; Kolers, 1966; Poulisse, 1999), we anticipated we would observe greater control difficulties on function vs. content words – but of particular interest here – was whether grammatical encoding would affect function more than content word switches, or perhaps function but not content word switches.

Importantly, though the read-aloud task elicits production via reading, and therefore aspects of the task will at least partially reflect mechanisms underlying reading comprehension, intrusion errors produced in this task primarily (if not exclusively) reflect the mechanisms of speech production (Gollan et al., 2014). Indeed Kolers (1966) found that language mixed text had no effect on comprehension measures, but robust effects on production times and errors produced in the read-aloud task. Several aspects of the errors produced in the read-aloud task supported this conclusion. First, most intrusions were produced with targets that do not share sounds or letters with their translation equivalent targets (e.g., *pero*—*but*). The production of intrusion errors even in the presence of substantial formal differences between translations, suggests that bilinguals are not simply mis-reading the paragraphs when they produce these errors. Most importantly, eye-tracking measures in Gollan et al. (2014; see Table 4–5, Figure 3) revealed that although bilinguals skipped function more than content words when reading mixed-language paragraphs, in the majority of cases bilinguals were fixating the target word when they produced the intrusion error, and part of speech effects on intrusion errors remained even when limiting the analysis to targets that were never skipped, again suggesting independence between mechanisms underlying reading processes and production of intrusion errors in the read-aloud task. Finally, if the read-aloud task elicited intrusion errors via different processing mechanisms than those that underlie intrusions in more naturalistic speech, then it might be expected that the read-aloud task would elicit a greater rate of intrusions than found in spontaneous speech. However, intrusions were produced in the read-aloud task approximately or just under 1% of the time, which is the same rate as reported for spontaneous speech (Poulisse, 1999), and in other experimental tasks that measured production more purely (e.g., the verbal fluency task; Gollan et al., 2011).

The read-aloud task appears to elicit speech production errors via the same, or similar processes that give rise to errors in spontaneous speech production. Specifically, we assume that multiple representations are activated in parallel during formulation of the planned utterance. Errors arise because random noise during processing allows non-target representations to become more active than the target (e.g., Dell, 1986). In this task, the code mixed utterance presented for reading is the most strongly activated target representation, along with co-activated representational elements from alternative code mixed utterances (e.g., including translation equivalents of the target words). When the target is weakly activated—e.g., ungrammatical, haphazard mixing sentences that do not fit with the bilingual speaker's experience—random noise is more likely to give rise to the production of alternative utterances (leading to intrusion errors). Similar phenomena have been observed in monolingual language production, in which phonological structures that violate constraints on sound structure are more likely to result in speech errors than those that respect such constraints (see Goldrick, 2011, for a review).

While such mechanisms provide one means of generating intrusion errors within production processes, other aspects of eye-tracking measures reported in Gollan et al. (2014; Figure 3) did suggest that function words are more vulnerable to failures of language control at least in part because they draw less attention than content words (e.g., function, but not content words, were more likely to involve an intrusion error if bilinguals were looking at words in the nontarget language when they produced the error). To address the possible contribution of reading processes to the results in the present study, we included part of speech as a factor in our analyses; since readers skip function words more than content words – part of speech effects should be differentially sensitive to grammaticality manipulations if these are driven by reading rather than by production mechanisms.

In parallel to analysis of these language-specific factors—which we assumed to be reflecting relatively automatic control processes specialized for language processing — we examined the extent to which language selection might also be driven by more attention demanding control processes subject to individual differences in domain general control abilities. To this end, we measured individual differences in three commonly administered measures of executive control. We hypothesized that bilinguals with stronger general cognitive control ability might also exhibit more efficient ability to mix languages in the read-aloud task. Of particular interest, however, was if individual difference measures would interact with our manipulation of grammaticality, as well as part of speech and language dominance effects. The presence of such interactions would suggest that bilingual language control mechanisms are fully integrated with, and broadly reliant on, domain general control mechanisms, whereas their absence would imply the existence of some language specific control mechanisms that function relatively independently. In Experiment 1 we tested young Spanish-English bilinguals, and in Experiment 2 we tested a group of older Spanish-English bilinguals on the same tasks that were administered in Experiment 1, to provide a second test of the individual differences question. We hypothesized that interactions between executive control measures and grammaticality might be most likely to emerge in older bilinguals who exhibit a broader range of individual differences in executive control ability, including some individuals with relatively impaired executive control ability.

Experiment 1 – Young Bilinguals

Method

Participants—Ninety-seven Spanish-English bilinguals participated in the study in exchange for course credit through undergraduate classes in the Psychology department at the University of California, San Diego (UCSD). Table 1 shows self-reported participant characteristics and ability to name pictures in each language on the Multilingual Naming Test or MINT (see below; Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012).

Materials and procedure—Stimuli were presented using PsyScope X software (Build 57; Cohen, MacWhinney, Flatt, & Provost, 1993; http://psy.ck.sissa.it) on an iMac 7 computer with a 20-inch color monitor. On each trial an entire paragraph appeared centered on the screen and participants were instructed to read the paragraph aloud as accurately as possible at a comfortable pace. Each bilingual read 16 paragraphs, four in each of four conditions: (a) grammatical, low-switch; (b) grammatical, high-switch; (c) ungrammatical, low-switch; (d) ungrammatical, high-switch. On average, paragraphs had 109 (*SD*=10) words (range 88–134 words), and on average 72% (*SD*=3%) were in English (and the rest in Spanish).

Paragraphs were modified from Gollan et al., (2014) so that all were presented in English word order, and varied with respect to switch rate and grammaticality of switches. An example of each paragraph is presented in the Appendix. Low switch paragraphs had 13 language switches in them, and high switch paragraphs had twice as many (i.e., 26 switches). A native Spanish-English bilingual adapted the paragraphs (which were originally composed to contain "haphazard" language switches) to manipulate switch rate and grammaticality without following any specific rules about what types of switches are or are not natural. A second native Spanish-English bilingual read through the paragraphs and confirmed the grammaticality manipulation (any disagreements were discussed and settled and paragraphs modified accordingly). Note that classifying switches as grammatical or not is a topic of some debate in the field (see e.g., Bullock & Toribio, 2009, for discussion). Our manipulation here was deliberately not restricted to one type of constraint on switching, and was intended as a first step towards examining the role of naturalness in facilitating control over language selection in production of connected speech.

However, to consider the extent to which the grammaticality manipulation might generally match theoretical constraints on switching we compared grammatical and ungrammatical paragraphs in frequency of occurrence of three constraints. First, we asked if ungrammatical paragraphs violate the Equivalence Constraint (Poplack, 1980), a structural constraint on language switching. Switches are grammatical, satisfying this constraint, if the word order matches across languages (e.g., a switch on a noun in a phrase like *the man* would be more acceptable than a switch in a *the old man* because the Spanish translation equivalents would appear in the same order *el hombre* for the former but not in the latter *el hombre viejo* which word for word translates to *the man old*). To consider this possibility we identified all phrases in the paragraphs with a Determiner Modifier Noun structure (e.g., *an old car*, *an old carro*), and then counted the number of switches on the noun. We excluded those cases with pronominal adjectives in Spanish (e.g., *misma*) along with their translation equivalents. We also excluded two examples where Spanish words were in the incorrect word order (e.g, *su fresca sombra*). These counts revealed that violations of the equivalence constraint occurred just 3 times in grammatical paragraphs (the total number of opportunities for switches in such constructions was n=47, for a violation rate of 6.4%), but 9 times in ungrammatical paragraphs (with n=45 and a violation rate of 20%); thus, ungrammatical paragraphs violated this particular equivalence constraint three times as often as grammatical paragraphs.

Though these numbers are relatively small, each subsequent comparison revealed a similar pattern. Another commonly accepted constraint on switching was mentioned above; this is that switches on a single function word are rare (Muysken, 2000; though, paradoxically, these are the most common targets of naturally occurring intrusion errors, Poulisse, 1999). To consider if such switches occurred more often in ungrammatical than in grammatical paragraphs, we counted the number of single function word switches (e.g., Spanish word, English function word, Spanish word). About 1.9% of all function word switches in grammatical paragraphs fit this profile (9/483 switches on function words), and more than twice as often, or 4.2% in (13/309) in ungrammatical paragraphs. Finally, it is well known that switches on single isolated words most commonly involve nouns (Myers-Scotton, 2002; Muysken, 2000), and much less often involves verbs (which may be switched within a sentence less often because they are less easily integrated into syntactic structure across different spoken languages; Myers-Scotton & Jake, 1995; but see Emmorey, Borinstein, Thompson, & Gollan, 2008, for contrasting patterns in speech-sign bilinguals). In ungrammatical paragraphs singleton switches on verbs (n=16/111 total verb switches, 14.4%) outnumbered singleton noun switches (n =10/173, 5.8%), whereas in grammatical paragraphs singleton switches on nouns (n=27/65, 41.5%) outnumbered singleton switches on verbs (n =15/63, 23.8%), as would be expected in naturally occurring code-switches. Thus, though grammatical paragraphs had some switches that violated constraints on switching proposed in the literature, for each and every comparison considered, there was a greater number of violations in ungrammatical than in grammatical paragraphs. Thus, the intuitions of the bilinguals who designed the paragraphs for the present study, generally match what has been reported in the literature for characterizing commonly occurring switches.

Paragraphs appeared in one of four different fixed orders with conditions counterbalanced across subjects and items in a Latin square design. Within each fixed order, all four conditions were presented on each group of four consecutive trials but in different condition orders across subjects so that the average list position of paragraphs was the same for every condition, and participants never saw paragraphs in the same condition on consecutive trials. Each trial began with a fixation point (+) that appeared at the location on the screen where the first word appear and remained on the screen until the participant pressed the space bar, after which the fixation was replaced by the paragraph which remained on the screen until the participant pressed the space bar again. A 1,500 ms ITI was presented and then replaced by the fixation point for the following trial. Before beginning bilinguals completed 4 shorter practice paragraphs (approximately 30 words in each) including one of each type of paragraph (counterbalanced in the same way as the critical paragraphs).

To provide an objective measure of English and Spanish proficiency, all bilinguals completed the Multilingual Naming Test (MINT; Gollan et al., 2012). This consists of 68 black-and-white line drawings, administered in order of progressing difficulty (e.g., item #1 is *hand*, and item # 68 is *axle*). This test was designed to assess picture-naming ability in four languages (English, Spanish, Mandarin, Hebrew), and also both dominant and non-dominant language proficiency.

A number of tests of executive function were also administered to provide an objective measure of domain-general executive control ability. Choice of tests was motivated by previous reports of an association between bilingualism and performance on these measures (for review see Bialystok et al., 2009). Means (and standard deviations) by condition for these tasks are shown in Table 1, which also shows that all three tests exhibited significant interference effects typically reported for these measures.

Stroop (Stroop, 1935): Target stimuli consisted of four color names (*black*, *blue*, *green*, and *red* in English; *negro*, *azul*, *verde*, and *rojo* in Spanish) and four neutral words (*big*, *prime*, *deep*, and *legal* in English; *boda*, *dulce*, *mesa*, and *usar* in Spanish). Each bilingual completed the task in both languages in counterbalanced order between participants as described below. The use of neutral words as a baseline and unblocked presentation of trial types (congruent, neutral, incongruent; see below) was based on the recommendations of Spieler, et al. (1996). There were 36 congruent trials, 36 incongruent trials, and 32 neutral trials. In the congruent trials, each of the four color names appeared nine times in its corresponding color. In the incongruent trials, each color name appeared three times in each of the three non-matching colors. In the neutral trials, each of the four neutral words appeared twice in each of the four colors. Sixteen randomized lists were created for each language, with the restriction that a word or color was not repeated more than twice on consecutive trials. Participants were instructed to produce the name of the color of ink in which the words were presented. On each trial, the target stimulus was presented until a response was recorded. The interval between trials was 500 ms. Before the critical trials, participants completed 16 practice trials in the same language as the critical trials. Participant responses were coded for accuracy online by an experimenter, and were also audio-taped for subsequent verification of accuracy.

Trail Making Test A and B: (TMT A and TMT B; from the Halstead Reitan Neuropsychological Test Battery; see Reitan, 1958; cf. Mickes et al., 2007): In Part A (TMT A), participants draw a line to connect the numbers 1–25 in consecutive order as quickly as possible within a 150 second time-limit. In Part B (TMT B), participants draw a line to connect 25 numbers and letters in alternating, consecutive order as quickly as possible within a 300 second time-limit. Time-to-complete each task is scored.

Flanker task: The flanker task was adapted from Fan et al. (2002). Stimuli consisted in a central arrow with two flankers on each side, and appeared in three conditions: congruent (central arrow pointing in the same direction as the flanker), incongruent (central arrow pointing in the opposite direction of the flanker), and neutral (central arrow flanked by lines without arrowheads). Direction of arrows was counterbalanced in each condition. There were 16 trials in each condition. Trials began with a central fixation point for 900 ms, then the target stimulus replaced the fixation point until a response was made, and there was an ITI of 1000 ms. Participants pressed a button on the left or right of a button box to indicate the direction of the central arrow. Practice blocks began with six neutral trials, followed by six congruent trials, then six incongruent trials, and ended with six trials with equal numbers of the different trial types presented in a random order. Feedback “correct” or “incorrect” was included for practice trials only.

The paragraph reading task was completed first after which bilinguals completed the remaining tasks in a counterbalanced order. Some bilinguals completed English MINT and English Stroop next or Spanish MINT and Spanish Stroop next (counterbalanced between participants). The two different language versions of MINT and Stroop were always separated in time by the Flanker Task or by the Trail Making Test (see below). An equal number of bilinguals completed Flanker or Trails first (counterbalanced between participants), followed by English or Spanish Stroop and MINT separated by whichever executive task they had not yet completed.

Results

Following the methods of Gollan et al., (2014) a native Spanish-English bilingual research assistant transcribed the errors which were classified as intrusions ($n = 899$) e.g., saying *el* instead of *he*), partial intrusions ($n = 208$; starting to produce an intrusion but self-correcting before producing the error), accent errors (e.g., $n = 844$; saying the correct word with the accent of the non-target language), or within-language errors ($n = 2{,}271$; e.g., saying *such* instead of *much*). All but one bilingual produced at least one intrusion and up to as many as 24 (M=9; SD=5). Accent errors are necessarily subjective (see also Kolers, 1966), and other error types are difficult to interpret given indeterminacy with respect to error content (i.e., omissions, $n = 291$ and insertions, $n = 222$); thus we focused our analyses primarily on intrusions and within-language errors. A small number of accent errors ($n = 5$) and intrusion errors ($n = 6$) were classified as more than one type of error (e.g., accent and within, or intrusion and accent). These cases were classified by default as intrusions rather than as within-language errors, with only one type of language control failure coded if both intrusion and accent error were produced. For example, if the intended target was *monastery* and the speaker said *monasterio* and then self-corrected to *monastery* but produced the English word with a Spanish accent, these would be counted as a single intrusion error.

Partial intrusions, i.e., intrusion errors self-corrected in mid-utterance, exhibited a noticeably different pattern from other intrusion errors. As discussed below, intrusions in this paradigm (as in naturally occurring errors) are more likely to target function than content words. Partial intrusion errors exhibit the opposite pattern. At switch sites (the primary location of intrusion errors; see below), partial intrusions occurred on 1.2% of content words (95% CI 0.9%, 1.4%), significantly higher than the rate of 0.1% of function words (95% CI 0.008%, 0.2%). This likely reflects the relatively difficulty of monitoring for errors on function vs. content words (Gollan et al., 2014; Poulisse & Bongaerts, 1994). Given this clear difference in sensitivity to word class across the two error types, we excluded partial intrusions from our analysis of intrusion errors.

Analysis methods—Behavioral data were analyzed using mixed effects regressions (linear for reading times and logistic for error rates; Jaeger, 2008). All categorical predictors (e.g., switching type, grammatical vs. ungrammatical) were contrast coded, and continuous predictors (e.g., Stroop interference) were centered. Random intercepts were included for participants; more complex random effects structures failed to converge for some analyses, so this simplified random effects structure was used in all cases. Significance was assessed via model comparisons (Barr, Levy, Scheepers, & Tily, 2013).

Whole Paragraph Reading times—We first examined how whole paragraph reading times were affected by switching rate (low vs. high), and paragraph type (grammatical vs. ungrammatical), along with the interactions of these factors. As shown in Figure 1, bilinguals read paragraphs with ungrammatical switches more slowly than those with grammatical language switches ($\beta = 3.75$, $SE\,\beta = 0.26$, $\chi^2(1) = 199.67$, $p < .0001$), and paragraphs with a high switch rate more slowly than those with a low switch rate ($\beta = -2.16$, $SE\,\beta = 0.26$, $\chi^2(1) = 70.08$, $p < .0001$). Reading times were especially slow in ungrammatical paragraphs with a high switching rate; i.e., switching rate and paragraph type interacted ($\beta = -1.81$, $SE\,\beta = 0.51$, $\chi^2(1) = 12.53$, $p < .0005$). Follow-up regressions at each switching rate showed that while grammaticality had a significant effect on readings time in both types of paragraphs, grammaticality effects were stronger at high switching rates ($\beta = 4.64$, SE $\beta = 0.34$, $\chi^2(1) = 159.44$, $p < .0001$) than at low switch rates ($\beta = 2.83$, SE $\beta = 0.35$, $\chi^2(1) = 60.77$, $p < .0001$). Additional, follow-up regressions within each paragraph type showed that while there were switch costs on readings times in both types of paragraphs (i.e., a significant effect of switch rate), switch costs were stronger in ungrammatical ($\beta = -3.07$, $SE\,\beta = 0.35$, $\chi^2(1) = 72.21$, $p < .0001$) than in grammatical paragraphs ($\beta = -1.26$, $SE\,\beta = 0.34$, $\chi^2(1) = 13.94$, $p < .0005$).

Intrusion Errors—Bilinguals were much more likely to produce intrusion errors at switch sites (mean: 1.4%, 95% CI 1.2%, 1.6%) than at non-switch positions in the paragraphs (mean: 0.3%, 95% CI 0.3%, 0.4%). We therefore focused our analysis at switch points, analyzing the likelihood of intrusion errors with a logistic mixed effects regression. In addition to the experimental factors (grammaticality and switch rate), we examined if intrusions were modulated by part of speech, and allowed part of speech to interact with grammaticality (to examine if grammaticality effects were driven exclusively or primarily by function vs. content words). A random intercept was included for participants.

As shown in Figure 2, bilinguals produced more intrusion errors when reading paragraphs with ungrammatical than with grammatical switches ($\beta = 0.66$, $SE\,\beta = 0.12$, $\chi^2(1) = 30.6$, $p < .0001$). While there was no main effect of switch rate ($\beta = 0.14$, $SE\,\beta = 0.11$, $\chi^2(1) = 1.62$, $p < .21$), there was a significant interaction between rate and grammaticality ($\beta = -0.74$, $SE\,\beta = 0.21$, $\chi^2(1) = 11.5$, $p < .001$). Follow-up regressions at each switching rate revealed grammaticality effects on intrusions only in paragraphs with high switch rates ($\beta = 1.13$, $SE\,\beta = 0.16$, $\chi^2(1) = 61.1$, $p < .0001$), whereas in paragraphs with low switch rates grammaticality effects were not significant ($\beta = 0.14$, $SE\,\beta = 0.19$, $\chi^2(1) = 0.58$, $p < .45$). Additional follow-up regressions within each paragraph type revealed switch costs in ungrammatical paragraphs, such that bilinguals produced more intrusion errors in high than in low switch paragraphs ($\beta = -0.26$, $SE\,\beta = 0.13$, $\chi^2(1) = 3.89$, $p < .05$), whereas in grammatical paragraphs, switch costs were absent, in fact there was a significant effect in the unexpected direction (i.e., more intrusion errors in low than in high switch paragraphs; $\beta = 0.51$, $SE\,\beta = 0.17$, $\chi^2(1) = 8.53$, $p < .005$). We do not interpret this effect any further, given the presence of a significant effect in the opposite direction in the reading times (see above), possibly suggesting a speed-accuracy trade-off.

There was a nonsignificant trend for bilinguals to produce more errors on function words (mean 1.5%, 95% CI 1.2%, 1.7%) compared to content words (mean 1.3%, 95% CI 1.1%, 1.6%; $\beta = 0.20$, $SE\ \beta = 0.12$, $\chi^2(1) = 2.82$, $p < .10$). The absence of robust part of speech effects on intrusion errors resembles results reported for English word order paragraphs as used here and in Gollan et al., (2014), and contrasts notably with highly robust part of speech effects on within-language errors (see below). Crucially, the effect of grammaticality was not significantly stronger for function vs. content words ($\beta = 0.20$, $SE\ \beta = 0.12$, $\chi^2(1) = 2.82$, $p < .10$[1]).

To examine if intrusion errors were modulated by the language of production, we repeated the analysis above, focusing on the 88 English-dominant speakers (those with higher English than Spanish MINT scores). Language of production was included as an additional predictor, interacting with part of speech (to examine if dominance effects were driven by function vs. content words). As shown in Figure 3, there was a significant reversed-dominance effect; bilinguals were more likely to produce intrusions in the dominant language (replacing written English words with spoken Spanish translations), than in the non-dominant language (replacing written Spanish words with spoken English translations; $\beta = 0.41$, $SE\ \beta = 0.11$, $\chi^2(1) = 13.7$, $p < .0005$). Interestingly (and contrasting with a result reported below for within-language errors), this did not interact with part of speech ($\beta = 0.22$, $SE\ \beta = 0.22$, $\chi^2(1) = 0.98$, $p < .35$).

Though our primary focus was on language switching ability and language intrusion errors, as a point of contrast we repeated our analyses of intrusion errors with errors produced within a single language as the dependent measure. To the extent that our manipulations targeted language control specifically, they should have no effect on within-language errors.

Within-Language Errors—Within-languages errors were equally likely to occur at switch sites (mean: 1.3%, 95% CI 1.1%, 1.6%) and other positions in paragraphs (mean: 1.3%, 95% CI 1.2%, 1.5%). We therefore analyzed errors at all points. Regression analyses were structured following those above.

We first examined the experimental factors (grammaticality and rate) and part of speech. As shown in Figure 4, bilinguals were no more likely to produce within-language errors when reading paragraphs with ungrammatical than with grammatical switches ($\beta = -0.02$, $SE\ \beta = 0.04$, $\chi^2(1) = .24$, $p < .65$). There was a non-significant trend for more errors to occur at high switch rates ($\beta = 0.07$, $SE\ \beta = 0.04$, $\chi^2(1) = 2.83$, $p < .10$). In contrast with the results reported above for intrusion errors, there was no significant interaction between rate and grammaticality ($\beta = 0.11$, $SE\ \beta = 0.08$, $\chi^2(1) = 1.58$, $p < .21$). Also in contrast with results reported for intrusions, bilinguals produced within-language errors much *less* frequently with function word targets (mean 0.9%, 95% CI 0.7%, 1.0%) compared to content word targets (mean 2.1%, 95% CI 1.8%, 2.4%; $\beta = -0.92$, $SE\ \beta = 0.04$, $\chi^2(1) = 458.39$, $p < .0001$). However, as for intrusion errors, there was no interaction of part of speech with grammaticality ($\beta = -0.09$, $SE\ \beta = 0.09$, $\chi^2(1) = 1.02$, $p < .32$).

[1]Given the marginal interaction, follow-up regressions were performed for each part of speech. These confirmed a reliable effect of grammaticality for both content ($\beta = 0.45$, $SE\ \beta = 0.21$, $\chi^2(1) = 4.79$, $p < .03$) and function words ($\beta = 0.84$, $SE\ \beta = 0.13$, $\chi^2(1) = 6.44$, $p < .0001$).

We then repeated the analysis above for the 88 English-dominant speakers, including language of production as an additional predictor, interacting with part of speech. As shown in Figure 3, there was a significant dominance effect; bilinguals were less likely to produce within-language errors in the dominant language (English), than in the non-dominant language (Spanish; $\beta = -0.92$, $SE\ \beta = 0.05$, $\chi^2(1) = 338.62$, $p < .0001$). Additionally, unlike results reported above for intrusions, language dominance effects on within-language errors were stronger for content than for function words, i.e., dominance interacted with part of speech ($\beta = 0.28$, $SE\ \beta = 0.09$, $\chi^2(1) = 9.54$, $p < .005$). Follow-up regressions showed that while there was a significant dominance effect for both parts of speech, the effect was stronger for content ($\beta = -1.03$, $SE\ \beta = 0.06$, $\chi^2(1) = 315.28$, $p < .0001$) as compared to function words ($\beta = -0.74$, $SE\ \beta = 0.07$, $\chi^2(1) = 97.32$, $p < .0001$).

To further investigate these effects, we re-ran these regressions, incorporating a continuous measure of each participant's language dominance (the ratio of the English to Spanish MINT scores), allowing this to interact with the factors above. The continuous dominance measure not only modulated the overall effect of language of production (such that the tendency for within-language errors to occur on Spanish vs. English targets was stronger for more English-dominant participants; $\beta = -0.73$, $SE\ \beta = 0.11$, $\chi^2(1) = 49.46$, $p < .0001$); it also modulated the interaction between language of production and part of speech ($\beta = 0.56$, $SE\ \beta = 0.21$, $\chi^2(1) = 7.36$, $p < .01$). Follow-up regressions revealed that the general effect of the dominance measures—the tendency for more English-dominant participants to produce more within-language errors on Spanish vs. English targets—was much stronger for content ($\beta = -1.03$, $SE\ \beta = 0.14$, $\chi^2(1) =63.39$, $p < .0001$) vs. function words ($\beta = -0.42$, $SE\ \beta = 0.15$, $\chi^2(1) = 7.59$, $p < .01$).

Executive function effects: We next examined the extent to which individuals with relatively poor performance on executive function tasks had greater difficulties with language control. In particular, we were interested if relatively poor executive function abilities enhanced the effects observed above; such interactions would support the integration of language-specific and domain general control mechanisms.

To address these issues, we constructed three separate new regression models, with each one including one of three executive function measures:

- *Stroop*: The language with the faster mean reaction times (RTs) on neutral trials was used to calculate each participant's degree of Stroop interference, relative to neutral trial RT: (mean RT on inconsistent trials – mean RT on neutral trials)/mean RT on neutral trials)
- *Flanker*: Similar to Stroop, this measure was interference relative to the neutral trial baseline: (mean RT on inconsistent trials – mean RT on neutral trials)/mean RT on neutral trials)
- *Trails*: Increase in completion time for Part B relative to Part A: Trail Marking Test Part B completion time/Trail Marking Test Part A completion time

Each (centered) executive function measure was included as a main effect, allowing us to examine its contributions to language control. To examine whether language control mechanisms are fully integrated with domain-general mechanisms, we also allowed each measure to interact with the significant effects observed above: grammaticality and its interaction with rate. Two participants were excluded from these analyses for failure to complete one of the executive function tasks. The results of the regressions for the 95 remaining bilinguals are summarized in Table 2.

If language control relies on domain-general mechanisms, individuals with lower executive function performance should show a greater rate of intrusion errors. There is some evidence consistent with this. While there were no main effects of either Stroop or Flanker interference, individuals with relatively worse performance on Trails produced more intrusion errors. To consider if this effect was specific to language control, we examined whether any of these executive function measures were correlated with the rate of within-language errors; as shown in Table 2, there were no significant effects.

If language control mechanisms are fully integrated with domain-general control mechanisms, we expect these differences in executive function to interact with grammaticality. There is some evidence consistent with this. Individuals with greater Stroop and Trails interference effects showed exaggerated effects of grammaticality on whole paragraph reading times. However, there were no such interactions in our analyses of intrusion errors. Furthermore, there was a marginal interaction in the unexpected direction for Flanker performance (such that individuals with greater Flanker interface showed decreased effects of grammaticality).

Discussion

The results of Experiment 1 revealed a number of key findings. First, grammaticality of switches had robust effects on both reading times and production of intrusion errors; this implies that grammatical encoding plays an important role in helping bilinguals maintain control over language selection. We also observed robust switch rate effects, so that switch costs increased with switch rate, and interactions with grammaticality, so that bilinguals found it particularly difficult to produce larger numbers of switches if those switches were also ungrammatical. Importantly, grammatical switches were not cost free; though intrusions did not reveal switch costs in grammatical paragraphs (they revealed an unexpected effect in the opposite direction), bilinguals' whole paragraph reading times were significantly slower for high- than for low-switch paragraphs. Additionally, intrusion errors exhibited robust part of speech effects, consistent with theories that posit distinct control mechanisms for retrieval of targets that assign thematic roles versus those that do not (Myers-Scotton, 2006), the finding of robust part of speech effects in naturally occurring intrusion errors (Poulisse, 1999), and inconsistent with the absence of part of speech effects in other recently reported experimental paradigms (Declerck & Philipp, 2015a).

In contrast, within language errors exhibited no sensitivity to our language control manipulations, including grammaticality and switch rate. Also replicating previous work (Gollan et al., 2014), intrusions and within-language errors exhibited different part of speech and language dominance effects. Function word targets were equally likely to elicit intrusions as content words, but were far less likely to elicit within language errors relative to content words, and language dominance effects were significantly reversed for intrusions relative to within-language errors. Moreover, of great interest, in the current study, part of speech effects interacted with language dominance in our analysis of within-language errors (dominance effects were driven more by content than by function words), but not in our analysis of intrusion errors. These contrasting effects across different error types imply different underlying processing mechanisms, a point that will be elaborated in the General Discussion section.

Finally, our analyses of individual differences in executive control ability revealed that bilinguals who exhibited greater interference effects on the Trails measure produced a greater number of intrusion errors but no interaction with grammaticality implying independence of language-specific and domain-general control mechanisms. Whole paragraph reading times revealed some evidence for integration of control mechanisms; bilinguals who exhibited greater interference effects in both Trails and Stroop tests were particularly slow to read paragraphs with ungrammatical switches in them. However, we defer further discussion of these contrasting patterns to the General Discussion, after we present the results of Experiment 2.

Experiment 2 – Older Bilinguals

Our analyses in Experiment 1 revealed some evidence for language control mechanisms, related to grammatical encoding, that facilitate language switches in mixed-language connected speech, but some ambiguity as to whether they function independently of general executive control or not. In these analyses we assumed that there is a continuous relationship between domain general control and language switching. However, an alternative possibility is that language control does systematically rely on domain general control ability, but that only a minimal degree of executive control is recruited for this job of controlling language selection. On this view, a constant relationship between domains would be found regardless of individual differences in executive control, and evidence for interactions would be especially obvious when executive control ability dipped below the minimal level needed to serve language switching. Consistent with this possibility, some have suggested that bilingualism enhances executive control ability, but that this bilingual advantage is more apparent in older age, when individuals are no longer at peak levels of functioning in executive control ability (Bialystok et al., 2009; see Valian, 2015 for a comprehensive discussion; see Hilchey et al., 2014; Paap, et al., 2015 for a differing perspective). To test this possibility, in Experiment 2 we tested a group of older bilinguals with the same materials and procedure as in Experiment 1 – aiming to answer two key questions: (a) Do aging bilinguals exhibit significantly larger grammaticality effects, and (b) do aging bilinguals reveal stronger evidence for a relationship between domain-general executive control and language control?

We anticipated that older bilinguals would exhibit a broader range of abilities in executive function measures than in younger bilinguals. Although there is some disagreement as to the underlying cognitive mechanism (Salthouse, 1996, 2010; Verhaegen, 2011), it is widely reported that older individuals have more difficulty with measures of executive control ability, including the measures used herein. Consistent with this view, the prefrontal cortex, which is thought to control planning of complex behaviors, declines earlier, and more rapidly, than other parts of the brain as part of the normal aging process (see reviews in Raz, 2000; West, 1996). Additionally, aging related declines affect cognitive processes related to inhibition; e.g., older adults appeared to lose the ability to suppress activity in the 'default mode network' during word generation (Meinzer et al., 2012). Possibly related aging deficits are found in measures of selective attention, including Stroop and Flanker tasks (Spieler, Balota, & Faust, 1996; West & Baylis, 1998). According to the *Inhibitory Deficit Hypothesis*, activation of possible responses to stimulus cues remains intact, but aging leads to specific impairments in the ability to inhibit irrelevant information from attention and working memory, and difficulty with suppression of inappropriate responses (Hasher & Zacks, 1988; Zacks & Hasher, 1994).

The notion that aging leads to deficits in inhibitory control is not universally accepted, and one problem with trying to answer this question is that the construct of inhibition is defined differently by different investigators and in different subfields (Aron, 2007). Empirically, the hypothesis is difficult to test because there is disagreement as to which tasks constitute valid measures of inhibition (Burke & Osborne, 2007). In the context of language processing, attempts to test specific versions of the inhibitory deficit hypothesis have failed to confirm its predictions (Burke, 1997, 1999; Burke & MacKay, 1997). Experiment 2 aimed to contribute unique evidence to these long-standing questions by investigating aging effects on bilingual language control. This may offer a more ecologically valid measure of inhibitory control than those typically investigated. As discussed in the introduction, dominance reversal effects have been attributed to the use of inhibitory control in bilingual language processing (Gollan & Ferreria, 2009; Kroll, et al., 2008); assuming this analysis is correct, such effects may offer a more naturalistic and powerful way to measure age related decline in inhibitory control.

Thus, in Experiment 2 we hypothesized we might find (a) larger effects of grammaticality for older than for younger bilinguals. Several possible factors could give rise to such effects. For example, if grammaticality is related to the degree of experience with language switching, a longer lifetime would provide greater experience and expectations that switches should follow typical patterns of language use. Greater grammaticality effects might also be expected in aging bilinguals if reduced ability to rely on domain-general executive control to plan, produce, and monitor language switches) leads older bilinguals to rely more heavily on language specific control mechanisms. Alternatively, if typical switch patterns arise because some types of switches are motivated by syntactic structure, and therefore easier than others, older bilinguals have more difficulty with difficult tasks in a generalized way (but see Gollan et al., 2008). For similar reasons, older bilinguals might be expected to show larger part of speech and language dominance effects. Finally, we also predicted that (b) older bilinguals would show greater evidence for interaction between executive function measures and our experimental manipulation of grammaticality, on the assumption that such interactions become apparent only when minimal levels of executive control ability are absent.

Previous studies of aging and language control—Several studies have examined switch costs in out-of-context speech and revealed switching deficits in aging. For example, older bilinguals were more likely to fail to switch when cued to do so than young bilinguals (Hernandez & Kohnert, 1999). A more recent study confirmed the presence of larger cued language switching costs for older than for young bilinguals (Weissberger, Wierenga, Bondi, & Gollan, 2012), but also demonstrated that age-related slowing was far smaller in language switching than it was in a comparable measure of non-linguistic task-switching. Another related study used cued switching tasks configured a bit differently relative those in Weissberger et al. (2012), and found no aging related increase in language switching costs, but greater task-switching costs in older than in younger bilinguals (Calabria, Branzi, Marne, Hernández, & Costa, 2015). Though the precise pattern of results in the two studies was different, in both cases dissociations between aging effects across linguistic and non-linguistic domains were reported, and language control appeared to be relatively more intact than non-linguistic control in aging bilinguals. This in turn implies that at least partially different cognitive mechanisms support non-linguistic task-switching versus language-switching. Finally, very limited aging effects were found in another study, in which bilinguals were instructed to choose for themselves on each trial which language to use to name the pictures. If the "bilingual switch mechanism" were impaired in aging, then older bilinguals should have chosen to switch less often than young bilinguals. Instead, older bilinguals chose to switch voluntarily as often as young bilinguals, and though they responded more slowly overall than young bilinguals, switch costs were similarly sized across age groups (Gollan & Ferreira, 2009). This again implies relatively intact ability to produce intended language switches in aging bilinguals.

Another study examined unintended language switching in young and older bilinguals and produced mixed results (Gollan, Sandoval, & Salmon, 2011). In this study, bilinguals completed a verbal fluency task, on each trial producing as many members of semantic and phonemic categories (e.g., *animals*) as they could when given a minute for each category. Increased age was associated with a higher rate of intrusion errors, however, the total number of intrusion errors produced was very low even for older bilinguals who exhibited substantial difficulty with a non-linguistic control task. More recently, limited aging effects were again found in a verbal fluency task even though bilinguals had to switch languages on every trial, repeating the same category but in the other language (e.g., first they produced as many *animal* names as they could in English, and then they completed the same category in Spanish; Ivanova, Montoya, Murillo, & Gollan, in press). These studies again imply some limited aging effects, on ability to prevent unintended language switches. However, the verbal fluency task does not does not elicit production of full sentences, and elicits only a small number of intrusions, and therefore, leaves many open questions about how aging may or may not influence language control when switches are produced in connected speech, and with a greater number of observations of both intended and unintended language switches.

Method

Participants—Twenty-five older bilinguals participated. Fifteen were participants in the UCSD Alzheimer's Disease Research Center and were diagnosed as cognitively intact by two neurologists on the basis of medical, neurological, and neuropsychological exams in the same year as they participated in our study. The remaining participants were recruited from the local community (e.g., flyers, health fairs) and were classified as cognitively intact on the basis of reported daily activities, and scores on the Dementia Rating Scale (DRS; Mattis, 1988). Table 1 shows self-reported participant characteristics and ability to name pictures in each language on the Multilingual Naming Test or MINT (Gollan, et al., 2012) for 20 young and 20 older bilinguals (6 males in each group) who were matched for ability to name pictures in English and Spanish. Two of the older bilinguals tested had 6 and 7 years of education and could not be matched to any young bilingual for education level; these, and three other older bilinguals were excluded from analyses (to enable proficiency matching).

Materials and procedure—These were the same as in Experiment 1.

Results

Errors produced by older bilinguals were transcribed as in Experiment 1 and classified as: intrusions ($n = 345$; matched younger bilinguals: $n = 176$); partial intrusions ($n = 46$; matched group: $n = 52$); accent errors ($n = 227$; matched group: $n = 180$); within-language errors ($n = 864$; matched group: $n = 403$); omissions ($n = 79$; matched group: $n = 50$); and insertions ($n = 54$; matched group: $n = 37$). Those classified as more than one error type (accent: older bilinguals: $n = 1$; intrusion: older bilinguals: $n = 5$) were classified as in Experiment 1. Also as in Experiment 1, our analyses focused on intrusions and within-language errors. All of the older bilinguals produced at least 5 intrusions and up to as many as 48 (M=17; SD=11). All of the matched younger bilinguals produced at least 3 intrusions and up to as many as 22 ($M = 9$; SD=6).

As in Experiment 1, we excluded partial intrusions, as both older and younger bilinguals showed opposite effects of grammatical category for partial compared to full intrusions. Specifically, and unlike the pattern reported above for intrusions, at switch sites bilinguals produced partial intrusions significantly more often with content (older: M=1.2%, 95% CI (0.8%, 1.6%); matched younger: M=1.4%, 95% CI (0.8%, 2.0%)) than with function word targets (older: M=0.1%, 95% CI (0.02%, 0.2%); matched younger: M=0.2%, 95% CI (0.07%, 0.3%)). If part of speech effects on partial intrusion errors should be attributed to self-monitoring processes and attention, these might exhibit monitoring impairments in aging bilinguals. To test this hypothesis, we compared the rate at which intrusions at any position were repaired [(partial intrusions/(partial intrusions + full intrusions)] across age groups for both content and function words (similar results were found when restricting the analysis to switch positions). The results of this comparison are shown in Figure 6. Consistent with a monitoring impairment, older adults showed a significantly lower rate of repairs for content (older: M=31.2%, 95% CI (21.5%, 42.3%); matched younger: M=41.0%, 95% CI (27.3%, 56.7%)) as well as function words (older: M=1.4%, 95% (CI 0.3%, 2.6%); matched younger: M=5.4%, 95% CI (1.8%, 10.1%)). A logistic mixed effects regression on the rate of self-repairs (with random intercepts for participants and contrast coded effects of group and part of speech) revealed main effects of group (less repairs by older speakers; $\beta = -0.93$, $SE\,\beta = 0.36$, $\chi^2(1) = 6.34$, $p < .02$) and part of speech (less repairs for function words; $\beta = -2.81$, $SE\,\beta = 0.32$, $\chi^2(1) = 113.45$, $p < .0001$) but no significant interaction ($\beta = -0.48$, $SE\,\beta = 0.65$, $\chi^2(1) = 0.54$, $p < .47$).

Analysis methods—Reading times and error rate analyses were structured as in Experiment 1. We examined how aging[2] influenced performance (in paragraph reading times and errors) by comparing older bilinguals to their proficiency-matched younger bilinguals (see Table 1). We then considered whether any aging-related differences can be attributed to executive function differences by examining executive function effects within the subset of 20 older adults alone. For analyses of language dominance effects we again excluded Spanish-dominant bilinguals. Like the younger bilinguals in Experiment 1, most of the older bilinguals in Experiment 2 (17/20) were English dominant (one older bilingual had identical MINT scores in English and Spanish and was assumed to be English-dominant due to immersion in the English-dominant environment; Gollan et al., 2012). To maintain equal sample sizes, and proficiency matching, between young and older bilinguals, for these analyses we excluded 3 young bilinguals with the lowest English MINT scores. Note that overall, on average bilinguals in Experiment 2 had relatively more balanced proficiency in Spanish and English than bilinguals in Experiment 1 who tended to be more English dominant (see Table 1).

Whole-paragraph Reading times—The structure of regressions analyzing reading times followed Experiment 1, but were augmented by including aging group, as well as interactions between aging group, grammaticality, and rate. The results are shown in Figure 7. Collapsing across aging groups, the results were similar to Experiment 1; slower reading times for paragraphs with ungrammatical switches ($\beta = 4.49$, $SE\ \beta = 0.45$, $\chi^2(1) = 93.0$, $p < .0001$), slower reading times with higher rates of switching ($\beta = -2.95$, $SE\ \beta = 0.45$, $\chi^2(1) = 42.09$, $p < .0001$), and a significant interaction (reflecting stronger grammaticality effects at high switch rates; $\beta = -2.40$, $SE\ \beta = 0.89$, $\chi^2(1) = 7.19$, $p < .01$).

Comparing across age groups, older bilinguals read more slowly than young bilinguals ($\beta = 9.85$, $SE\ \beta = 2.51$, $\chi^2(1) = 13.05$, $p < .0005$), but aging did not magnify grammaticality and switch effects, i.e., there aging group did not interact with any of the other factors: neither grammaticality ($\beta = 0.88$, $SE\ \beta = 0.89$, $\chi^2(1) = 0.97$, $p < .35$), nor rate ($\beta = 0.03$, $SE\ \beta = 0.89$, $\chi^2(1) < .01$, $p < .98$); nor the grammaticality by rate interaction ($\beta = -0.83$, $SE\ \beta = 1.79$, $\chi^2(1) = 0.22$, $p < .65$).

To determine whether any aging-related effects were due to differences in executive function across the groups, we repeated the executive function analyses of Experiment 1 for the older adults. As shown in Table 3, unlike in Experiment 1, there were no significant interactions between grammaticality and any executive function measure; this suggests the main effect of aging group on reading times is not due to executive function impairments.

Intrusion Errors—As in Experiment 1, older bilinguals were much more likely to produce intrusion errors at switch sites (mean: 3.2%, 95% CI 2.2%, 4.3%) than other positions in

[2]Since our manipulation of age was dichotomous (young vs. old), the effect of age was treated as a categorical variable in all analyses. Note as well that the two participant groups show a clear separation on executive function measures (see Table 1). Given that older adults exhibited levels of impairment on executive function that were never observed in our younger adult sample, age and executive function are necessarily confounded in any between-group analysis. We therefore performed analyzed continuous measures of executive functions within (but not between) aging groups.

paragraphs (mean: 0.5%, 95% CI 0.4%, 0.7%). We therefore focused our analysis at these points.

Analyses followed the structure of Experiment 1, augmented with an effect of aging group. To verify that aging effects were not specific to one part of speech, we included an interaction of group by this factor. Parallel to the analysis of reading times, we also examined interactions between aging group, grammaticality, and rate. As shown in Figure 8, collapsing across aging group the results paralleled Experiment 1: greater intrusions in ungrammatical paragraphs ($\beta = 0.62$, $SE\ \beta = 0.17$, $\chi^2(1) = 13.3$, $p < .00052$); no overall rate effect ($\beta = -0.05$, $SE\ \beta = 0.14$, $\chi^2(1) = 0.1$, $p < .80$); and a significant interaction of rate and grammaticality ($\beta = -0.71$, $SE\ \beta = 0.29$, $\chi^2(1) = 5.72$, $p < .03$, reflecting stronger grammaticality effects at high than at low switch rates). In addition, bilinguals produced significantly more intrusion errors with function than with content words targets ($\beta = 0.38$, $SE\ \beta = 0.16$, $\chi^2(1) = 5.78$, $p < .02$). This effect was only marginally significant in Experiment 1, and may suggest that more balanced bilinguals (as were bilinguals in Experiment 2 relative to Experiment 1; see Spanish versus English MINT scores in Table 1) exhibit this effect regardless of word order (i.e., Spanish or English; see also Gollan et al., 2014).

Comparison across age groups showed that older adults produced significantly more intrusions than young bilinguals ($\beta = 0.86$, $SE\ \beta = 0.26$, $\chi^2(1) = 10.22$, $p < .005$). However, aging did not magnify any of the above-reported significant effects; there was no interaction of aging group with grammaticality ($\beta = 0.18$, $SE\ \beta = 0.30$, $\chi^2(1) = 0.34$, $p < .60$), rate ($\beta = -0.03$, $SE\ \beta = 0.29$, $\chi^2(1) = 0.34$, $p < .56$), part of speech ($\beta = -0.02$, $SE\ \beta = 0.28$, $\chi^2(1) = 0.29$, $p < .59$), or the grammaticality by rate interaction ($\beta = 0.90$, $SE\ \beta = 0.58$, $\chi^2(1) = 2.37$, $p < .13$).

To analyze age-related changes to language dominance effects, we analyzed data from our groups of 17 young and 17 older English-dominant proficiency matched bilinguals. Following Experiment 1, the regression models above were augmented to include language of production and its interaction with part of speech. These factors were allowed to interact with aging group. When collapsing across aging group, the results paralleled Experiment 1. There was a reversed dominance effect (replacing written English words with spoken Spanish translations more likely than vice versa; $\beta = 0.36$, $SE\ \beta = 0.15$, $\chi^2(1) = 5.53$, $p < .02$), which did not interact with part of speech ($\beta = 0.21$, $SE\ \beta = 0.30$, $\chi^2(1) = 0.44$, $p < .55$).

Inconsistent with the notion of an inhibitory control deficit in aging, or with the notion that reversed dominance effects reflect inhibitory control mechanisms, comparison across groups revealed no significant interaction of age group with language of production ($\beta = -0.35$, $SE\ \beta = 0.31$, $\chi^2(1) = 1.34$, $p < .25$), nor a significant effect on the interaction of language of production and part of speech ($\beta = 0.02$, $SE\ \beta = 0.61$, $\chi^2(1) < .01$, $p < .98$).

To examine whether more robust effects of executive functions might be found at lower levels of executive control ability, we repeated the executive function analyses of Experiment 1 for the 20 older bilinguals tested in Experiment 2 (excluding the 20 matched young bilinguals). As shown in Table 3, as in Experiment 1, there was a significant main effect of Trails performance on the rate of intrusions errors, but no significant effects of Stroop or Flanker performance. However, the interactions of executive functions with grammaticality were not replicated. As shown in Table 3, reading times exhibited no significant effects (recall that in Experiment 1, Stroop and Trails interacted with grammaticality in this analysis). Additionally, in Experiment 2 we found Stroop interference effects were correlated with the rate of within-language errors, but this effect was not significant in Experiment 1. It might seem that null effects in Experiment 2 could be caused by reduction in power (we had just 20 older bilinguals, and more than four times as many young bilinguals in Experiment 1). However, note that more robust, or at least equally robust, effects might have nevertheless been expected given the greater variation in executive control abilities present in the older bilingual group, and the hypothesis that the role of executive control might become apparent only at lower levels of function. In our discussion below, we will interpret only the one effect that appeared consistently in both Experiments 1 and 2 (the main effect of Trails on intrusion errors). Finally, executive function measures did not significantly interact with grammaticality[3].

Within-Language Errors—Shown in Figure 9, as in Experiment 1, within-languages errors were roughly equally likely to occur at switch sites (mean: 2.8%, 95% CI 1.5%, 4.3%) and other positions in paragraphs (mean: 2.4%, 95% CI 1.5%, 3.6%). We therefore analyzed errors at all points. Regression analyses were structured following those of above.

As shown in Figure 9, collapsing across aging groups, the results paralleled Experiment 1; significantly fewer within-language errors on function vs. content words ($\beta = -0.77$, $SE\ \beta = 0.06$, $\chi^2(1) = 155.87$, $p < .0001$); no effect of grammaticality ($\beta = -0.01$, $SE\ \beta = 0.06$, $\chi^2(1) < .01$, $p < .98$); a marginal tendency for more within-language errors at higher switch rates ($\beta = 0.12$, $SE\ \beta = 0.06$, $\chi^2(1) = 3.65$, $p < .06$); no interaction of rate and grammaticality ($\beta = 0.13$, $SE\ \beta = 0.12$, $\chi^2(1) = 1.09$, $p < .30$); and no interaction of grammaticality and part of speech ($\beta = -0.02$, $SE\ \beta = 0.11$, $\chi^2(1) = 0.04$, $p < .85$).

Comparing across age groups, older bilinguals produced significantly more within- language errors than younger bilinguals ($\beta = 0.62$, $SE\ \beta = 0.26$, $\chi^2(1) = 5.29$, $p < .05$). However, aging group did not interact with the part of speech effect ($\beta = 0.12$, $SE\ \beta = 0.12$, $\chi^2(1) = 1.01$, $p < .35$). (It also failed to interact with any other factor in the model: grammaticality ($\beta = 0.14$, $SE\ \beta = 0.12$, $\chi^2(1) = 1.33$, $p < .25$); rate ($\beta = -0.07$, $SE\ \beta = 0.12$, $\chi^2(1) = 0.37$, $p < .55$); and the grammaticality by rate interaction ($\beta = 0.04$, $SE\ \beta = 0.24$, $\chi^2(1) = 0.03$, $p < .90$)).

In the analysis of language dominance effects (with 17 young and 17 older proficiency matched English-dominant bilinguals), collapsing across age groups, the effect of language of production paralleled Experiment 1, bilinguals produced more within-language errors in the non-dominant than in the dominant language ($\beta = -0.95$, $SE\ \beta = 0.07$, $\chi^2(1) = 198.18$, $p < .0001$). In contrast to Experiment 1, this was not modulated by part of speech ($\beta = 0.08$, $SE\ \beta = 0.13$, $\chi^2(1) = 0.37$, $p < .55$).

[3]As shown in Table 3 there was a non-significant trend for individuals with worse Trails performance to show a reduced effect of grammaticality. Given the absence of interactions with aging group, and the absence of a significant effect in Experiment 1, we do not believe this trend reflects a reliable pattern.

Comparing across age groups, there was no significant interaction of group with language of production ($\beta = -0.111$, $SE\ \beta = 0.13$, $\chi^2(1) = 0.70$, $p < .45$) nor a significant effect on the interaction of language of production and part of speech ($\beta = 0.39$, $SE\ \beta = 0.26$, $\chi^2(1) = 2.21$, $p < .14$). The lack of interaction of part of speech and dominance in Experiment 2 likely reflects the fact that these participants were relatively more balanced in their relative proficiency in English and Spanish compared to the full set of young adults in Experiment 1. As discussed in Experiment 1, the strength of dominance effects (including the interaction with part of speech) is correlated with relative proficiency; more English-dominant bilinguals show stronger part of speech effects than more balanced bilinguals. Thus, for the more balanced bilinguals in Experiment 2, the difference between effects for content and function words should be smaller.

Discussion

Our main goals in Experiment 2 were to see if grammaticality effects on bilingual language control might be modulated by aging, and if relationships between language specific control and executive control might be more apparent in a group of bilinguals with relatively impaired executive control (i.e., older bilinguals). The results suggest that the answer to both of these questions is no. Though we found robust effects of grammaticality and switch rate that interacted, these effects were not significantly greater in older than in proficiency matched younger bilinguals. Additionally, we found, if anything, reduced evidence for integration between language control and executive control (i.e., young, but not older bilinguals, exhibited greater grammaticality effects on reading times for individuals with larger Stroop and Trails interference effects), and similar, but not greater evidence for involvement of executive control for preventing unintended language switches (in both young and older bilinguals, see Tables 2 and 3, there was a main effect of Trails on intrusion errors, but no other effects were significant).

On the other hand, aging bilinguals did not exhibit fully intact ability to switch languages in connected speech; they produced significantly more intrusion errors relative to younger bilinguals, and a lower rate of self-corrections of intrusion errors (i.e., rate of partial relative to full intrusion errors). Thus, we conclude that language specific control mechanisms remain relatively intact and unchanged in aging bilinguals, and that the main effect of aging (especially the increased rate of intrusion errors which can not simply be attributed to age-related slowing), should be attributed to decline in executive control ability, and a role this plays in monitoring upcoming speech after it is planned (Gauvin, De Baene, Brass, & Hartsuiker, 2016; Nozari, Dell, & Schwartz, 2011; Postma, 2000). On this view, relatively automatic aspects of language control remain intact in aging (e.g., grammatical encoding and the propensity to allow language switches in some, perhaps cross-linguistically more syntactically congruous contexts, and for some parts of speech, more than others). In contrast, more deliberative processes, triggered after formulation of sentence frames and lexical retrieval, are impaired in aging (e.g., monitoring, which may be more reliant on domain general executive control abilities). Consistent with this account, our analysis of partial intrusions revealed that older bilinguals had a lower proportion of self-repairs of the intrusion errors. Similar mechanisms might explain why older bilinguals also produced a greater rate of within-language errors than younger bilinguals, but further research will be needed to identify these relationships more precisely (as it was Stroop interference that seemed to predict within-language errors only in aging, not in young bilinguals, but Trails that seemed to predict production of intrusions in both young and older bilinguals).

General Discussion

In the current study, young and older Spanish-English bilinguals read aloud paragraphs with language switches that varied in the extent to which they conformed to constraints on naturally occurring code-switches (i.e., grammatical versus ungrammatical), and in the rate of language switches (high-switch had twice as many switches as low-switch paragraphs). Replicating prior work (Gollan et al., 2014; Kolers, 1966) bilinguals produced speech errors when reading aloud mixed-language paragraphs, including both within-language errors (such as saying *everything* instead of *everyone*) and language control failures (such as saying *juntos* instead of *together*). Of great interest, we observed robust effects of grammaticality, such that both young and older bilinguals produced switches in grammatical paragraphs relatively quickly and accurately but by contrast, needed more time, and were more likely to exhibit total failures of language control (i.e., intrusions), when they attempted to produce paragraphs with an abundance of ungrammatical switches (i.e., grammaticality effects were especially strong when switch rates were high). Importantly, intrusion errors revealed significant grammaticality effects for retrieval of both function and content word targets, thereby indicating that grammaticality effects are unlikely to be an artifact of the read-aloud task (i.e., intrusions did not arise exclusively because of skipping of short high-frequency targets during reading), and that grammatical encoding likely influences naturally occurring code switches in a manner similar to that observed herein.

Grammaticality and switch rate manipulations also appeared to affect bilingual language selection specifically, as they had absolutely no effect on production of within-language speech errors. The two error types also exhibited dissociations in part-of-speech and language dominance effects. Intrusion errors were produced equally often with function word as with content word targets, whereas within language errors were produced much less often with function than with content word targets. Language dominance effects also trended in opposite directions for the two types of speech errors. Specifically, bilinguals produced significantly more within-language errors in their non-dominant than in their dominant language (particularly for content rather than function words). By contrast, bilinguals were more likely to replace a dominant language target with its non-dominant language translation than vice-versa, i.e., intrusion errors exhibited *reversed language dominance effects* (and these dominance effects did not interact with part of speech). Finally, further suggesting a dissociation in the mechanisms underlying error types, intrusion errors were concentrated at switch sites, whereas within language errors were produced at switch and non-switch sites equally often.

The dissociations in effects observed across errors that require control over bilingual language selection (i.e., intrusions) versus those that do not (i.e., within-language errors) suggest that at least partially non-overlapping processing mechanisms support bilingual language control and speech production more generally. To further explore the cognitive mechanisms that support bilingual language selection we examined if grammaticality effects were modulated by individual differences in commonly administered tests of executive control abilities (in Experiment 1), and across younger versus older bilinguals (in Experiment 2). Older bilinguals read paragraphs more slowly, produced more intrusions and within-language errors than proficiency matched younger bilinguals, and were less likely to self-correct an intrusion in mid-utterance than younger bilinguals. In Experiment 1, bilinguals with stronger general shifting ability as measured by performance on Trails B relative to Trails A, produced fewer language intrusion errors than bilinguals with weaker general shifting ability. Trails performance was similarly related to production of intrusion errors in Experiment 2 with older bilinguals, and in both experiments Trails performance did not similarly influence the number of within-language errors produced. These results imply a role for nonlinguistic executive control ability, as measured by the Trails task, in ensuring that the intended language is produced during mixed-language speech. Importantly, this effect did not interact with grammaticality, and instead appeared to function independently from language specific control mechanisms.

Trails performance, and also Stroop interference effects, also appeared to modulate grammaticality effects on whole paragraph reading times in Experiment 1 such that bilinguals who exhibited greater interference on these tasks also read paragraphs more slowly, particularly more difficult to produce ungrammatical paragraphs. However, these results did not replicate in Experiment 2, even though older bilinguals exhibited significant deficits on executive control measures relative to young bilinguals (see Table 1). Finally, and even though older bilinguals produced paragraphs more slowly and with more errors than younger bilinguals, aging did not modulate the size of grammaticality, part-of-speech, and language dominance effects. Considering all the tests conducted across both experiments, together these provided some support for a limited role of executive control in preventing language intrusion errors, but no consistent evidence for integration between language specific control and domain general executive control mechanisms in controlling bilingual language selection.

Language-specific Control Mechanisms

These results support an emerging consensus in the field that that some language control mechanisms are likely to be relatively automatic and language-specific, that is, independent of domain-general executive control mechanisms. In particular, they demonstrate that syntactic structure has robust effects on the size of language switch costs, in general agreement with proposals of greater automaticity for specific types of switches (reflecting structural principles; Myers-Scotton, 1993; 2002; 2005; 2006), and with other recent studies that arrived at the same conclusion using different methods and via examination of a very limited set of structural alternations (Declerck & Philipp, 2015a; Dussias, 2003; Fairchild & Van Hell, in press; Rossi et al., submitted). Unique to our study was the finding that grammaticality modulated switch costs not only in production speed, but also in total failures of language control (intrusion errors), and in diverse speaker groups (including both young and older bilinguals). A question that arises is why do habitual switching patterns facilitate switching. Presumably, bilinguals tend to switch – at least in part – at points within sentences in which there is greater overlap in underlying syntactic structure between languages, or greater flexibility for other reasons (e.g., part of speech) with respect to whether or not translation equivalent terms can or cannot be readily exchanged. The finding that bilinguals have more difficulty switching languages at points that violate these constraints demonstrates either that experience with certain types of switches matter, or that – for independent reasons – some types of switches are naturally easier than others. We prefer the latter interpretation because this could motivate why certain types of switches are more frequent than others in the first place.

Our finding of contrasting part of speech effects on intrusions versus within-language errors further implied that syntactic structure has a unique influence on bilingual language control (relative to other aspects of speech production). This result might seem incongruent with failures to observe part of speech effects on switch times in sentence production in other production tasks (Declerk & Philipp, 2015). However, apparent differences across studies might reflect methodological differences; e.g., bilinguals in Declerk and Philipp produced a very small number of structures repeatedly, were forced to pause for over a second in between each word within the sentences, and part of speech effects were measured in response times rather than in errors. These differences may play a critical role in producing the different patterns. Other results suggest that multiple factors contribute to such effects: for example, in Gollan et al., 2014 part of speech effects were strongest on English targets in paragraphs with Spanish word order; here we tested only paragraphs with English word order and observed greater part of speech effects on intrusions with more balanced bilinguals in Experiment 2). Though the read-aloud task clearly differs from spontaneous speech, it arguably resembles fluent connected speech to a much greater extent than the bulk of studies in the extant literature which focused on production of single words (in the absence of context), or allowed greater preparation time for upcoming switches than would be present in more naturally paced speech (Declerk & Phillip, 2015).

Of great interest, part of speech effects in the current study interacted with language dominance in our analysis of within-language errors, but not in our analysis of intrusion errors. Furthermore, relative proficiency in English versus Spanish, significantly modulated the interaction between part of speech and dominance in Experiment 1. Language dominance effects were greater for content words (which tend to be lower frequency) than function words (which tend to be relatively higher frequency). This result is expected based on previous findings that frequency effects tend to be larger in the non-dominant than in the dominant languages, likely resulting from reduced frequency of use of the non-dominant relative to the dominant languages (e.g., Gollan, et al., 2008; Gollan, Slattery, Goldenberg, van Assche, Duyck, & Rayner, 2011; Duyck, Vanderelst, Desmet, & Hartsuiker, 2008). The absence of an analogous interaction for intrusion errors – which exhibited reversed dominance effects – implies a different processing locus for these dominance effects. This is consistent with our previous suggestion that reversed dominance effects reflect global application of inhibition to the dominant language (Gollan & Ferreira, 2009; see also Misra, Guo, Bobb, & Kroll, 2012; Guo, Liu, Chen, Y Li, 2013; Guo, Ma, & Liu, 2013; Van Assche, Duyck, & Gollan, 2013)—yielding consistent effects across parts of speech (and, therefore, no interaction).

Aging also did not modulate part of speech effects or language dominance effects. Because function words tend to be short and arguably communicate less meaning than content words (for a recent discussion, see Bell, Brenier, Gregory, Girand, & Jurafsky, 2009), it might seem that their frequent elicitation of intrusion errors in bilingual speech reflects temporary lapses in attention or monitoring (Poulisse & Bongaerts, 1994), or even temporary forgetfulness as to which language was the target language. However, such explanations would predict greater part of speech effects in older than in younger bilinguals, and we observed no hint of such an effect (see results section of Experiment 2). Similar, and perhaps even broader, implications can be drawn from the absence of any dramatic modulation of language-dominance effects by aging. Assuming that dominance reversal reflects operation of a general mechanism of inhibitory control (Green, 1998), and if inhibitory control is impaired in aging (Hasher & Zacks, 1988; Zacks & Hasher, 1994), older bilinguals should have been less able to apply inhibition, and should have exhibited significantly smaller reversal of language dominance effects. Though only young bilinguals exhibited a significant reversal of dominance effects on intrusion errors (in Experiment 1), older bilinguals exhibited trends in this direction, and there was no hint of an interaction between age-group and dominance effects on intrusion errors.

The absence of an interaction between aging and grammaticality effects in our data also support the notion of language specific control mechanisms that operate independently of executive control mechanisms, and therefore remain relatively preserved in aging (see similar arguments in Burke & Osbourne, 1997). Tests of executive control ability in the same young and older bilinguals revealed robust aging effects that were particularly large in more difficult processing conditions (e.g., incongruent trials of the Flanker Task; see Table 1). Thus, there is an apparent dissociation in aging effects on language control versus executive control. In the linguistic domain, aging generally increases processing time and errors regardless of the relative difficulty of the language-control problem. In contrast, in other executive control tasks, aging effects were most apparent in more difficult control conditions (i.e., robust interactions were observed). This implies either that executive control mechanisms are sufficiently intact in aging to support language processing (e.g., Kramer, Humphrey, Larish, Logan, & Strayer, 1994; Mayr, 2001), or that the language system is equipped with its own specialized control mechanisms that remain sheltered from aging related declines in domain-general control. This apparent dissociation in aging effects across domains is consistent with several recent studies. Direct comparison of aging effects on cued task and language switching has revealed different patterns of aging effects in each domain (Calabria et al., 2015; Weissberger et al., 2012). In the literature comparing task and language switching there is growing consensus that there is relatively little overlap in the underlying processing mechanisms that support switches across domains (Branzi, Calabria, Boscarino, & Costa, 2016; Calabria, et al., 2012; Gollan, Kleinman, & Wierenga, 2014; Prior & Gollan, 2013).

Though we did not observe fully cost free switches in the present study, it seems notable that doubling the number of switches within grammatical paragraphs did not lead bilinguals to produce more intrusion errors (contrasting high versus low switch rate in grammatical paragraphs). These grammatical switches elicited costs only in full paragraph reading times, and this might imply independence between mechanisms that elicit language switches (and costs in time) and those that function to prevent language control failures (and costs in errors). However, such an interpretation could be premature given the possibility of speed accuracy trade-offs (in Experiment 1 bilinguals actually produced fewer intrusions in high-switch than in low-switch grammatical paragraphs). Moreover, it is possible that fully cost-free switches could be observed in paragraphs with even lower switch rates, and with more fine-grained manipulations of grammaticality than implemented here. While acknowledging this possibility, pursuit of cost-free switches might be less productive than attempting to understand factors that appear to significantly modulate switch costs.

An additional implication is more methodological in nature. The original read-aloud study (Kolers, 1966) was criticized for requiring bilinguals to produce haphazard language switches (Sridhar & Sridhar, 1980). However, in the current study, bilinguals produced intrusions even when paragraphs contained mostly grammatical and relatively low switch rate, consistent with our proposal that the read-aloud task engages the language production system in ways that reflect the workings of the same control mechanisms that operate when bilinguals produce code-mixed speech spontaneously. Moreover, though function words are skipped significantly more often than content words when bilinguals read-aloud mixed-language passages, our observation of significant grammaticality effects on both content and function word targets provides further support for our proposal that intrusions produced in the read-aloud task reflect mechanisms of the language production system (much more than mechanisms of reading comprehension – though these are necessarily implicated at least partly in the read aloud task). Thus, the read aloud task appears to engage the production system in ways that reveal the mechanisms underlying natural speech production. Further investigation of both reading times and error rates in this task will provide a unique form of evidence about the mechanisms of bilingual language control (Gollan et al., 2014).

The Role of Domain-general Executive Control in Bilingual Language Selection

A primary motivation for the present study, was our hypothesis that if language control mechanisms are broadly reliant on domain general executive control, we should observe significant interactions between our grammaticality manipulation and individual difference measures of executive function, or aging, or both (and similar predictions could be made for part of speech and language dominance effects). In contrast, the absence of such interactions would imply the existence of some language specific control mechanisms that function relatively independently of domain-general control. In both Experiments 1 and 2 we observed little to no evidence for integration of control mechanisms across linguistic and nonlinguistic domains.

Perhaps most notably, older bilinguals showed no enhanced sensitivity to experimental manipulations targeting language control (i.e., switching rate and grammaticality), or to other signatures of language control (part of speech and language-dominance), even though they read paragraphs more slowly, and produced more speech errors, including both within-language errors and intrusion errors than young bilinguals. Importantly, young and older bilinguals were carefully matched for ability to name pictures in English and in Spanish. Thus, age differences in the read-aloud task could not be attributed to between group differences in bilingual proficiency level. This raises a question: if putative language-specific control mechanisms are preserved in older age, why did older bilinguals have more difficulty with the task than younger bilinguals?

Relevant to this discussion, both young and older bilinguals who had greater relative difficulty completing Trails B than Trails A, exhibited a higher overall rate of intrusion errors but there was little evidence of interactions between Trails and other signatures of language control (i.e., we found an interaction with grammaticality only in Experiment 1 in reading times, not in intrusion errors). Furthermore, the Trails measure showed no analogous relationship to within-language errors. This implies recruitment of domain general executive control specifically for the purpose of maintaining control over bilingual language selection, and not for prevention of speech errors in general (note that Stroop interference effects did appear to modulate language dominance effects on within-language errors but this is perhaps not surprising given that ability to control Stroop interference increases with language proficiency, and this task is relatively more linguistic than Trails and Flanker). To account for this pattern of results, we suggest that language specific and domain-general control mechanisms operate relatively independently to jointly maintain language control (Gollan et al., 2011). This hypothesis raises several questions, including (a) why Trails, but not the other measures, exhibited this evidence of recruitment, (b) why age-effects appeared to be more broad (applying to both within- and between-language errors), and most interestingly, (c) what different roles in speech production might be subserved by language-specific versus domain-general control mechanisms.

It might be tempting to interpret the Trails effect found in the younger bilinguals (Experiment 1) as suggesting that domain general switching ability (Yehene & Meiran, 2007), possibly measured by Trails (but see Tao et al., 2015) relatively more than by the Flanker or Stroop tasks, is more critical to code-switching than other aspects of executive control (e.g., ability to sustain attention in the face of distraction). Such a result would be consistent with previous suggestions of a bilingual advantage in switching (Prior & MacWhinney, 2010; Prior & Gollan, 2011; Tao et al., 2015; and an association between habitual switching and more efficient executive control, Hartanto & Yang, in press; Verreyt et al., 2016). However, we would caution against such an interpretation until a clearer framework emerges as to precisely which aspects of bilingual language use might lead to a switching advantage (see, e.g., some recent failures to replicate the switching advantage; Hernández, Martin, Barceló, & Costa, 2014; Paap & Greenberg, 2013; Paap & Sawi, 2014; Wisehart, Viswanathan, & Bialystok, 2016). Similarly, it might be argued that perhaps only certain types of language switches, a type not measured here, are elicited by domain general executive control, or that a certain type of bilingual, a type not tested here, relies to a greater extent on domain general executive control to achieve and maintain language selection. Even if possibly true, the present study provided ample opportunities to reveal links between language selection and executive control (having manipulating switch types, switch rates, testing multiple age groups, and multiple measures of executive control). Thus, at minimum our study suggests substantial limitations on the degree of cross-talk across domains.

As to why aging effects appear primarily in the form of general main effects (slower reading time, and increased error rates, without interacting with grammaticality), it is possible that aging affects numerous different cognitive processes, including some that affect bilingual language selection, and perhaps others that lead to greater production of speech errors in general (MacKay & James, 2004; but see Vousden & Maylor, 2006). One notable exception in this regard was older bilinguals' lower rate of self-corrections of intrusion errors (i.e., rate of partial relative to full intrusion errors; see Figure 6). This result led us to suggest that executive control ability may play an important role in monitoring upcoming speech after it is planned (see Discussion section in Experiment 2; Gauvin et al., 2016; Nozari et al., 2011). Also of particular interest here was the absence of an interaction between aging and language dominance on within-language errors; both aging effects and dominance effects were highly robust in Experiment 2, but we obtained no evidence that older bilinguals had particular difficulty controlling speech in a less proficient language. Though language control is clearly important for production of the non-dominant language, other aging related benefits could offset these expected challenges with producing the nondominant language in older age (e.g., Gollan, et al., 2008) reported an advantage in production of low-frequency words in the non-dominant language for older relative to younger bilinguals).

Finally, the question of whether the language system relies on domain-specific control mechanisms also has important implications for studies of cognitive training and the hypothesis of cognitive reserve. On this view, certain life-long experiences confer some degree of better brain health, which then transfers into better maintenance of cognitive status into older age, and delay of onset of symptoms of progressive diseases leading to dementia, such as Alzheimer’s disease (Stern, 2002, 2009). One such life-long experience could include bilingualism. This has been supported by reports of later age of onset of Alzheimer’s in bilinguals, as well as better preservation of executive functioning in older age in bilinguals relative to monolinguals (Bialystok et al., 2009; but see Hilchey et al., 2014; Paap et al., 2015). Although our results suggest that the language system is equipped with its own control mechanisms that remain sheltered from such declines, we also obtained evidence for a unique role for domain-general control in maintaining bilingual language control. A challenge for future work will be to identify the extent to which monitoring processes are trainable, and what other aspects of bilingualism could potentially provide a source of more general cognitive benefits. Several investigators have cautioned that transfer of training effects across tasks tend to be quite restricted (Consensus Statement, 2008; Hulme & Melby-Lervag, 2012; Melby-Lervag & Hulme, 2013; Owen et al., 2010; Redick et al., 2012), thus, it would be important to seek empirical evidence for transfer in monitoring abilities to test the feasibility of this possible explanation. Until the role of domain-general monitoring (or control more broadly) in bilingual language processing can be made more explicit it will remain unclear precisely which aspect of bilingual language use provides “brain-training” of the sort that might provide general cognitive benefits. Even though life-long bilingualism unquestionably involves substantial practice in a form of cognitive control, there appears to substantial automaticity and domain-specificity of functioning of these mechanisms, and further work is needed to identify more precisely which forms of control are needed to function in different aspects of speech processing.

Acknowledgments

The authors thank for comments on earlier versions of this manuscript, Rosa Montoya and Mayra Murillo for composition of the paragraphs and error coding, and Clayton Bryson for assistance with data coding.

Funding

This research was supported by grants from the National Institute on Deafness and Other Communication Disorders (011492), the National Institute of Child Health and Human Development (050287, 051030, 079426) and the

National Science Foundation (BCS1344269, BCS1457159). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NIH or NSF.

References

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

Aron AR. The neural basis of inhibition in cognitive control. Neuroscientist. 2007; 13:214–228. [PubMed: 17519365]

Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language. 2013; 68:255–278.

Bell A, Brenier JM, Gregory M, Girand C, Jurafsky D. Predictability effects on durations of content and function words in conversational English. Journal of Memory and Language. 2009; 60:92–111.

Bialystok E, Craik FIM, Green DW, Gollan TH. Bilingual minds. Psychological Science in the Public Interest. 2009; 10:89–129. [PubMed: 26168404]

Bobb SC, Wodniecka Z. Language switching in picture naming: What asymmetric switch costs (do not) tell us about inhibition in bilingual speech planning. Journal of Cognitive Psycholology. 2013; 25:568–585.

Bullock, BE.; Toribio, AJ. Themes in the study of code-switching. In: Bullock, BE.; Toribio, AJ., editors. The Cambridge Handbook of Linguistic Code-switching. Cambridge, UK: Cambridge University Press; 2009. p. 1-17.

Burke DM. Language, aging, and inhibition deficits: Evaluation of a theory. Journal of Gerontology. 1997; 52B(6):254–264.

Burke, DM. Language production and aging. In: Kemper, S.; Kliegl, R., editors. Constraints on language: Aging, grammar and memory. Boston: Kluwer; 1999. p. 3-28.

Burke, DM.; Osborne, GL. Aging and inhibition deficits: Where are the effects?. In: Gorfein, D.; MacLeod, C., editors. On the place of inhibitory processes in cognition. Washington, DC: American Psychological Association Press; 2007.

Burke DM, MacKay DG. Memory, language and ageing. Philosophical Transactions of the Royal Society: Biological Sciences. 1997; 352:1845–1856. [PubMed: 9460069]

Branzi FM, Calabria M, Boscarino ML, Costa A. On the overlap between bilingual language control and domain general executive control. Acta Psychologica. 2016; 166:21–30. [PubMed: 27043252]

Calabria M, Branzi FM, Marne P, Hernández M, Costa A. Age-related effects over bilingual language control and executive control. Bilingualism: Language and Cognition. 2015; 18:65–78.

Calabria M, Hernández M, Branzi FM, Costa A. Qualitative differences between bilingual language control and executive control: Evidence from task switching. Frontiers in Psychology. 2012; 2:399. [PubMed: 22275905]

Christoffels IK, Firk C, Schiller NO. Bilingual language control: An event-related brain potential study. Brain Research. 2007; 1147:192–208. [PubMed: 17391649]

Cohen JD, MacWhinney B, Flatt M, Provost J. PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. Behavior Research Methods, Instruments, & Computers. 1993; 25:257–271.

A Consensus on the Brain Training Industry from the Scientific Community. 2014. http://longevity3.stanford.edu/blog/2014/10/15/the-consensus-on-the-brain-training-industry-from-the-scientific-community-2/

Costa A, Santesteban M. Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. Journal of Memory and Language. 2004; 50:491–511.

Declerck M, Koch I, Philipp AM. The minimum requirements of language control: Evidence from sequential predictability effects in language switching. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2015; 41:377–394. DOI: 10.1037/xlm0000021

Declerck M, Philipp AM. A sentence to remember: Instructed language switching in sentence production. Cognition. 2015a; 137:166–173. DOI: 10.1016/j.cognition.2015.01.006 [PubMed: 25659539]

Declerck M, Philipp AM. A review of control processes and their locus in language switching. Psychonomics Bulletin & Review. 2015b; 22:1630–1645.

Declerck M, Philipp AM, Koch I. Bilingual control: Sequential memory in language switching. Journal of Experimental Psychology: Learning, Memory & Cognition. 2013; 39:1793–1806. DOI: 10.1037/a0033094

Dell GS. A spreading-activation theory of retrieval in sentence production. Psychological Review. 1986; 93:283–321. [PubMed: 3749399]

Dell GS, Oppenheim GM, Kittredge AK. Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. Language and Cognitive Processes. 2008; 23:583–608. [PubMed: 20622975]

Dussias PE. Spanish-English code-mixing at the auxiliary phrase: Evidence from eye-movements. Revista Internacional de Lingüística Iberoamerican. 2003; 2:7–34.

Duyck W, Vanderelst D, Desmet T, Hartsuiker RJ. The frequency-effect in second-language visual word recognition. Psychonomic Bulletin & Review. 2008; 15:850–855. [PubMed: 18792515]

Emmorey K, Borinstein HB, Thompson R, Gollan TH. Bimodal bilingualism. Bilingualism: Language and Cognition. 2008; 11:43–61.

Fairchild, S.; van Hell, JG. Determiner-noun code-switching in Spanish-English heritage speakers. Bilingualism: Language and Cognition. in presshttp://dx.doi.org/10.1017/S1366728915000619

Fan J, McCandliss BD, Sommer T, Raz A, Posner MI. Testing the efficiency and independence of attentional networks. Journal of Cognitive Neuroscience. 2002; 14:340–347. [PubMed: 11970796]

Festman T, Münte TF. Cognitive control in Russian-German bilinguals. Frontiers in Psychology. 2012; 3:115. [PubMed: 22529831]

Fink A, Goldrick M. Pervasive benefits of preparation in language switching. Psychonomic Bulletin and Review. 2015; 22:808–814. [PubMed: 25257712]

Finkbeiner M, Almeida J, Janssen N, Caramazza A. Lexical selection in bilingual speech production does not involve language suppression. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2006; 32:1075–1089.

Fricke M, Kroll JF, Dussias PE. Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. Journal of Memory and Language. in press.

Garrett, MF. The analysis of sentence production. In: Bower, GH., editor. The psychology of learning and motivation. Vol. 9. New York: Academic Press; 1975. p. 133-177.

Gauvin HS, De Baene W, Brass M, Hartsuiker RJ. Conflict monitoring in speech processing: An fMRI study of error detection in speech production and perception. NeuroImage. 2016; 126:96–105. [PubMed: 26608243]

Goldrick M. Linking speech errors and generative phonological theory. Language and Linguistics Compass. 2011; 5:397–412.

Gollan TH, Ferreira VS. Should I Stay or Should I Switch? A Cost–Benefit Analysis of Voluntary Language Switching in Young and Aging Bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:640–665.

Gollan TH, Kleinman D, Wierenga CE. What's easier: Doing what you want, or being told what to do? Cued versus voluntary language and task switching. Journal of Experimental Psychology: General. 2014; 143:2167–2195. [PubMed: 25313951]

Gollan TH, Montoya RI, Cera CM, Sandoval TC. More use almost always means smaller a frequency effect: Aging, bilingualism, and the weaker links hypothesis. Journal of Memory and Language. 2008; 58:787–814. [PubMed: 19343088]

Gollan TH, Sandoval T, Salmon DP. Cross-language intrusion errors in aging bilinguals reveal the link between executive control and language selection. Psychological Science. 2011; 22:1155–1164. [PubMed: 21775653]

Gollan TH, Schotter ER, Gomez J, Murillo M, Rayner K. Multiple Levels of Bilingual Language Control: Evidence From Language Intrusions in Reading Aloud. Psychological Science. 2014; 25:585–595. [PubMed: 24367061]

Gollan TH, Slattery TJ, Goldenberg D, van Assche E, Duyck W, Rayner K. Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. Journal of Experimental Psychology: General. 2011; 140:186–209. [PubMed: 21219080]

Gollan TH, Weissberger G, Runnqvist E, Montoya RI, Cera CM. Self-ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. Bilingualism: Language and Cognition. 2012; 15:594–615.

Green DW. Mental control of the bilingual lexicosemantic system. Bilingualism: Language and Cognition. 1998; 1:67–81.

Gullifer J, Kroll JF, Dussias PE. When language switching has no apparent cost: Lexical access in sentence context. Frontiers in Psychology. 2013; 4:1–13. [PubMed: 23382719]

Guo T, Liu F, Chen B, Li S. Inhibition of non-target languages in multilingual word production: Evidence from Uighur-Chinese-English trilinguals. Acta Psychologica. 2013; 143:277–283. [PubMed: 23688401]

Guo T, Ma F, Liu F. An ERP study of inhibition of non-target languages in trilingual word production. Brain and Language. 2013; 127:12–20. [PubMed: 23994766]

Hartanto A, Yang H. Disparate Bilingual Experiences Modulate Task-switching Advantages: A Diffusion-model Analysis of the Effects of Interactional Context on Switch Costs. Cognition. in press.

Hasher L, Zacks RT. Working memory, comprehension, and aging: A review and a new view. The Psychology of Learning and Motivation. 1988; 22:193–225.

Hernandez AE. Language switching in the bilingual brain: What's next? Brain and Language. 2009; 109:133–140. [PubMed: 19250662]

Hernandez AE, Kohnert KJ. Aging and language switching in bilinguals. Aging, Neuropsychology, and Cognition. 1999; 6:69–83.

Hilchey, MD.; Saint-Aubin, J.; Klein, RM. Cambridge Handbook of Bilingual Processing. Cambridge: CUP; 2015. Does bilingual exercise enhance cognitive fitness in non-linguistic executive processing tasks.

Hulme C, Melby-Lervag M. Current evidence does not support the claims made for CogMed working memory training. Journal of Applied Research in Memory and Cognition. 2012; 1:197–200.

Ibáñez A, Macizo P, Bajo M. Language access and language selection in professional translators. Acta Psychologica. 2010; 135:257–266. DOI: 10.1016/j.actpsy.2010.07.009 [PubMed: 20705277]

Ivanova I, Murillo M, Montoya RI, Gollan TH. Does bilingual language control decline in older age? Linguistic Approaches to Bilingualism. in press.

Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language. 2008; 59:434–446. [PubMed: 19884961]

King G, Zeng L. Logistic regression in rare events data. Political Analysis. 2001; 9:137–163.

Kolers P. Reading and talking bilingually. American Journal of Psychology. 1966; 3:357–376. [PubMed: 5968473]

Kleinman D, Gollan TH. Speaking two languages for the price of one: Bypassing language control mechanisms via accessibility-driven switches. Psychological Science. in press.

Kramer AF, Humphrey DG, Larish JF, Logan GD, Strayer DL. Aging and inhibition: Beyond a unitary view of inhibitory processing in attention. Psychology and Aging. 1994; 9:491–512. [PubMed: 7893421]

Kroll JF, Bobb SC, Misra M, Guo T. Language selection in bilingual speech: Evidence for inhibitory processes. Acta Psychologica. 2008; 128:416–430. [PubMed: 18358449]

MacKay DG, James LE. Sequencing, speech production, and selective effects of aging on phonological and morphological speech errors. Psychology and Aging. 2004; 19:93–107. [PubMed: 15065934]

Mattis, S. Dementia rating scale: Professional manual. Odessa, FL: Psychological Assessment Resources; 1988.

Mayr U. Age differences in the selection of mental sets: The role of inhibition, stimulus ambiguity, and response-set overlap. Psychology and Aging. 2001; 16:96–109. [PubMed: 11302371]

Meinzer M, Seeds L, Flaisch T, Harnish S, Cohen ML, McGregor K, et al. Impact of changed positive and negative task-related brain activity on word-retrieval in aging. Neurobiology of Aging. 2012; 33:656–669. [PubMed: 20696496]

Melby-Lervag M, Hulme C. Is working memory training effective? A meta-analytic review. Developmental Psychology. 2013; 49:270–291. [PubMed: 22612437]

Meuter RF, Allport A. Bilingual language switching in naming: Asymmetrical costs of language selection. Journal of Memory and Language. 1999; 40:25–40.

Mickes L, Wixted JT, Fenema-Notestine C, Galasko D, Bondi MW, Thal LJ, Salmon DP. Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. Neuropsychology. 2007; 21:696–705. [PubMed: 17983283]

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

Muysken, P. Bilingual speech: A typology of code-switching. Oxford: Cambridge University Press; 2000.

Myers-Scotton, C. Duelling languages: Grammatical structure in codeswitching. Oxford, England: Oxford University Press; 1993,1997. Original work published 1993

Myers-Scotton, C. Contact linguistics: Bilingual encounters and grammatical outcomes. Oxford: Oxford University Press; 2002.

Myers-Scotton, CM. Supporting a differential access hypothesis: Code switching and other contact data. In: Kroll, J.; de Groot, A., editors. Handbook of bilingualism: psycholinguistic approaches. New York: Oxford University Press; 2005. p. 326-348.

Myers-Scotton CM. Natural codeswitching knocks on the laboratory door. Bilingualism: Language and Cognition. 2006; 9:203–212.

Myers-Scotton C, Jake JL. Matching lemmas in a bilingual competence and production model: Evidence from intrasentential code-switching. Linguistics. 1995; 33:981–1024.

Nozari N, Dell GS, Schwartz MF. Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. Cognitive Psychology. 2011; 63:1–33. [PubMed: 21652015]

Owen A, et al. Putting brain training to the test. Nature. 2010; 465:775–8. [PubMed: 20407435]

Paap KR, Greenberg ZI. There is no coherent evidence for a bilingual advantage in executive processing. Cognitive Psychology. 2013; 66:232–258. [PubMed: 23370226]

Paap KR, Johnson HA, Sawi O. Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. Cortex. 2015; 65:265–278. [PubMed: 26048659]

Paap KR, Sawi O. Bilingual advantages in executive functioning: problems in convergent validity, discriminant validity, and the identification of the theoretical constructs. Frontiers in Psychology. 2014; 5:962. [PubMed: 25249988]

Peeters D, Runnqvist E, Bertrand D, Grainger J. Asymmetrical switch costs in bilingual language production induced by reading words. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2014; 40:284–292.

Philipp AM, Gade M, Koch I. Inhibitory processes in language switching: Evidence from switching language-defined response sets. European Journal of Cognitive Psychology. 2007; 19:395–416.

Philipp AM, Koch I. Inhibition in language switching: What is inhibited when switching between languages in naming tasks? Journal of Experimental Psychology: Learning, Memory, and Cognition. 2009; 35:1187–1195.

Poplack S. Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of codeswitching. Linguistics. 1980; 18:581–618.

Postma A. Detection of errors during speech production: a review of speech monitoring models. Cognition. 2000; 77:97–31. [PubMed: 10986364]

Poulisse, N. Slips of the tongue: Speech errors in first and second language production. Amsterdam/ Philadelphia: John Benjamins; 1999.

Poulisse N, Bongaerts T. First language use in second language production. Applied Linguistics. 1994; 15:36–57.

Prior A, Gollan TH. Good Language-Switchers are Good Task-Switchers: Evidence from Spanish–English and Mandarin–English Bilinguals. Journal of the International Neuropsychological Society. 2011; 17:682–691. [PubMed: 22882810]

Prior A, Gollan TH. The elusive link between language control and executive control: A case of limited transfer. The Journal of Cognitive Psychology. 2013; 25:622–645. [PubMed: 24688756]

Prior A, MacWhinney B. A bilingual advantage in task switching. Bilingualism: Language and Cognition. 2010; 13:253–262.

Raz, N. Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In: Craik, FIM.; Salthouse, TA., editors. The Handbook of Aging and Cognition. 2. Lawrence Erlbaum Associates Publishers; Mahwah, NJ, USA: 2000. p. 1-90.

Redick TS, Shipstead Z, Harrison TL, Hicks KL, Fried DE, Hambrick DZ, Kane MJ, Engle RW. No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. Journal of Experimental Psychology: General. 2012; 142:1–21. [PubMed: 22612770]

Reitan RM. Validity of the Trail Making Test to organic brain damage. Perceptual Motor Skills. 1958; 8:271–276.

Rossi E, Ting C, Diaz M, Newman S, van Hell GJ, Dussias PE. Functional neuroimaging and codeswitching: Neural basis of switching words in meaningful sentences. in preparation.

Salthouse TA. The processing speed theory of cognitive aging. Psychological Review. 1996; 103:403–428. [PubMed: 8759042]

Salthouse TA. Is flanker-based inhibition related to age? Identifying specific influences of individual differences on neurocognitive variables. Brain & Cognition. 2010; 73:51–61. [PubMed: 20303636]

Soveri A, Rodriguez-Fornells A, Laine M. Is there a relationship between language switching and executive functions in bilingualism? introducing a within group analysis approach. Frontiers Psychology. 2011; 2:183.

Spieler DH, Balota DA, Faust ME. Stroop performance in healthy younger and older adults and in individuals with Dementia of the Alzheimer's type. Journal of Experimental Psychology: Human Perception & Performance. 1996; 22:461–479. [PubMed: 8934854]

Sridhar SN, Sridhar KK. The syntax and psycholinguistics of bilingual code mixing. Canadian Journal of Psychology. 1980; 34:407–416.

Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. Journal of the International Neuropsychological Society. 2002; 8:448–460. [PubMed: 11939702]

Stern Y. Cognitive reserve. Neuropsychologia. 2009; 47:2015–2028. [PubMed: 19467352]

Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology. 1935; 18:643–662.

Tao L, Taft M, Gollan TH. The bilingual switching advantage: Sometimes related to bilingual proficiency, sometimes not. The Journal of the International Neuropsychological Society. 2015; 21:531–544. [PubMed: 26527242]

Tarlowski A, Wodniecka Z, Marzecová A. Language switching in the production of phrases. Journal of Psycholinguistic, Research. 2013; 42:103–180. [PubMed: 22450881]

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. Role of left prefrontal cortex in retrieval of semantic knowledge: A re-evaluation. Proceedings of the National Academy of Science. 1997; 94:14792–14797.

Valian V. Bilingualism and cognition. Bilingualism: Language and Cognition. 2015; 18(1):3–24.

Verhagen P. Aging and executive control: Reports of a demise greatly exaggerated. Psychological Science. 2011; 20:174–180.

Verhoef K, Roelofs A, Chwilla DJ. Role of inhibition in language switching: evidence from event-related brain potentials in overt picture naming. Cognition. 2009; 110:84–99. [PubMed: 19084830]

Verreyt N, Woumans E, Vandelanotte D, Szmalec A, Duyck W. The influence of language-switching experience on the bilingual executive control advantage. Bilingualism: Language and Cognition. 2016; 19:181–190.

Vousden JI, Maylor EA. Speech errors across the lifespan. Language and Cognitive Processes. 2006; 21:48–77.

Yehene E, Meiran N. Is there a general task switching ability? Acta Psychologica. 2007; 126:169–195. [PubMed: 17223059]

Weissberger GH, Wierenga CE, Bondi MW, Gollan TH. Partially overlapping mechanisms of language and task control in young and older bilinguals. Psychology and Aging. 2012; 27:959–974. [PubMed: 22582883]

West RL. An Application of Prefrontal Cortex Function Theory to Cognitive Aging. Psychological Bulletin. 1996; 120:272–292. [PubMed: 8831298]

West R, Baylis GC. Effect of increased response dominance and contextual disintegration on the Stroop interference effect in older adults. Psychology and Aging. 1998; 13:206–217. [PubMed: 9640582]

Wisehart M, Viswanathan M, Bialystok E. Flexibility in task switching by monolinguals and bilinguals. Bilingualism: Language and Cognition. 2016; 19:141–146.

Zacks, RT.; Hasher, L. Directed ignoring: Inhibitory regulation of working memory. In: Dagenbach, D.; Carr, TH., editors. Inhibitory Processes in Attention, Memory, and Language. San Diego, CA: Academic Press; 1994. p. 241-264.

Appendix

Grammatical Low-switch

He then lit it by striking un cerillo debajo del asiento de su chair. The truly meticulous manera en que hacía papá his cigarettes was indeed an art. He took his first puff, detuvo la respiración, and then exhaled smoke through his nose with a healthy satisfaction. Blowing smoke through his nose siempre me fascinaba. For me it was nothing short of a miracle. Me pregunté, how did he do it? Someday I would find out. Someday yo aprendería, porque todos los hombres learn how, and I would get to be a man como mi padre.

Grammatical High-switch

Luego lo prendió by striking a match debajo del asiento of his chair. The truly meticulous manner en que hacía papá his cigarettes was indeed un arte. He took his first puff, detuvo his breath, and then exhaled smoke through las narices with a healthy satisfaction. Blowing humo through his nose siempre fascinated me. Para mi it was nothing short of a miracle. I asked myself, cómo lo hacía? Someday I would find out. Algún día I would learn how, because all los hombres learn how, and I would get to be un hombre just like my padre.

Ungrammatical Low-Switch

He then lit it by striking un cerillo debajo del asiento of his chair. The truly meticulous manner in which Dad rolled his cigarettes was an art. He took his first chupazo, detuvo su respiración, y luego exhaled smoke through his narices con una healthy satisfaction. Blowing humo por la nose always fascinated me. For me it was nothing short of a miracle. I asked myself, how lo hacía? Algún día I would find out. Someday I would learn how, because all hombres aprenden, y I would get to be a hombre como mi papá.

Ungrammatical High-switch

Luego lo lit by striking a match debajo del seat of his chair. The verdadera meticulous manner in which Dad rolled his cigarrillos era un art. He dio el primer puff, held his breath, and luego echó humo through his nose with a healthy satisfaction. Blowing humo through his nose always me fascinaba. For me it was nothing short de un miracle. I asked myself, ¿cómo did he do it? Someday yo would find out. Someday yo aprendería how, because all hombres learn how, and I would get to be a hombre como mi papá.

Highlights

- Grammatical encoding facilitates language switching for content and function words.
- Executive control facilitates switching independent from language-specific control.
- Grammatical encoding mechanisms are language-specific and remain intact in aging.
- Aging impairs language selection at a post-lexical monitoring stage.
- Bilingual language control is largely, but not entirely, domain specific.

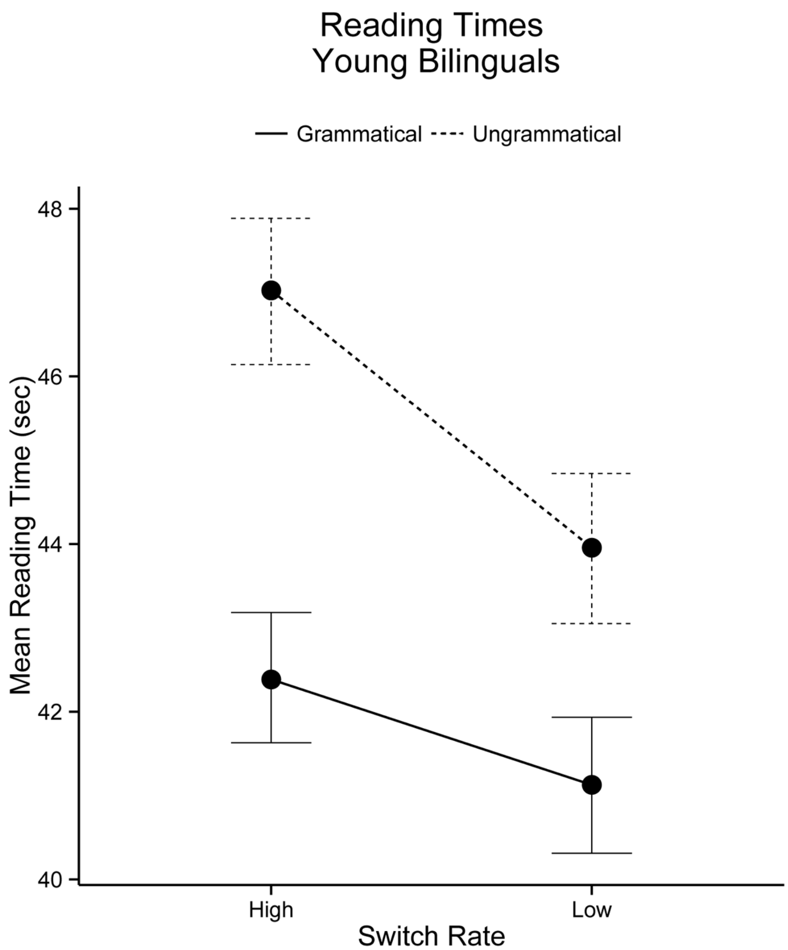

Figure 1.

By-participant mean whole paragraph reading times by condition, Experiment 1. Error bars depict bootstrapped 95% confidence intervals for means (estimated with 1,000 bootstrap samples[4]).

[4]In a bootstrap procedure, the distribution of a statistic is estimated by repeatedly re-sampling from the observations with replacement. This does not require assuming the statistic is normally distributed. Given that most of our dependent measures are (non-normally distributed) proportions, we elected to use this method throughout.

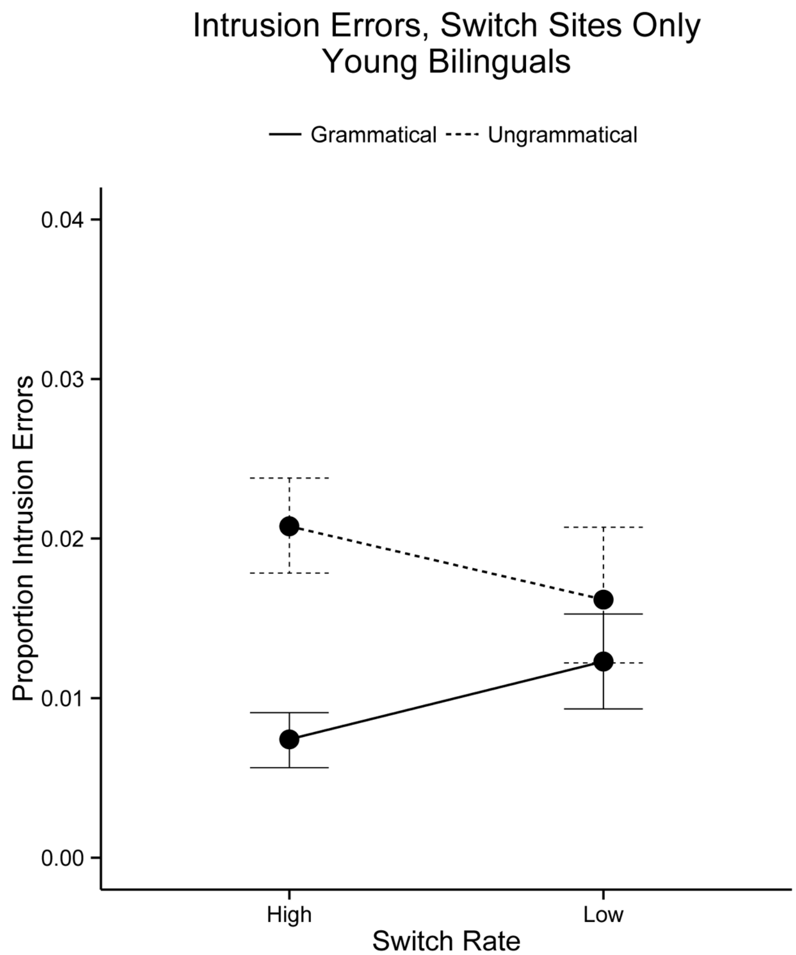

Figure 2.

By-participant mean intrusion error rate at switch sites within each experimental condition, Experiment 1. Error bars depict bootstrapped 95% confidence intervals for means.

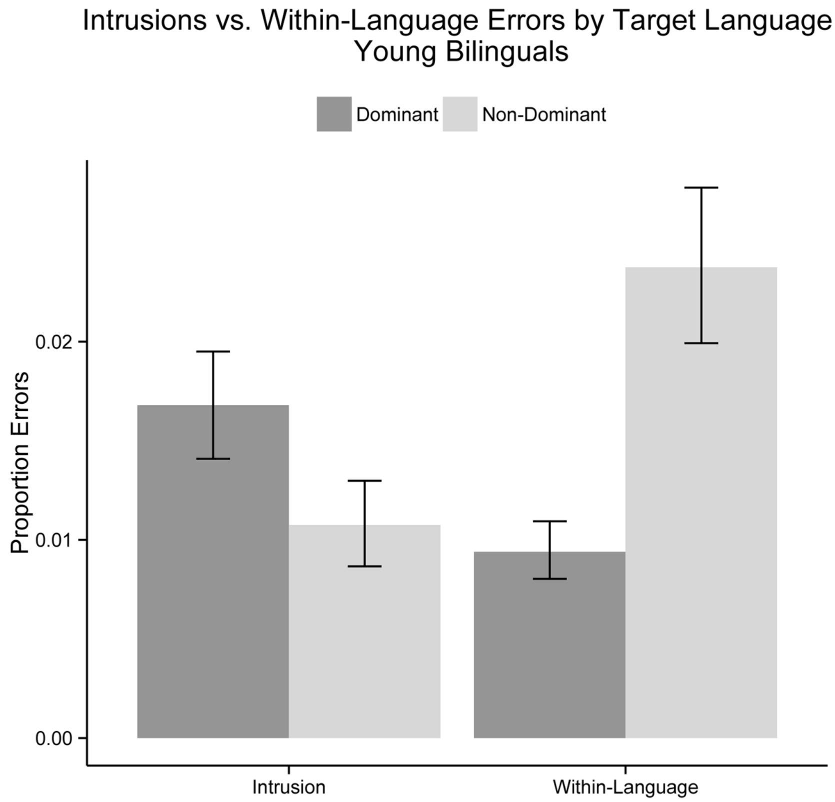

Figure 3.

By-participant mean intrusion error rate (at switch sites) and mean within-language error rate (all positions), English dominant participants. Error bars depict bootstrapped 95% confidence intervals for means.

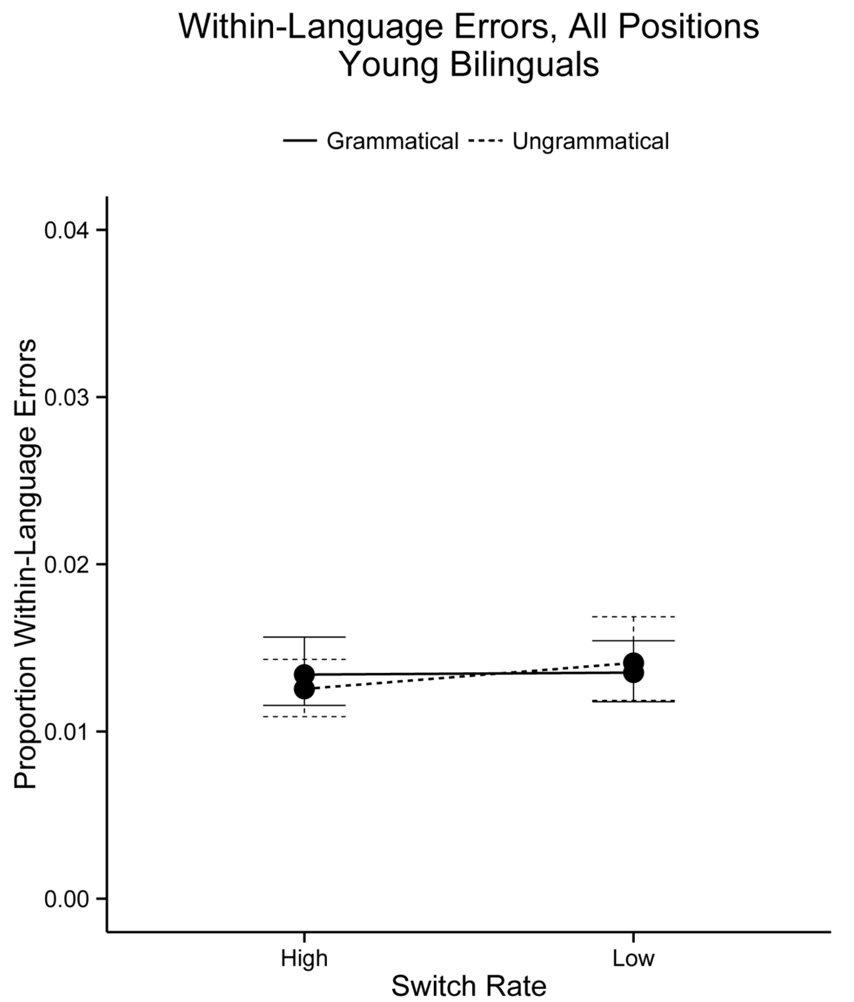

Figure 4.
Mean within-language error rate (across participants) within each experimental condition, Experiment 1. Error bars depict bootstrapped 95% confidence intervals for means.

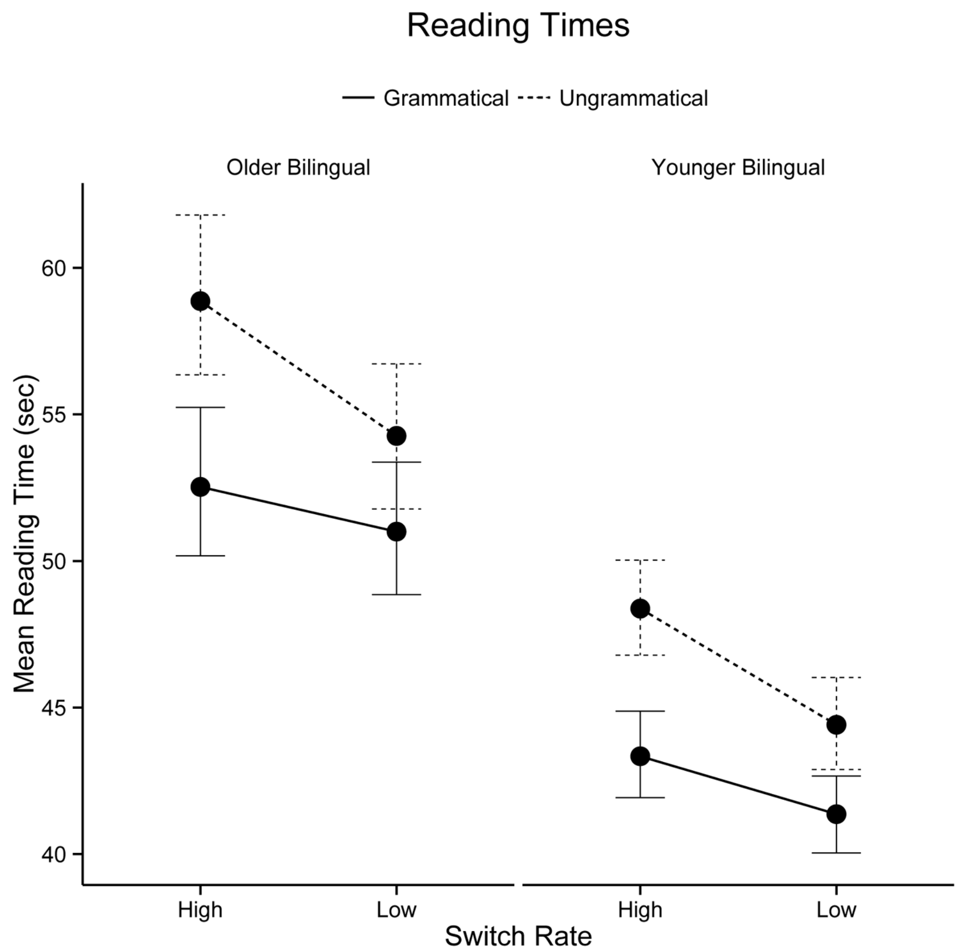

Figure 5.

By-participant mean whole paragraph reading times by condition (performance of matched younger adults from Experiment 1 provided for comparison). Error bars depict bootstrapped 95% confidence intervals for means.

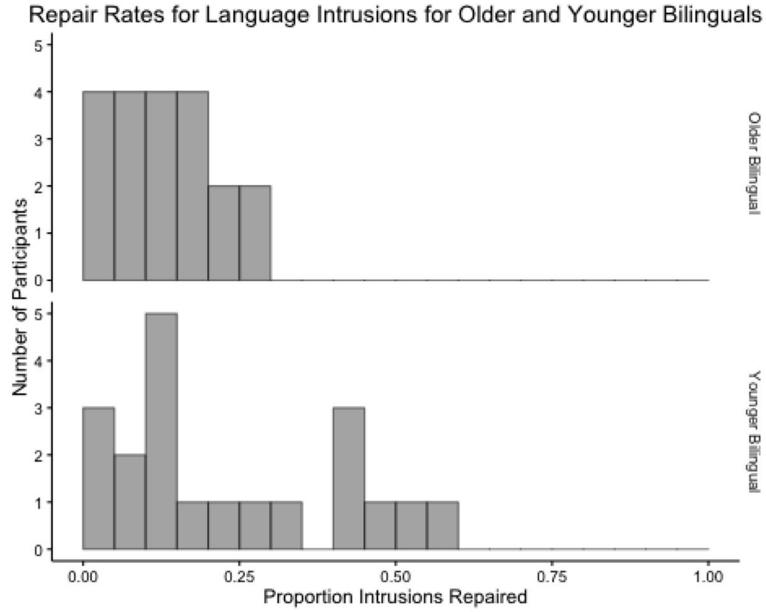

Figure 6.

Distribution of by-participant repair rates for older and matched younger bilinguals, Experiment 2.

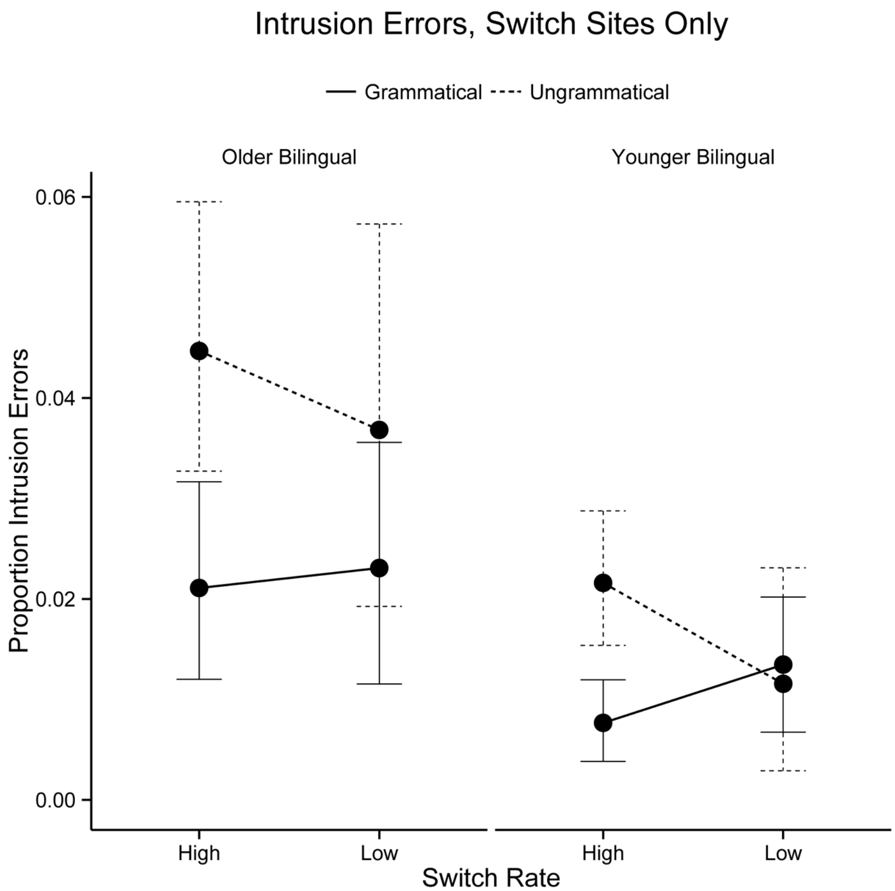

Figure 7.

By-participant mean intrusion error rate at switch sites within each experimental condition (older bilinguals and matched younger bilinguals). Error bars depict bootstrapped 95% confidence intervals for means.

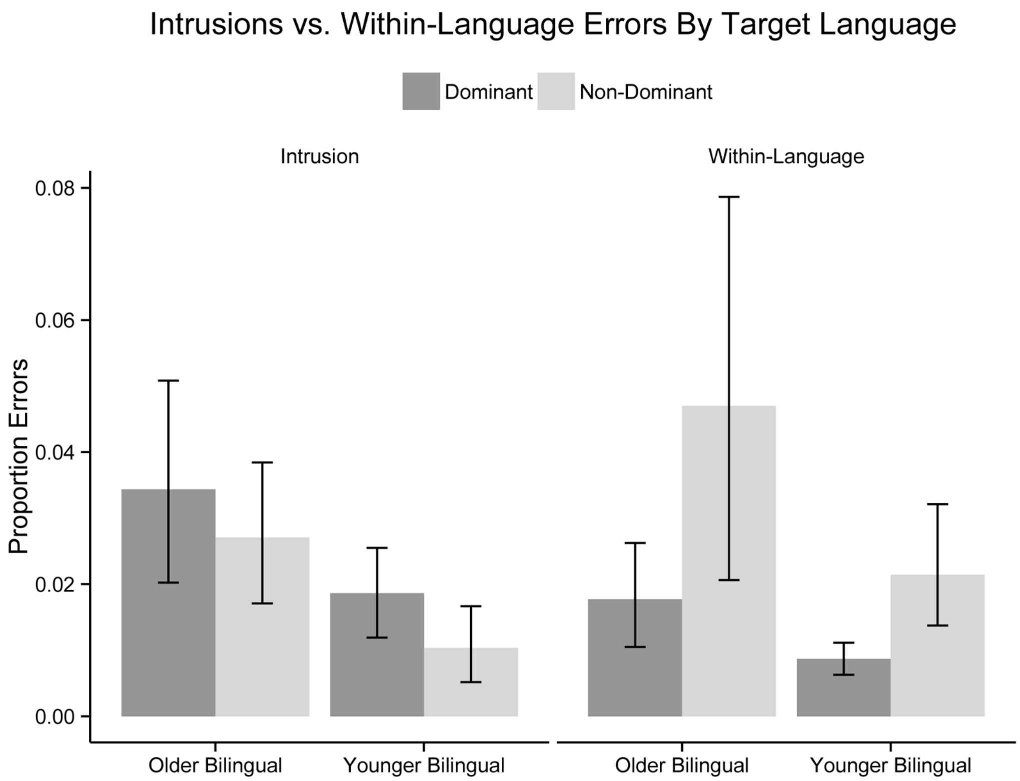

Figure 8.

By-participant mean intrusion error rate (at switch sites) and mean within-language error rate (all positions), English dominant participants (older bilinguals and matched younger bilinguals). Error bars depict bootstrapped 95% confidence intervals for means.

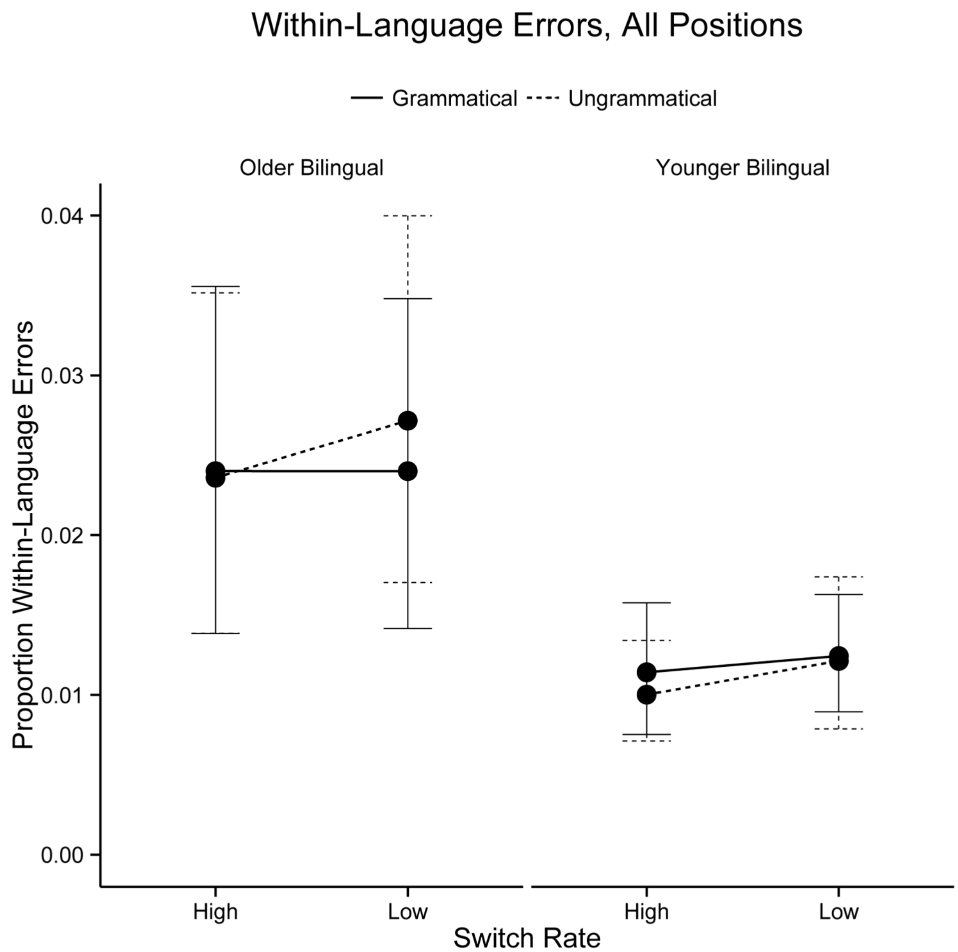

Figure 9.
Mean within-language error rate (across participants) within each experimental condition (older bilinguals and matched younger bilinguals). Error bars depict bootstrapped 95% confidence intervals for means.

Table 1

Participant characteristics.

	Younger bilinguals (n = 97[a])		Older bilinguals (n = 20)		Matched Younger bilinguals (n = 20)		Comparison of Older and Matched Younger	
	M	*SD*	*M*	*SD*	*M*	*SD*	*t*	*p*
Age	20.3	2.0	75.7	8.8	20.2	1.2	106.19	<.01
Dementia Rating Scale score	--	--	137.8	3.4	--	--	--	--
Mini Mental Status Exam score	--	--	28.6	1.8	--	--	--	--
Education	13.8	1.2	14.3	2.7	14.1	1.0	<1	.82
Age of English Acquisition	3.1	2.5	4.8	3.7	3.9	2.5	<1	.35
English proficiency self-rated[b]	6.5	0.8	6.5	0.6	6.4	0.7	<1	.68
Spanish proficiency self-rated[b]	5.8	0.9	5.5	1.3	5.9	0.7	<1	.28
Current percent English use	80.9	15.6	77.3	22.9	86.9	9.5	1.40	.09
Percent English use during childhood	56.7	17.9	52.4	31.3	62.0	11.9	<1	.21
English MINT score[c]	60.4	3.8	63.1	4.0	63.0	2.1	1.23	.88
Spanish MINT score[c]	45.8	10.2	52.4	7.8	52.9	5.9	<1	.82
Trail Making Test[d]								
Trails A	26.1	7.0	42.0	12.7	26.1	5.6	5.76	<.01
Trails B	54.5	15.2	122.2	53.3	55.6	14.2	6.20	<.01
Trails B errors	0.4	0.7	1.1	1.4	0.4	0.8	1.42	.09
B minus A	28.4	13.5	80.3	49.9	29.5	11.9	4.75	<.01
B/A	2.2	0.6	3.0	1.2	2.2	0.5	2.69	.01
Flanker Task[e]								
Congruent	527	165	801	202	515	191	4.99	<.01
Neutral	510	160	730	170	493	173	4.68	<.01
Incongruent	571	185	897	276	540	186	5.28	<.01
Flanker Effect[f]	61	64	167	169	48	24	3.06	<.01
Flanker Measure[f]	12.1	10.2	22	21	10	5	2.52	.02
Stroop[g]								
Congruent	732	105	884	160	751	108	3.00	<.01
Neutral	756	104	953	147	779	112	4.45	<.01

	Younger bilinguals (n = 97[a])		Older bilinguals (n = 20)		Matched Younger bilinguals (n = 20)		Comparison of Older and Matched Younger	
	M	***SD***	***M***	***SD***	***M***	***SD***	***t***	***p***
Incongruent	852	141	1119	180	884	175	4.41	<.01
Stroop Effect[f]	97	68	166	93	105	86	1.89	.04
Stroop Measure[f]	12.7	8.4	18	11	13	10	1.48	.15

[a]For executive function measures, means exclude the two individuals that were excluded from analysis due to failure to complete executive function tasks.

[b]Self-rated proficiency level was averaged across ratings for speaking, comprehension of spoken speech, reading, and writing on a scale from 1 (“little to no knowledge”) to 7 (“like a native speaker”).

[c]Maximum possible score is 68.

[d]Error rates were extremely low for both young and older bilinguals and there were no significant aging effects in any condition

[e]Error rates were extremely low (<1%) in all conditions except for older adults in the incongruent condition (2%).

[f]Flanker and Stroop effects were calculated by subtracting neutral from incongruent times. Flanker and Stroop measures were calculated as follows (incongruent minus neutral)/neutral

[g]Stroop performance is reported for the language with the fastest incongruent trial reaction times. Error rates were very low in all conditions except for the incongruent condition in which both young and older bilinguals produced 5% errors. Aging effects on error rates were not significant.

Table 2

Summary of regressions examining effects of individual differences in executive function in younger bilinguals. Significant effects are shown in bold; the second column summarizes the predicted effect in words (if observed) and the cells to the right provide coefficient values (standard errors in parentheses; *: $p < 0.05$, **: $p < 0.005$, ***: $p < 0.0005$). Non-bold cells give coefficients with $0.15 > ps > 0.05$. Cells with dashes have coefficients with $ps > 0.15$.

		Stroop	Flanker	Trails
Reading Times	Main Effect: Longer reading times (RTs)	--	--	2.03 (1.13)
	Interaction with Type: Reading Time Even Longer on Ungrammatical	**10.96 (3.1)*** **	−4.63 (2.54)	**1.79 (0.43)*** **
	Interaction with Type, Rate	--	--	--
Intrusion Errors	Main Effect: Increased intrusions	--	--	**0.31 (0.12)***
	Interaction with Type	--	--	--
	Interaction with Type, Rate	--	--	--
Within Errors	Main Effect	--	--	--

Table 3

Summary of regressions examining effects of individual differences in executive function in older bilinguals. Significant effects are shown in bold; the second column summarizes the predicted effect in words and the cells to the right provide coefficient values (standard errors in parentheses; *: $p < 0.05$, **: $p < 0.005$, ***: $p < 0.0005$). Non-bold cells give coefficients with $0.15 > ps > 0.05$. Cells with dashes have coefficients with $ps > 0.15$.

		Stroop	Flanker	Trails
Reading Times	Main Effect: Longer reading times (RTs)	--	--	--
	Interaction with Type: Reading Time Even Longer on Ungrammatical	--	--	--
	Interaction with Type, Rate	--	--	--
Intrusion Errors	Main Effect: Increased intrusions	--	--	**0.42 (0.13)****
	Interaction with Type	--	--	−0.26 (0.14)
	Interaction with Type, Rate	--	--	--
Within Errors	Main Effect: Increased within-language errors	**4.08(1.66)***	--	--