Research Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2014, Vol. 25(2) 585–595

DOI: 10.1177/0956797613512661
pss.sagepub.com
SAGE

Multiple Levels of Bilingual Language Control: Evidence From Language Intrusions in Reading Aloud

Tamar H. Gollan[1], Elizabeth R. Schotter[2], Joanne Gomez[1], Mayra Murillo[1], and Keith Rayner[2]
[1]Department of Psychiatry and [2]Department of Psychology, University of California, San Diego

Abstract

Bilinguals rarely produce words in an unintended language. However, we induced such intrusion errors (e.g., saying *el* instead of *he*) in 32 Spanish-English bilinguals who read aloud single-language (English or Spanish) and mixed-language (haphazard mix of English and Spanish) paragraphs with English or Spanish word order. These bilinguals produced language intrusions almost exclusively in mixed-language paragraphs, and most often when attempting to produce dominant-language targets (accent-only errors also exhibited reversed language-dominance effects). Most intrusion errors occurred for function words, especially when they were not from the language that determined the word order in the paragraph. Eye movements showed that fixating a word in the nontarget language increased intrusion errors only for function words. Together, these results imply multiple mechanisms of language control, including (a) inhibition of the dominant language at both lexical and sublexical processing levels, (b) special retrieval mechanisms for function words in mixed-language utterances, and (c) attentional monitoring of the target word for its match with the intended language.

Keywords

bilingualism, language control, reading aloud, speech error, intrusion error, eye movements, lexical access, phonology

Received 10/18/12; Revision accepted 10/21/13

Proficient bilinguals have remarkable control over two language systems, maintaining separation between them when speaking to monolinguals, fluently mixing languages when they wish to, and almost never producing words in an unintended language by mistake. Even if rare, unintentional language switches—*cross-language intrusion errors*—provide a unique, rich, and largely unexplored source of evidence about how bilinguals maintain control over language selection. Few investigations of intrusions have been conducted, perhaps because such errors are difficult to induce in experimental settings. In one study of category fluency, young bilinguals naming category members produced intrusions less than 1% of the time, and aging bilinguals at most 3% of the time (Gollan, Sandoval, & Salmon, 2011). The rarity of intrusions in aging bilinguals implies language-specific control mechanisms that remain relatively unaffected by aging-related cognitive decline. However, aging bilinguals with deficits in a nonlinguistic flanker task produced the most intrusions in that study. Thus, language control may be maintained both by language-specific mechanisms and by domain-general mechanisms that support both linguistic and nonlinguistic tasks (Weissberger, Wierenga, Bondi, & Gollan, 2012).

Two prominent hypotheses about bilingual language control could play a key role in explaining intentional and unintentional language mixing. One view assumes that bilinguals inhibit the dominant language (Green, 1986, 1998) to enable switching to the nondominant language.

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

Supporting this view are studies showing that cued-language-switching tasks produce larger switch costs for the dominant than for the nondominant language (Meuter & Allport, 1999). *Dominance reversal*, in which bilinguals respond more slowly in their dominant than in their nondominant language, provides further evidence of inhibition. Reversal has been reported both for cued language mixing (e.g., Christoffels, Firk, & Schiller, 2007; Costa & Santesteban, 2004; Verhoef, Roelofs, & Chwilla, 2009) and for voluntary language mixing (Gollan & Ferreira, 2009).

Another, mutually compatible, asymmetry in bilingual language control was hypothesized by Myers-Scotton (1993, 1997, 2002), who suggested that in mixed-language utterances, one language functions as the *matrix language*, providing syntactic frames and the majority of words, morphemes, and inflections, as well as dictating word order. Within this view, function words in the matrix language should be retrieved relatively automatically, which would reduce or prevent intrusion errors. In addition to function words, other grammatical elements (e.g., language-specific requirements for word order) would be expected to come from the matrix language, and mixed-language utterances that violate these constraints should be difficult to produce.

Surprisingly from this perspective, Poulisse (1999; Poulisse & Bongaerts, 1994) found that most of Dutch-English bilinguals' intrusions involved function word targets (articles, pronouns, conjunctions, and editing expressions, such as "I mean"); this was true in object naming, design description, story retelling, and a short interview. These bilinguals produced intrusions at most 1% of the time, and fewer intrusions when speaking Dutch than when speaking English (their late-learned, nondominant language). Kolers (1966) reported a similar result for proficient French-English bilinguals who read aloud paragraphs that alternated "haphazardly" (p. 358) between languages; the bilinguals sometimes inadvertently said the translation of a written word rather than the word itself. Again, most of these intrusion errors involved function word targets.

We explored bilingual language control by investigating language-dominance and word-order effects on intrusion errors for function versus content words, using Kolers's (1966) paradigm. Although different from natural language production, the reading-aloud paradigm allows elicitation of connected speech and rapid production of many words, which increases statistical power for observing patterns in intrusions (which are normally infrequent).

We hypothesized that when bilinguals mix languages voluntarily, they inhibit the dominant language (Gollan & Ferreira, 2009). If similar mechanisms support reading aloud mixed-language passages, intrusion errors in this task might exhibit dominance reversal (i.e., for English-dominant bilinguals, English words would slip into Spanish more often than the reverse). We further hypothesized that function words would be relatively immune to intrusion errors when they matched the matrix language (e.g., English function words would be less likely to be accidentally replaced by Spanish words in paragraphs with English word order, and vice versa). Kolers (1966) reported that bilinguals were equally likely to substitute English for French words and French for English words, but did not report whether intrusions were modulated by target-language word order. Also, he tested only a small number of bilinguals (n = 12), both French-dominant and English-dominant, which could have obscured dominance effects.

Given the apparently consistent vulnerability of function words to intrusion errors across multiple paradigms (Hartsuiker & Declerck, 2009; Kolers, 1966; Poulisse, 1999), we used eye movement data to explore whether this vulnerability could be attributed to differences in attention given to accessing and monitoring content versus function words (Poulisse & Bongaerts, 1994). In silent reading, where a person looks (i.e., overt attention) is an indicator of where the person attends (i.e., covert attention; Rayner, 2009). Monolinguals often skip words (Schotter, Angele, & Rayner, 2012), allocating less overt attention to them, and function words are skipped more often than content words (35% vs. 15%, respectively; e.g., Carpenter & Just, 1983; Rayner, 1998). During reading aloud, the eyes are often ahead of the voice (Buswell, 1922; Inhoff, Solomon, Radach, & Seymour, 2011), such that overt and covert attention are separated. In mixed-language paragraphs, this could lead bilinguals to plan production of a word in one language while looking at a word in the other language; increased vulnerability to intrusions in such cases could reveal the role of attention in maintaining language control.

Method

Participants

Thirty-two Spanish-English bilinguals at the University of California, San Diego, participated for course credit. Table 1 summarizes participants' characteristics. Most were English-dominant bilinguals, according to their picture-naming test scores; 2 were closely balanced bilinguals.

Materials and procedure

Participants completed a language-history questionnaire and a picture-naming test before the reading-aloud task. Stimuli for the picture-naming test were 33 pictures (even-numbered items from the Multilingual Naming Test; Gollan, Weissberger, Runnqvist, Montoya, & Cera, 2012). For the reading-aloud task, we selected 16 paragraphs, 108

Table 1. Participants' Characteristics

Characteristic	*M*	*SD*
Age (years)	20.5	2.0
Daily use of English (%)	81.5	14.7
Daily use of English during childhood (%)	58.1	17.2
Age of first exposure to English (years)	3.7	3.0
Age of first exposure to Spanish (years)	0.7	1.4
English picture-naming score[a]	29.4	1.8
Spanish picture-naming score[a]	23.6	4.8
Years lived in Spanish-speaking country	1.2	2.7
Frequency of saying a word in the other language without meaning to[b]	2.3	0.9
Self-rated English proficiency[c]		
Speaking	6.4	1.2
Reading	6.3	1.0
Writing	6.3	1.0
Listening	6.5	1.1
Self-rated Spanish proficiency[c]		
Speaking	5.9	1.4
Reading	5.8	0.9
Writing	5.3	1.0
Listening	6.5	0.9

[a]The maximum possible score on these tests was 33. [b]The 6-point rating scale included the following anchors: 1 (*never*), 2 (*very infrequently*), 3 (*occasionally*), and 6 (*constantly*). [c]Ratings were on a scale from 1 (*little to no knowledge*) to 7 (*like a native speaker*).

words long on average (*SD* = 11), from short stories that had been published in both English and Spanish. We created two mixed-language versions of each story that matched the mixing frequency in the example published by Kolers (1966). Each of three Spanish-English bilinguals read 16 paragraphs, 4 in each of the four conditions: (a) English only, (b) Spanish only, (c) mixed language with English word order, and (d) mixed language with Spanish word order (see the Supplemental Material available online). Paragraphs were rotated across conditions between subjects in a Latin-square design. Paragraph order was randomized uniquely for each bilingual.

Eye movements were recorded via an SR Research Ltd. (Mississauga, Ontario, Canada) Eyelink 1000 eye tracker with a temporal resolution of 500 Hz. Head restraint was not used, but head position was monitored. After calibration, eye-position error was less than 1°. Participants were seated 60 cm from a 20-in. CRT monitor with a resolution of 1,280 × 1,024 pixels. Although viewing was binocular, only movements of the right eye were recorded. At the start of the experiment, participants completed a nine-point calibration and validation procedure to allow monitoring of both horizontal and vertical eye movements. At the start of each trial, a black box (65 pixels wide and 85 pixels tall) appeared in the top-left corner of the screen, where the first word would appear. When a fixation was detected in this box, it disappeared, and the paragraph came on-screen. Paragraphs were presented as black letters on a white background in 32-point Courier New font, with 1.7 letters equaling 1° of visual angle. Participants were instructed to read the paragraphs aloud as accurately as possible at a comfortable pace. A bilingual experimenter recorded errors and later checked the coding using audiovisual recordings with the audio time-locked to a video of the eye movement record.

Results

Errors were classified as either (a) intrusions (e.g., saying *pero* instead of *but*), (b) partial intrusions (starting to produce an intrusion but self-correcting before producing the error), (c) accent errors (e.g., saying the correct word with the accent of the nontarget language), and (d) within-language errors (e.g., saying *such* instead of *much*). All participants produced at least one intrusion; the maximum number was 30. Accent errors (see also Kolers, 1966) were identified as such by two Spanish-English bilingual assistants; none of the bilinguals tested had a strong accent in either language in spontaneous speech. We focus primarily on intrusions and report analyses on the effects of language mixing, part of speech (function vs. content), language of the target word, and word order (English vs. Spanish) on production of errors. In a final section, we report analyses of eye movement data to consider if part-of-speech effects were modulated by word skipping and by looking at words in the nontarget language. Data were analyzed using logistic regressions (Jaeger, 2008), in which *b* values represent effect sizes in logit space (for details on the regression models and corresponding analyses of variance, see the Supplemental Material).

Factors that elicited intrusion errors

Table 2 shows the number of errors produced in each language in each condition. Participants produced intrusion errors almost exclusively in mixed-language paragraphs—significantly more often than in single-language paragraphs ($b = 9.39$, $SE = 1.41$, $z = 6.64$, $p < .001$). This suggests that it is difficult for bilinguals to mix languages haphazardly (Dussias, 2003; Sridhar & Sridhar, 1980). However, language mixing did not increase errors in a generalized way; bilinguals produced *fewer*[1] within-language errors in the mixed-language than the single-language paragraphs ($b = -1.40$, $SE = 0.16$, $z = 8.86$, $p < .001$). Because our primary goal was to characterize intrusion errors, we did not test for an interaction (but see the Supplemental Material). However, the significant effect of condition was in opposite directions for the two types of errors (note the sign difference in the *b* values).

Numbers of errors are shown separately for content and function word targets in Table 3. Overall, the majority of intrusions involved function words (mostly articles,

Table 2. Total Number of Errors of Each Type Produced in Each Language in Each Condition

Language of target word and error type	Single-language paragraphs[a]		Mixed-language paragraphs		Total
	English only	Spanish only	English word order	Spanish word order	
English					
Intrusion	1	—	104	97	202
Partial intrusion	1	—	14	25	40
Accent error	15	—	75	58	148
Within-language error	43	—	33	27	103
Spanish					
Intrusion	—	2	50	36	88
Partial intrusion	—	0	4	2	6
Accent error	—	2	15	36	53
Within-language error	—	113	54	56	223

[a]Paragraphs written in English presented no opportunities to err on Spanish target words, and paragraphs written in Spanish presented no opportunities to err on English target words.

pronouns, prepositions, conjunctions, and quantifiers) rather than content words (mostly nouns, some adjectives and verbs; $b = 0.72$, $SE = 0.21$, $z = 3.41$, $p < .001$). In contrast, and as observed previously for speech errors produced by monolinguals (Garrett, 1982; Levelt, 1989), within-language errors showed an effect in the opposite direction: fewer errors with function than with content word targets ($b = -0.60$, $SE = 0.15$, $z = 4.15$, $p < .001$). Note that proper nouns did not induce any intrusion errors, though they did elicit some accent errors.

Among content word targets, intrusions were about 5 times more likely to involve cognates[2] (e.g., *familia* → *family*) than to involve noncognates (e.g., *amigo* → *friend;* $b = -3.96$, $SE = 0.76$, $z = 5.20$, $p < .001$), but function words showed a trend in the opposite direction (Table 3). Because we did not manipulate cognate status (only 9% of words were cognates, excluding proper nouns) and few function words are Spanish-English cognates, we do not consider these effects in detail but note that if cognates induce intrusions, this may involve mechanisms similar to those that cause cognates to trigger intentional language switches on subsequent words (Broersma, 2009; Broersma, Isurin, Bultena, & de Bot, 2009; Kootstra, van Hell, & Dijkstra, 2012).

Do bilinguals inhibit the dominant language to achieve language mixing?

Figure 1 shows the effect of the target language on the number of errors, collapsed across conditions. Demonstrating their English dominance, participants produced significantly more within-language errors when attempting to produce Spanish, compared with English, targets ($b = 0.96$, $SE = 0.20$, $z = 4.87$, $p < .001$). In contrast, the pattern for intrusions was in the opposite direction; participants produced intrusion errors significantly *less* often when

Table 3. Average Number of Errors of Each Type Produced With Content and Function Word Targets (Collapsed Across the Four Conditions)

Cognate status	Intrusion		Partial intrusion		Accent error		Within-language error	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Content word targets								
Cognate	2.6	1.7	1.2	1.4	1.0	1.3	1.5	2.3
Noncognate	0.5	1.0	0.2	0.5	1.3	1.7	4.7	5.5
Proper noun	0	—	0	—	1.3	1.7	0.0	0.2
Total[a]	3.1	2.3	1.3	1.6	3.7	2.8	6.3	7.7
Function word targets								
Cognate	1.0	0.9	0.0	0.0	0.1	0.2	0.2	0.5
Noncognate	5.0	4.2	0.1	0.3	2.5	1.3	3.7	4.3
Total[a]	6.0	4.5	0.1	0.3	2.6	1.4	3.9	4.6

[a]Total means were calculated separately (i.e., they are not simply the sums of column means).

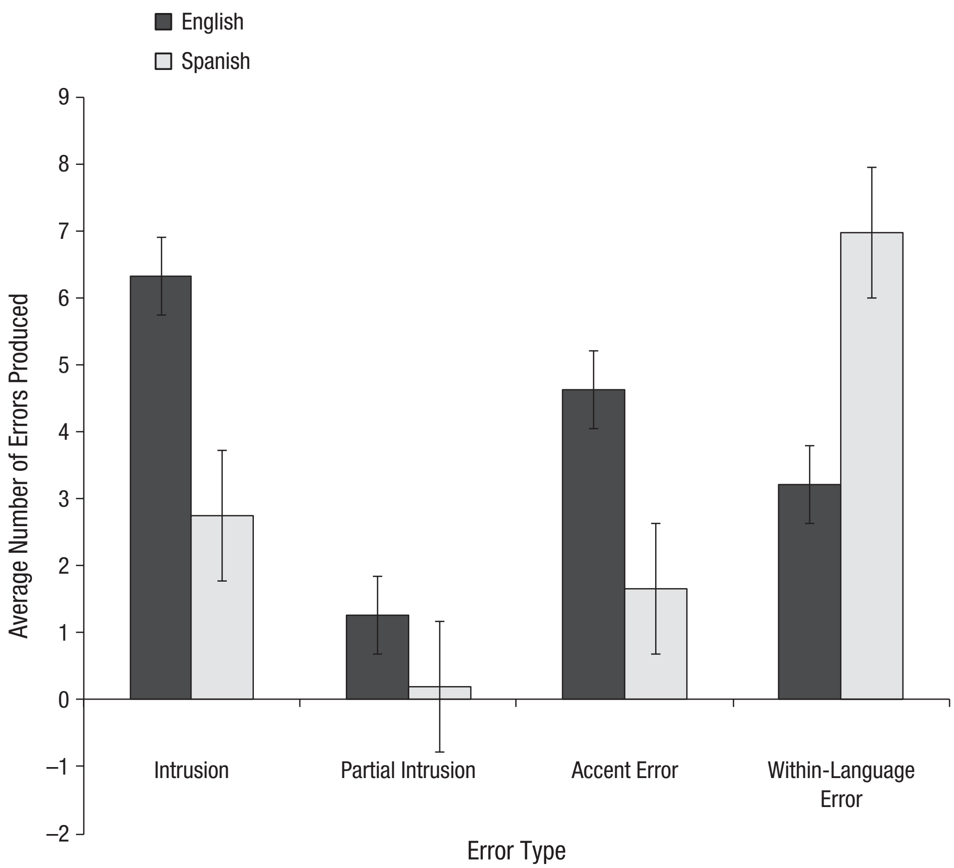

Fig. 1. Average number of errors of each type produced in each target language (collapsed across conditions). Error bars show standard errors, calculated separately for each language.

attempting to produce Spanish, compared with English, targets ($b = -0.81$, $SE = 0.24$, $z = 3.43$, $p < .001$). This result suggests that participants inhibited English when reading mixed-language paragraphs. Accent errors showed the same pattern of more errors for English than for Spanish targets ($b = -1.25$, $SE = 0.45$, $z = 2.76$, $p < .01$).

Measures of reading fluency also confirmed the bilinguals' English dominance; total reading time was faster for English-only paragraphs ($M = 31.7$ s, $SD = 5.6$) than for Spanish-only paragraphs ($M = 40.3$ s, $SD = 7.5$), $t(31) = 8.78$, $p < .001$. Results for reading time also indicated that, as expected, it was difficult to read aloud text with a haphazard mixture of the two languages (Kolers, 1966); the single-language paragraphs were read faster than the mixed-language paragraphs with English word order ($M = 44.1$ s, $SD = 8.9$) and Spanish word order ($M = 45.6$ s, $SD = 8.4$; both ps $< .01$).

Does word order facilitate retrieval of function words?

Figure 2 illustrates how the rate of intrusions varied with the type of word target, the language of the target, and the word order of the paragraph. These rates were adjusted by the opportunity to err; that is, the calculation took into account how many function or content words there were in each language (English-word-order paragraphs had about twice as many English function words as Spanish-word-order paragraphs, and vice versa).

We analyzed the incidence of intrusions in mixed-language paragraphs as a function of matrix language (English word order, Spanish word order), target language (English, Spanish), and part of speech (function, content). Because intrusions were rare, there were many words for which an error was never observed, which led to difficulties using traditional logistic regressions (i.e., the log odds of 0 is undefined). Thus, we performed a linear regression with subjects and paragraphs (instead of individual words) as crossed random effects with an empirical logit transform (Barr, 2008), using the maximal-random-effects structure (Barr, Levy, Scheepers, & Tily, 2013). Given the large number of observations, the t distribution approached the normal distribution, and absolute t values of 1.96 or higher indicated significance at the .05 level.

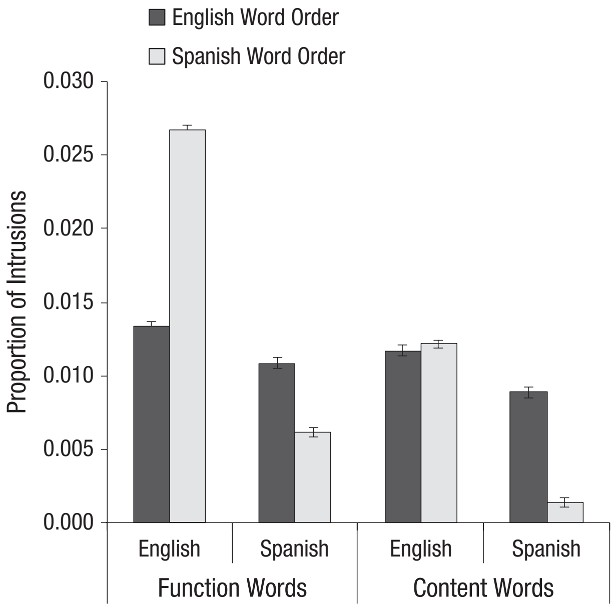

Fig. 2. Average proportion of intrusions for English and Spanish content and function words in paragraphs with English and Spanish word order. These proportions were calculated by dividing the number of intrusion errors by the number of opportunities to err (i.e., the number of content or function words in each language). Error bars show standard errors, calculated separately for each word order.

Participants produced more intrusions with English than with Spanish targets ($b = -0.24$, $SE = 0.11$, $t = 2.23$). A significant interaction between matrix language (i.e., word order) and target language ($b = 0.81$, $SE = 0.19$, $t = 4.18$) reflected that fact that there were more intrusions when the target language did not match the matrix language than when they did match. Analyzing function and content targets separately, we found that mismatch between the target language and matrix language had a significant effect for function words ($t = 5.43$), but not content words ($t = 1.05$). Thus, there was a three-way interaction of matrix language, target language, and part of speech ($b = -0.93$, $SE = 0.39$, $t = 2.39$). No other effects were significant (ts < 1.08).

Eye movements suggest vulnerability of function words to contextual distraction

An important question is whether skipping words or fixating on words in the wrong language can explain the observed part-of-speech effects (i.e., why function words were more vulnerable to intrusions than content words were) and reversed language-dominance effects. Table 4 shows skipping rates for function and content targets that were produced as intrusions and within-language errors and that were produced (by different participants) correctly.

Table 4. Average Percentage of Targets That Were Skipped When the Targets Were Produced Incorrectly (Intrusions and Within-Language Errors) and When They Were Produced Correctly (Collapsed Across the Four Conditions)

Error type and target type	Target word produced as an error		Target word produced correctly	
	M	*SD*	*M*	*SD*
Intrusion				
Content word	2.4	9.9	5.4	6.9
Function word	22.2	26.6	12.4	9.8
Within-language error				
Content word	2.0	6.2	1.5	7.0
Function word	19.8	32.4	2.6	7.4

For intrusions, there was no main effect of skipping ($p > .50$), but this null effect was qualified by a marginally significant interaction with part of speech ($b = 2.59$, $SE = 1.38$, $z = 1.88$, $p = .06$); bilinguals were significantly more likely to produce an error when they skipped rather than fixated on function words ($b = 0.90$, $SE = 0.28$, $z = 3.18$, $p < .005$), but not content words ($p > .33$). Within-language errors showed a similar pattern; there was no main effect of skipping ($p > .18$), but skipping affected production of function more than content words (a significant interaction; $b = -2.46$, $SE = 1.24$, $z = 1.98$, $p < .05$). Thus, skipping increased errors both within and across languages for function but not content word targets, a result that possibly reflects similar consequences of not allocating overt attention during word identification.

Note that skipping alone cannot explain differential part-of-speech effects on the two error types because the majority of targets produced as intrusions (84%) were fixated (i.e., not skipped). Part-of-speech effects on the proportion of intrusion errors were robust even after we excluded all skipped targets from the analysis ($b = 0.62$, $z = 3.05$, $p < .005$), as were reversed language-dominance effects ($b = -0.76$, $z = 3.09$, $p < .005$). However, part-of-speech effects may be partially explained by increased vulnerability of function words to contextual distraction. Whereas participants tended to look directly at content words (see Fig. 3), which prevented errors, they were more likely to not directly fixate function words. Figure 3 illustrates that a large proportion of function-word intrusion errors occurred when participants were looking at a word in the other language rather than the same language, but content words showed no such effect. Additionally, most correct responses and most intrusion errors with content word targets were produced when participants were looking directly at target words.

To determine if looking at words in the other language might explain part-of-speech effects, we assessed whether

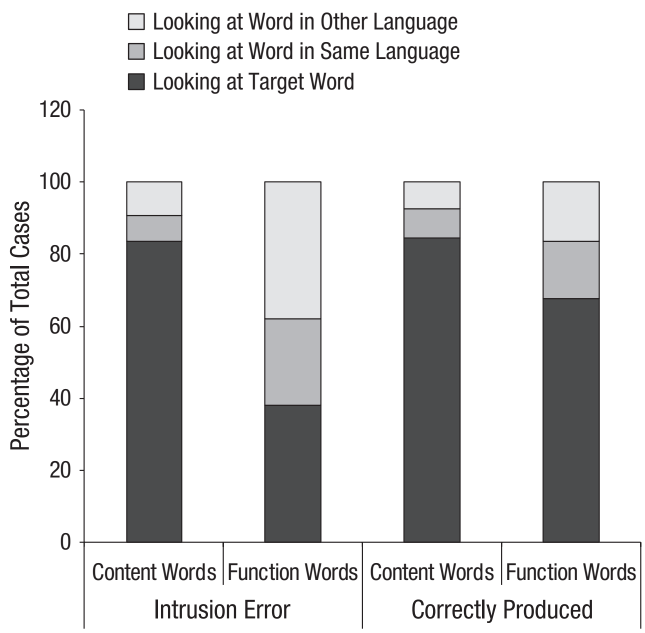

Fig. 3. Distributions of fixations when participants produced intrusions for content and function targets and when they produced those targets correctly. Each bar shows the percentage of cases in which participants were looking directly at the target word, at a different word in the same language, and at a word in the other language.

function words outnumbered content words as targets of intrusion errors only when participants looked at words in the wrong language during error production. Indeed, participants were less likely to produce an intrusion, and more likely to produce a correct response, when they looked at a word in the same language rather than a different language ($b = -1.21$, $SE = 0.34$, $z = 3.53$, $p < .001$), and this effect seemed to be driven primarily by function word targets (Table 5).

Although there was no interaction between the language of the word looked at and part of speech ($p > .82$; but see the analyses of variance in the Supplemental Material), this was likely due to the sparsity of data for content words (because participants more often than not looked directly at content words when they produced them and this analysis did not include cases in which participants looked at the target word). For function word targets, participants were significantly more likely to produce an intrusion error when looking at a word in the other language than when looking at a word in the same language ($b = 1.27$, $SE = 0.27$, $z = 4.66$, $p < .001$). The model for content word targets failed to converge (because of insufficient data).

Discussion

The results reported here imply both domain-general and language-specific mechanisms of bilingual language control, and reveal that reading aloud is useful for inducing speech errors. Although we observed a large number of language intrusions (290; compared with just 18 produced by young bilinguals in Gollan et al., 2011), the rate of errors (about 0.6%) was in line with that observed in other tasks (0.4% in Gollan et al., 2011). Reading aloud also elicited many monolingual-like errors (i.e., within-language errors; $n = 326$), and comparisons of these with intrusions revealed opposite patterns, implying distinct underlying mechanisms.

Table 5. Average Number of Content and Function Word Targets Produced as Intrusions When Participants Were Looking at the Target or at a Different Word in the Same Language or the Other Language

Word looked at during intrusion	Content word targets *M*	Content word targets *SD*	Function word targets *M*	Function word targets *SD*
Target word	2.5	2.2	2.3	2.2
Nontarget word in the same language	0.2	0.4	1.4	1.6
Nontarget word in the other language	0.3	0.5	2.3	2.1

Participants produced intrusion errors more often (a) in mixed-language as opposed to single-language paragraphs, (b) for words in their dominant rather than nondominant language (i.e., intrusions exhibited reversed language dominance), and (c) for function words as opposed to content words. In contrast, participants produced within-language errors more often (or equally often, after we adjusted for opportunities to err) (a) in single-language paragraphs, (b) for words in their nondominant language, and (c) for content words (even though paragraphs had about twice as many function as content words). A closer look at intrusion errors revealed that matrix language (word order) facilitated retrieval of the target language, particularly for function words. In addition, eye movement data revealed that part-of-speech effects on intrusion errors were partially (but not entirely) explained by absence of overt attention, and by distraction caused by looking at words in the nontarget language (in the case of function, but not content, word targets).

Reversed language-dominance effects imply that bilinguals inhibit the dominant language (Green, 1986, 1998; Misra, Guo, Bobb, & Kroll, 2012) when they intend to mix languages (Gollan & Ferreira, 2009). Thus, Spanish intruded into English more often than the reverse, even though our bilinguals were English-dominant. Reversed language dominance has been reported in cued (e.g., Christoffels et al., 2007; Costa & Santesteban, 2004; Verhoef et al., 2009) and voluntary (Gollan & Ferreira, 2009) language-switching paradigms.

The observed dominance reversal for both intrusion and accent-only errors implies that inhibition is applied at both lexical (Phillip & Koch, 2009) and postlexical processing levels because accent is specified independently from, and presumably after, lexical selection. Such processing might be easier to maintain if bilinguals represent a separate phonological inventory for each language (de Bot, 1992). Similar processing might explain how bilinguals can speak one language with the accent of the other (Grosjean, 1982). Consistent with the proposal that accent is specified postlexically is our finding that accent errors, unlike intrusion errors, were produced, on average, equally often (p = .66) with function (M = 2.6, SD = 1.4) and content (M = 2.3, SD = 2.4) word targets (or even more often with content than with function words if proper nouns are counted as content words; M = 1.3, SD = 1.7). Poulisse (1999) did not observe reversed dominance effects, perhaps because her bilinguals had a lower proficiency level, or because they were not mixing languages intentionally. Another possibility is that the combination of languages tested is critical (English could be more vulnerable to intrusions than Spanish, but see "Alternative Explanations" in the Supplemental Material).

In addition to inhibition, a second force that apparently facilitates selection of words in the intended language is syntactic and follows from Myers-Scotton's proposal (1993, 1997, 2002) that grammatical elements, including function words, are retrieved more automatically than content words in mixed-language utterances. Automatic retrieval prevents intrusions for words that match the matrix language, but induces errors for words that do not. According to this view, grammatical encoding can facilitate lexical selection; for bilinguals, syntactic frames may specify language-specific slots (which are perhaps analogous to syntactic category constraints; Dell, 1986; Garrett, 1975; Kootstra, van Hell, & Dijkstra, 2010). Although both content and function word targets tended to exhibit fewer intrusions when they matched (rather than mismatched) the matrix language's word order, the critical interaction of target language and matrix language was significant for function, but not content, word targets. These findings could imply that naturally occurring bilingual intrusions often involve function words because bilinguals temporarily lose top-down control over specification of the matrix language (not just over selection of individual words).

This possibility fits with the observation that intentional language mixing seldom involves production of a single function word (Bullock & Toribio, 2009). Muysken (2000) suggested that the restricted distribution of function words in mixed-language clauses reflects the grammatical nonequivalence of function words across languages. For example, mixing may be discouraged when translation-equivalent words occur in different orders within the sentence in the two languages (e.g., "it" in *lo quiero comer*, "it I-want to-eat," vs. *I want to eat it*). If nonequivalence discourages bilinguals from mixing languages on purpose, why does it not also prevent them from mixing languages by mistake? Poulisse and Bongaerts (1994) suggested that function words intrude more often than content words because function words in the dominant language are used with much higher frequency than their nondominant-language equivalents. This explanation does not work for our data given that the nondominant language intruded more often than the dominant language. Poulisse and Bongaerts also suggested that reduced automaticity of speech in beginning learners leaves little attention to spare for accessing and monitoring function words that convey little meaning. Confirming the role of attention, they reported that participants self-corrected a greater proportion of content than function word targets; similarly, our proficient bilinguals produced partial intrusions (i.e., self-corrections) primarily for content words (43 out of 46 cases).

Thus, part-of-speech effects could arise from an attention-based control mechanism—possibly a form-sensitive monitoring process (Levelt, 1989; Slevc & Ferreira, 2006) that checks planned utterances for their match with the intended target language, and that more easily misses short than long words in the nontarget language.[3] According to this view, function words should elicit intrusions more often than expected by chance; consistent with this prediction is our finding that 66% of the intrusion errors were function words (and 62% of words in the paragraphs were function words; this difference is small, but Poulisse & Bongaerts, 1994, found larger differences). This view also assumes that language control mechanisms are insensitive to frequency.

By contrast, we found that function words were less often targets of within-language errors than were content words, and the same pattern has been found for monolingual speech errors (Garrett, 1982; Levelt, 1989). Even though the majority of words produced in our task were function words, as is the case in spontaneous speech (55%–60%; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Poulisse & Bongaerts, 1994), function words may be less vulnerable to within-language errors because extremely high-frequency words are easier than other words to produce. Our eye movement data also support attention-based explanations of the part-of-speech effects. Only function words were vulnerable to contextual distraction, being replaced by intrusions more often when participants fixated a word in the other (as opposed to the same) language. One strategy bilinguals seemed to employ to avoid making errors was to look directly at the target words when they were producing them (see Fig. 3).

A potential limitation of this study is that we used reading aloud to elicit errors. Language selection in reading aloud will necessarily begin with word identification—and in mixed-language paragraphs, with determination of language membership (Kolers, 1966). But we hypothesize that after this point, language selection and control will be influenced primarily by processes that elicit language control in normal speech. Supporting this hypothesis is the fact that the majority of intrusion errors we observed (177/290, or 61%) involved noncognates, and thus could not possibly reflect misidentification (e.g., *pero* does not share any letters with its translation *but*). Also relevant is our replication of the previously reported predominance of function word targets in intrusions (Hartsuiker & Declerck, 2009; Poulisse, 1999) and relatively low rate of within-language errors with function word targets (Garrett, 1982; Levelt, 1989).

The proposal of multiple bilingual language-control mechanisms, including inhibition at both lexical and sublexical levels, grammatical encoding constraints, and monitoring, resembles proposals of multiple monitoring systems in monolingual language production (for a review, see Hartsuiker, Bastiaanse, Postma, & Wijnen, 2005), and might explain the rarity of language-selection failures. However, although we observed more intrusions than within-language errors in paragraphs with English and Spanish haphazardly mixed, it remains an open question if intrusions in normal bilingual speech should be classified as frequent or rare. If intrusions are produced less frequently than within-language errors involving meaning-related substitutions of whole words, this would provide strong support for models that restrict activation of intended targets via language tags (e.g., Green, 1998), which are perhaps analogous to syntactic constraints on lexical selection in monolingual production (Dell, Oppenheim, & Kittredge, 2008). Further investigation of bilingual speech errors—and how they differ from within-language semantic substitution errors—should lead to further insights about bilingual language control, and language production more generally.

Author Contributions

T. H. Gollan developed the study concept. All authors contributed to the study design. Testing and data collection were performed by J. Gomez and M. Murillo. T. H. Gollan and E. R. Schotter performed the data analysis and interpretation. T. H. Gollan drafted the manuscript, and E. R. Schotter and K. Rayner provided critical revisions. All authors approved the final version of the manuscript for submission.

Acknowledgments

The authors thank Paola (Giuli) Dussias, Gerrit-Jan Kootstra, and Vic Ferreira for helpful comments.

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.

Funding

This research was supported by the National Institute on Deafness and Other Communication Disorders (Grants 011492 and 000041) and the National Institute of Child Health and Human Development (Grants 050287 and 065829).

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Notes

1. This difference was not significant after we adjusted for opportunities to err in each language (e.g., all words were in English in English-only paragraphs, whereas there were only 60% and 40% English words, respectively, in English- and Spanish-word-order paragraphs).
2. The majority of cognates in the paragraphs were not identical across languages (e.g., *family* and *familia*); the few exceptions (e.g., *comparable*) were classified as produced correctly regardless of accent.
3. A similar process might explain why most content words that elicited intrusion errors were cognates (see Table 3), or cognate effects might instead resemble mixed-error effects in monolingual speech errors (e.g., Dell, 1986; Nozari, Dell, & Schwartz, 2011; Rapp & Goldrick, 2000).

References

Barr, D. J. (2008). Analyzing ‘visual world’ eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*, 457–474.

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

Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*, 92–111.

Broersma, M. (2009). Triggered codeswitching between cognate languages. *Bilingualism: Language and Cognition*, *12*, 447–462.

Broersma, M., Isurin, L., Bultena, S., & de Bot, K. (2009). Triggered code switching: Evidence from Dutch–English and Russian–English bilinguals. In L. Isurin, D. Winford, & K. de Bot (Eds.), *Multidisciplinary approaches to code switching* (pp. 85–102). Amsterdam, The Netherlands: John Benjamins.

Bullock, B. E., & Toribio, A. J. (Eds.). (2009). *The Cambridge handbook of linguistic code-switching*. Cambridge, England: Cambridge University Press.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago, IL: University of Chicago Press.

Carpenter, P. A., & Just, M. A. (1983). What your eyes do while your mind is reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 275–307). New York, NY: Academic Press.

Christoffels, I., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain Research*, *1147*, 192–208.

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

de Bot, K. (1992). A bilingual production model: Levelt's Speaking model adapted. *Applied Linguistics*, *13*, 1–24.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *3*, 283–321.

Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes*, *23*, 583–608.

Dussias, P. E. (2003). Spanish-English code-mixing at the auxiliary phrase: Evidence from eye-movements. *Revista Internacional de Lingüística Iberoamerican*, *2*, 7–34.

Garrett, M. F. (1975). The analysis of sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 133–177). New York, NY: Academic Press.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London, England: Academic Press.

Gollan, T. H., & Ferreira, V. S. (2009). 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*, *35*, 640–665.

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

Gollan, T. H., Weissberger, G., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). 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*, *15*, 594–615.

Green, D. W. (1986). Control, activation and resource: A framework and a model for the control of speech in bilinguals. *Brain & Language*, *27*, 210–223.

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

Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press.

Hartsuiker, R. J., Bastiaanse, R., Postma, A., & Wijnen, F. N. K. (Eds.). (2005). *Phonological encoding and monitoring in normal and pathological speech*. Hove, England: Psychology Press.

Hartsuiker, R. J., & Declerck, M. (2009, September). *Albert Costa y Julio Iglesias move up, but Fidel Castro stays put: Language attraction in bilingual language production*. Paper presented at the AMLaP 2009 conference, Barcelona, Spain.

Inhoff, A. W., Solomon, M., Radach, R., & Seymour, B. A. (2011). Temporal dynamics of the eye-voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, *23*, 543–558.

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

Kolers, P. (1966). Reading and talking bilingually. *American Journal of Psychology*, *3*, 357–376.

Kootstra, G. J., van Hell, J. G., & Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language*, *63*, 210–231.

Kootstra, G. J., van Hell, J. G., & Dijkstra, T. (2012). *Interactive alignment drives lexical triggering of code-switching in bilingual dialogue*. Manuscript submitted for publication.

Levelt, W. J. M. (1989). *Speaking*. Cambridge, MA: MIT Press.

Meuter, R. F. I., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, *40*, 25–40.

Misra, M., Guo, T., Bobb, S. C., & Kroll, J. F. (2012). When bilinguals choose a single word to speak: Electrophysiological evidence for inhibition of the native language. *Journal of Memory and Language*, *67*, 224–237.

Muysken, P. (2000). *Bilingual speech: A typology of code-switching*. Oxford, England: Cambridge University Press.

Myers-Scotton, C. (1993). *Duelling languages: Grammatical structure in codeswitching*. Oxford, England: Oxford University Press.

Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching* (2nd ed.). Oxford, England: Oxford University Press.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford, England: Oxford University Press.

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension the basis for error detection? A conflict-based theory of error detection in speech production. *Cognitive Psychology*, *63*, 1–33.

Phillip, A. M., & Koch, I. (2009). Inhibition in language switching: What is inhibited when switching between languages in naming tasks? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1187–1195.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Amsterdam, The Netherlands: John Benjamins.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, *15*, 36–57.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*, 460–499.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye movements and attention during reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*, 5–35.

Slevc, L. R., & Ferreira, V. S. (2006). Halting in single-word production: A test of the perceptual-loop theory of speech monitoring. *Journal of Memory and Language*, *54*, 515–540.

Sridhar, S. N., & Sridhar, K. K. (1980). The syntax and psycholinguistics of bilingual code mixing. *Canadian Journal of Psychology*, *34*, 407–416.

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

Weissberger, G. H., Wierenga, C. E., Bondi, M. W., & Gollan, T. H. (2012). Partially overlapping mechanisms of language and task control in young and older bilinguals. *Psychology and Aging*, *27*, 959–974.