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

Tamar H. Gollan¹, Elizabeth R. Schotter², Joanne Gomez¹, Mayra Murillo¹, and Keith Rayner²
¹Department of Psychiatry and ²Department 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 be) 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.
Supporting this view are studies showing that cued-language-switching tasks produce larger switch costs for
the dominant than for the nondominant language (Meuter & Alport, 1999). Dominance reversal, in which bilinguals
respond more slowly in their dominant than in their non-
dominant language, provides further evidence of inhibi-
tion. 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 auto-
matically, 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 lan-
guage, 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). Kolars (1966) reported a similar
result for proficient French-English bilinguals who read
aloud paragraphs that alternated “haphazardly” (p. 358)
between languages; the bilinguals sometimes inadver-
tently 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 hypoth-
esized that function words would be relatively immune
to intrusion errors when they matched the matrix lan-
guage (e.g., English function words would be less likely
to be accidentally replaced by Spanish words in para-
graphs with English word order, and vice versa). Kolers
(1966) reported that bilinguals were equally likely to sub-
stitute English for French words and French for English
words, but did not report whether intrusions were modu-
lated 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 func-
tion words to intrusion errors across multiple paradigms
(Hartsuiker & Declerck, 2009; Kolars, 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 ver-
sus 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 atten-
tion; 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
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 (Duñazas, 2003; Sridhar & Sridhar, 1980). However, language mixing did not increase errors in a generalized way; bilinguals produced \( f_{ewer} \) 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,
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 (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>
<thead>
<tr>
<th>Language of target word and error type</th>
<th>Single-language paragraphs</th>
<th>Mixed-language paragraphs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English only</td>
<td>Spanish only</td>
</tr>
<tr>
<td>Intrusion</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>Partial intrusion</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>Accent error</td>
<td>15</td>
<td>—</td>
</tr>
<tr>
<td>Within-language error</td>
<td>43</td>
<td>—</td>
</tr>
</tbody>
</table>

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   |

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

<table>
<thead>
<tr>
<th>Cognate status</th>
<th>Intrusion</th>
<th>Partial intrusion</th>
<th>Accent error</th>
<th>Within-language error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Content word targets</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognate</td>
<td>2.6</td>
<td>1.7</td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td>Noncognate</td>
<td>0.5</td>
<td>1.0</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Proper noun</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>Total(^a)</td>
<td>3.1</td>
<td>2.3</td>
<td>1.3</td>
<td>1.6</td>
</tr>
</tbody>
</table>

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).

---

Language Intrusion Errors

589

attempts 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 $p$s < .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. 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.](image-url)
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.

![Graph showing proportion of intrusions for English and Spanish content and function words](image)

**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)**

<table>
<thead>
<tr>
<th>Error type and target type</th>
<th>Target word produced as an error</th>
<th>Target word produced correctly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Intrusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content word</td>
<td>2.4</td>
<td>9.9</td>
</tr>
<tr>
<td>Function word</td>
<td>22.2</td>
<td>26.6</td>
</tr>
<tr>
<td>Within-language error</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content word</td>
<td>2.0</td>
<td>6.2</td>
</tr>
<tr>
<td>Function word</td>
<td>19.8</td>
<td>32.4</td>
</tr>
</tbody>
</table>

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...
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 ($b = -1.21, SE = 0.34, z = 3.55, p < .001$), 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; $n = 320$), and comparisons of these with intrusions revealed opposite patterns, implying distinct underlying mechanisms.

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. 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, Girard, & 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 fixedated 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 familis); 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**


