Using what's there: Bilinguals adaptively rely on orthographic and color cues to achieve language control (preprint)

Article *in* Cognition · June 2019

4 authors:

Julie Fadlon
Tel Aviv University

Chuchu Li
University of California, San Diego

Anat Prior
University of Haifa

Tamar H Gollan
University of California, San Diego

Using what's there: Bilinguals adaptively rely on orthographic and color cues to achieve language control

Julie Fadlon[a*]***, Chuchu Li*** [a]***, Anat Prior*** [b] ***and Tamar H. Gollan*** [a]

[a] University of California, San Diego. 9500 Gilman Dr, La Jolla, CA 92093, USA.

[b] University of Haifa, Abba Khoushy Ave 199, Haifa 3498838, Israel.

Abstract

We examined if bilinguals of two different language combinations can rely on novel and arbitrary cues to facilitate switching between languages in a read-aloud task. Spanish-English (Experiment 1) and Hebrew-English (Experiment 2) bilinguals read aloud mixed-language paragraphs, known to induce language intrusion errors (e.g., saying *el* instead of *the*), to test if intrusion rates are affected by: language combination, color-cues, language dominance, and part of speech. For Spanish-English bilinguals, written input is not rich in visual cues to language membership, whereas for Hebrew-English bilinguals rich cues are present (i.e., the two languages have different orthographies and are read in opposite directions). Hebrew-English bilinguals made fewer intrusion errors than Spanish-English bilinguals, and color cues significantly reduced intrusions on switches to the dominant language but not to the nondominant language, to the same extent in both bilingual populations. These results reveal powerful effects of visual cues for facilitating production of language switches, and illustrate that switching mechanisms are highly adaptable and sensitive, in that they can both recruit language- and orthography-specific cues when available and also rapidly exploit novel arbitrary cues to

[*] Corresponding author. Present address: Corresponding author. Present address: The sentence Processing Lab, Department of Linguistics, Tel-Aviv University, Mordechai Levanon Street, Tel Aviv 69978. Israel
Email address: fadlonj@tauex.tau.ac.il

language membership when these are afforded. Finally, such incidental, experimentally induced cues, were recruited even in the presence of other already powerful cues, when task demands were high.

Key words: bilingualism, code-switching, language control, production, visual cues, adaptability

Bilingual speech involves the simultaneous activation of two languages, even when only one language is being used (Colomé, 2001; Costa, Caramazza, & Sebastian-Galles, 2000; Gollan, Sandoval, & Salmon, 2011; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Hoshino & Kroll, 2008; Kroll, Bobb, Misra, & Guo, 2008; Poulisse & Bongaerts, 1994; Van Hell & Dijkstra, 2002). At the same time, bilinguals are equipped with the remarkable ability to manage language choice via control processes. Over the years, several models of language control have been proposed (Costa, Miozzo, & Caramazza, 1999; Dijkstra & van Heuven, 2002; Green, 1998; Green & Abutalebi, 2013; Poulisse & Bongaerts, 1994). However, little attention has been paid to describing the way the language control system uses available cues in the input to guide language choice (for a similar discussion, see Woumans et al., 2015).

One model addressing cue detection explicitly is the adaptive control model (Green & Abutalebi, 2013). According to this model, one of the processes responsible for language control is salient cue detection, which triggers disengagement from using one language and engagement in using another one. Like other control processes, salient cue detection is assumed to adapt to meet the demands placed upon it by the interactional context in which bilinguals engage in speech. The present study set out to investigate this assumption by examining the system's ability to make use of a cue not habitually related to language and how it might be affected by the availability of other cues in the input. The experiments we report below asked how arbitrary but salient color cues would facilitate the performance of two bilingual populations, English-Spanish and Hebrew-English, producing spoken language switches while reading aloud mixed language paragraphs. Half the time these switches were presented in the same color as the preceding text (black), while the other half of switches appeared in red to create a detectable pop-out effect. This manipulation tested the language control system's ability to use this highly prominent, yet arbitrary, cue and whether this ability is modulated by the availability of cues that are intrinsic to language use.

Spanish-English versus Hebrew-English are two language combinations that are distinguished by the availability of naturally occurring visual language-choice cues in written input. As a result, the sensitivity of these populations to a color cue manipulation offers important insights into the relationship between the adaptability of cue detection processes and the need to adapt, based on relative presence or absence of available cues in the input. The extent to which language control is affected by a novel cue given different language control demands can further our understanding of how cue detection processes adjust to demands placed upon them, and the extent to which bilinguals can use any available cues from the environment.

Background: visual cues for language choice in bilinguals

Previous studies on the effects of visuals cues on bilingual performance consist of two prominent lines of research. The first line of research examined the influence of the sociocultural characteristics of a face, on language choice in production (for reviews, see Hartsuiker, 2015; Woumans et al., 2015). Li, Yang, Scherf, and Li (2013) found that Chinese-English bilinguals' response times in a picture-naming task were faster when preceded by the presentation of a face congruent with the target language. Zhang, Morris, Cheng, and Yap (2013) report the results of a study in which Chinese- English bilinguals participated in a computer-mediated conversion with a man introduced as “Michael Li”, speaking English. During the conversation, participant saw a picture of a Caucasian (language-congruent) or a Chinese (language-incongruent) male face. Fluency (estimated by subjective ratings and words per minute), was lower in the language incongruent condition. Hartsuiker and Declerck (2009), (see also Hartsuiker, 2015) found similar effects when face-language cultural associations were weaker, that is – when faces were distinguished only by prior knowledge of the native language of the person depicted: Dutch-English bilinguals described visually presented scenes involving pictures of famous native-Dutch or native-English people (actors, musicians, politicians) and produced more language intrusion errors when the target language was incongruent with the famous person’s native-language. Cultural cues were also found to affect production in a second experiment by Zhang et al. (2013) involving Chinese-English bilinguals. Rather than famous faces, participants saw pictures of Chinese and American cultural icons (e.g., the Great Wall in China, Mount Rushmore in the United States). Participants had to complete two tasks using English (1) describe these cultural icons (2) describe a set of culturally-neutral images. The authors observed that, in both tasks, participants who saw American cultural icons were more fluent (based on subjective rating and word per minute) than the ones who saw Chinese cultural icons. Finally, Woumans et al. (2015) tested whether the association of a familiar face to a language, when not related to cultural associations, can affect performance in a production task. They used simulated Skype conversations to familiarize Spanish-Catalan and Dutch-French bilinguals with faces and associate these faces with a language. Performance in a task demanding production of associations for words produced by familiar and unfamiliar faces revealed faster responses when face-language pairings were congruent with ones presented in the familiarization stage. This demonstrates that even brief exposure can generate associations between specific faces as a cue for one or the other language in bilingual speakers.

The second line of research examined the influence of orthographic cues on performance in language detection of written words. Using a speeded language recognition task, Vaid and Frenck-Mestre (2002) found that French-English late bilinguals are sensitive to language bigram frequencies such that they are faster to classify words with letter sequences marking them as belonging to one language. Van Kesteren, Dijkstra, and de Smedt (2012) tested the sensitivity of Norwegian–English bilinguals to language specific script and bigram frequency in a language decision and lexical decision tasks. They observed that both types of information facilitated response times with a bigger effect for language specific script. Casaponsa, Carreiras, and Duñabeitia (2014) tested how bigram frequency specificity of Spanish and Basque words affects the performance of balanced and unbalanced Spanish-Basque bilinguals and Spanish monolinguals in two tasks: (a) speeded language detection (b) progressive demasking, in which subjects had to identify and type-in words that were presented as masked targets with progressively increasing display times. In the language detection task, they observed bigram frequency effects, such that all groups were fastest at identifying Basque-marked words. In the progressive demasking task, this effect was only observed with the two bilingual groups (see also Casaponsa, Carreiras, & Duñabeitia, 2015). However, Degani, Prior, & Hajajra (2017) demonstrated that Arabic-Hebrew bilinguals activated lexical items from both lexicons when presented with written Hebrew words (Arabic and Hebrew orthographies are completely non-overlapping), suggesting that even fully reliable orthographic cues are not sufficient for limiting activation to the target language exclusively (see also Prior, Degani, Awawdy, Yassin, & Korem, 2017, for analogous findings with auditory presentation).

In sum, accumulated evidence indicates that bilinguals are sensitive to visual cues that are closely related to language use, such as the sociocultural characteristics of a face, orthography, and bigram frequency of written words. In addition, it appears that this sensitivity remains intact even when cue-language associations are based on brief exposure during a familiarization stage in the absence of any sociocultural information. Previous studies hence demonstrate that the language control system is resourceful in detecting available language-choice cues. But associating faces and orthography with a particular language is something bilinguals do very regularly. As stated above, in the current study, we further tested the bilingual language control system's resourcefulness by investigating its sensitivity to an association not typical to bilingual language use, namely a highly prominent change in the color of a written word from black to red, and how it is modulated by a variation in language control demands.

Current study

To examine the effects of language control processes we induced language *intrusion errors* by asking bilinguals to read aloud mixed-language paragraphs (Kolers, 1966; Gollan, Schotter, Gomez, Murillo, & Rayner, 2014), which are paragraphs written mostly in one language (Gollan & Goldrick, 2018), but also contain a small number of 'switch-out' words that are written in the other language. This paradigm is specifically useful for investigating language-control mechanisms in connected speech, because it consistently elicits language control failures (intrusion errors) in which bilinguals spontaneously translate some of the switch words, and produce them in the non-target language by mistake. Such errors provide a striking form of evidence that reveals which types of language switches are more difficult to plan and produce, and by contrast, what makes language control relatively easier (Gollan & Goldrick, 2016, 2018; Gollan, Schotter et al., 2014; Gollan, Stasenko, Li, & Salmon, 2017; Li & Gollan, 2018; Ratiu & Azuma, 2017; Schotter, Li, & Gollan, in press), making the paradigm useful for identifying factors which constrain the language control system. Two consistent findings in this paradigm include part of speech effects and reversed language dominance effects. Specifically, and in accordance with similar observations in natural speech (Poulisse & Bongaerts, 1994), part of speech plays a role in predicting intrusion errors as bilinguals produce many more intrusions (or failures to switch) with function word targets as compared to content word targets. Furthermore, and surprisingly, bilinguals are more likely to replace words written in the dominant language with the nondominant translation than the reverse; e.g., English-dominant Spanish-English bilinguals, who are more proficient in English, nonetheless tend to fail to switch to an English word embedded in a Spanish paragraph more often than they fail to switch to a Spanish word embedded in an English paragraph (reversed dominance effects are also found in other paradigms see Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Christoffels, Firk, & Schiller, 2007; Declerck, Thoma, Koch, & Philipp, 2015; Gollan & Ferreira, 2009; Gollan, Kleinman, & Wierenga, 2014). An additional finding in previous studies is that most intrusions occurred with switches out of the default language, but switching back to the default language was much easier (relatively error free; Gollan & Goldrick, 2018).

Naturally, since input is presented in the written modality, performance in the read-aloud task depends on successful completion of reading processes. Importantly, however, previous studies that combined this paradigm with eye-tracking found that correct production of switch-out words cannot be predicted based on gaze durations or skipping rates (Gollan, Schotter et al., 2014; Schotter et al., in press). This indicates that performance in this task does not merely or largely reflect misperception during reading, in which switch costs are observed less often (Declerck, Koch, Duñabeitia, Grainger, & Stephan, in press) but instead reflects speech planning and monitoring mechanisms.

Thus, in the current study we employed this paradigm to examine the sensitivity of the bilingual language control system to existing and novel cues to an upcoming switch given a variation in the availability of other cues, intrinsic to language use. To this end, we compared performance of two bilingual populations, namely English-dominant Spanish-English bilinguals and Hebrew-dominant Hebrew-English bilinguals. Spanish and English are both written left to right and also share the Latin alphabet (e.g. *hombre* is the Spanish equivalent of *man*). In contrast, the Hebrew and the English writing systems differ both in alphabet (e.g., גבר is the Hebrew equivalent of *man*) and in writing direction, as Hebrew is written right to left. This means that Hebrew-English bilinguals are experienced in linking visual cues in the writing system to language choice. Spanish-English bilinguals, on the other hand, rely on visual cues in the writing system less habitually, since distinctions between the two writing systems are far less pronounced.

Therefore, comparison of the sensitivity of these two populations to novel visual cues can be highly informative for learning whether and how sensitivity to cues for switching depends on the presence of other cues that are more intrinsic to language use. If only one of these populations benefits from the availability of color cues, or if one population benefits more than the other, it would suggest that sensitivity to novel cues is affected by the availability of already existing cues. This relationship could manifest in several different manners. First, since Hebrew-English bilinguals are already equipped with rich cues in the input that are used habitually (alphabet, writing direction), we may observe that (a) overall, they would produce less intrusion errors as compared to English-Spanish bilinguals (b) their control system might be less attuned to detecting novel cues. If (b) is correct, then color effects should be stronger for Spanish-English bilinguals as compared to Hebrew-English bilinguals. Alternatively, since Hebrew-English bilinguals are already practiced in relying on visual cues to language membership, the opposite might be found. If so, Hebrew-English bilinguals might be better in quickly recognizing and using a novel visual cue to avoid intrusion errors, showing stronger color-cue effects than Spanish-English bilinguals. A third possibility is that the color cue would equally help both populations to avoid intrusion errors. Such a finding would suggest that the language control system is extremely adaptive in its ability to use available cues in the input, and that this ability is totally independent of pre-existing cues, and the general difficulty of the control problem across different language combinations. Finally, a fourth possibility is that color-language associations are too arbitrary to facilitate language choices and no color effect will be observed.

Two experiments compared how efficient Spanish-English and Hebrew-English bilinguals were in utilizing an arbitrary visual language choice cue, namely color change from black to red. Accordingly, Experiment 1 (Spanish-English) and Experiment 2 (Hebrew-English) manipulated word color, such that in some paragraphs switch-out words appeared in the same color as the preceding text (black) while in other paragraph switch-out words exhibited color change (i.e. appeared in red). In addition, as this is the first study using the mixed-language read aloud paradigm with Hebrew-English bilinguals, we were also interested whether factors previously observed to impact intrusion rates (dominance, part of speech) are observed in this population as well. Hence, as reported in detail below, in addition to switch color (black/red), we also manipulated the language (dominant/nondominant) and part of speech (function/content) of the switched word, creating a fully crossed factorial design.

Experiment 1

Method

Participants. Forty eight Spanish-English bilingual undergraduates at the University of California, San Diego participated for course credit. Table 1 shows self-reported participant characteristics and Multilingual Naming Test scores in both languages (MINT; Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012). All but one participant obtained higher picture-naming scores in English, hence were classified as English-dominant. One bilingual obtained identical MINT scores in English and Spanish (60/68 in both languages) but was included in the analyses below.

Materials and procedure. The paragraphs were 32 adaptations of materials used in Gollan, Schotter, et al. (2014) and Gollan and Goldrick (2018). Each paragraph was further modified to appear between subjects in one of 8 conditions that were created by crossing the three manipulated factors: default language (dominant: English/ nondominant: Spanish), part of speech (POS, content/function) and color (black/red). Each paragraph contained 8 controlled switches out of the default language on free-standing words in the non-default language (i.e. free morphemes: morphemes that can occur as a word independently of other morphemes). These were inserted while making sure they occurred at points where both languages exhibit the same linear order, hence were also in accordance with Poplack's (1980) Equivalence Constraint (for more about constraints on code-switching see Declerck & Philipp, 2015; Deuchar, 2005; Hok-Shing Chan, 2009; Muysken, 2000). Further, insertions did not include cognates or proper names, were not repeated more than twice in a paragraph, and did not separate common or idiomatic multiword phrases.

Table 1: Self-reported participant characteristics and Multilingual Naming Test (MINT) scores, Experiments 1 and 2.

		Experiment 1			**Experiment 2**		
Measure		*M*	*SD*	Range	*M*	*SD*	Range
*Age		20.45	2.14	18 – 27	26.74	3.54	19 – 39
*Age of acquisition	*Dominant	3.64	2.45	0 – 10	1.05	0.47	0.01 – 3
	*Non-dominant	1.1	1.17	0 – 5	6.84	2.32	0.1 – 10
Age of learning to read	Dominant	5.74	2.33	3 – 16	5.72	0.66	4 – 7
	*Non-dominant	5.9	2.34	3 – 14	9.47	2.26	6 – 18
*% of time speaking	*Dominant	66.4	20.62	5 – 100	79.6	24.2	5 – 100
	*Non-dominant	32.71	20.33	0 – 95	18.75	20.8	0 – 90
% of time reading	Dominant	80.85	18.08	25 – 100	75.4	23.5	20 - 100
	Non-dominant	17.69	16.11	0 – 50	23.7	21.15	0 – 80
Self-rated proficiency: speaking	Dominant	9.66	0.65	7.5 – 10	9.65	0.57	8 – 10
	*Non-dominant	8.55	1.43	4 – 10	7.62	1.36	4 – 10
Self-rated proficiency: comprehension of speech	Dominant	9.77	0.51	8 – 10	9.75	0.52	8 – 10
	*Non-dominant	9.24	1.17	6 – 10	8.4	1.29	3 – 10
Self-rated proficiency: reading	Dominant	9.58	0.79	6 – 10	9.66	0.62	8 – 10
	Non-dominant	8.36	1.54	2 – 10	8.14	1.37	4 – 10
*MINT score (percentage)	*Dominant	0.90	0.04	0.8 – 1	0.93	0.03	
	*Non-dominant	0.67	0.14	0.36 - 0.95	0.75	0.09	
*Formal education, self (years):		13.83	1.48	12 – 17	14.62	2.22	12 - 22
*Formal education, mother (years)		10.54	4.31	2 – 18	15.21	3.58	12 – 30
*Formal education, father (years)		10.25	4.49	3 – 21	14.72	3.37	5 – 24

Note. *indicates significant difference between experiments at $p <. 05$ level.

We considered prepositions, determiners, quantifiers, relative pronouns (e.g., *that*, *which*), pronouns (e.g., *he*, *she*), and auxiliary verbs (e.g., *is*, *have*) as function words and other lexical categories as content words. Two Spanish-English bilinguals confirmed that controlled switches met these conditions and did not disrupt the narrative. A Latin square design was used to create 8 experimental lists distributed between participants, such that each participant saw 32 paragraphs, with no repetitions. Condition order was fully randomized at each session. Appendix A shows an example of how paragraphs were adapted so that the same paragraph was presented in all conditions (between subjects as part of counterbalancing).

Materials were presented using Psychopy on an iMac 7 computer with a 20-inch color monitor. Participants were tested individually in a quiet room with an experimenter, who recorded their responses for data coding. They were told that they will see paragraphs on the screen and were instructed to read each one aloud as accurately as possible. They then completed two practice items representing both values of the switch-out language and color manipulations, each containing 4 content and 4 function switches. On each trial, a fixation cross appeared at the position where the first word of the paragraph would appear. To replace the fixation cross with the paragraph participants pressed the spacebar. They then read the paragraph aloud and pressed the spacebar again once they finished.

Results

Two Spanish-English bilinguals transcribed the recordings coding intrusion errors, i.e. cases where participants produced the equivalent of a word in the non-target language (i.e., full intrusion, for example, saying *man* when the target was *hombre*), or initiated such a production but self-corrected in its midst (i.e., partial intrusion, for example, saying *ma…hombre* when the target was *hombre*). Following previous studies which focused on full intrusions only, the first analysis presented below also focus on these, and on switch-out points which included our factorial manipulations. Table 2 reports the number of full and partial intrusions on switch-out words, switch back to the default-language words, and words that were not switches (non-switches). Figure 1 demonstrates the distribution of full intrusions on switch-out words by target language, color and part of speech.

Table 2*:* The number of full and partial intrusion errors by target language, part of speech and word color, Experiment 1.

		Full intrusions				Partial intrusions			
Target		Dominant		Nondominant		Dominant		Nondominant	
Type	Color	Function	Content	Function	Content	Function	Content	Function	Content
Switch Out	Black	218	17	56	6	34	16	7	10
	Red	55	4	16	1	11	12	6	3
Switch Back	Black	17	3	9	3	11	15	1	9
Non-switch	Black	13	8	14	7	4	7	0	8

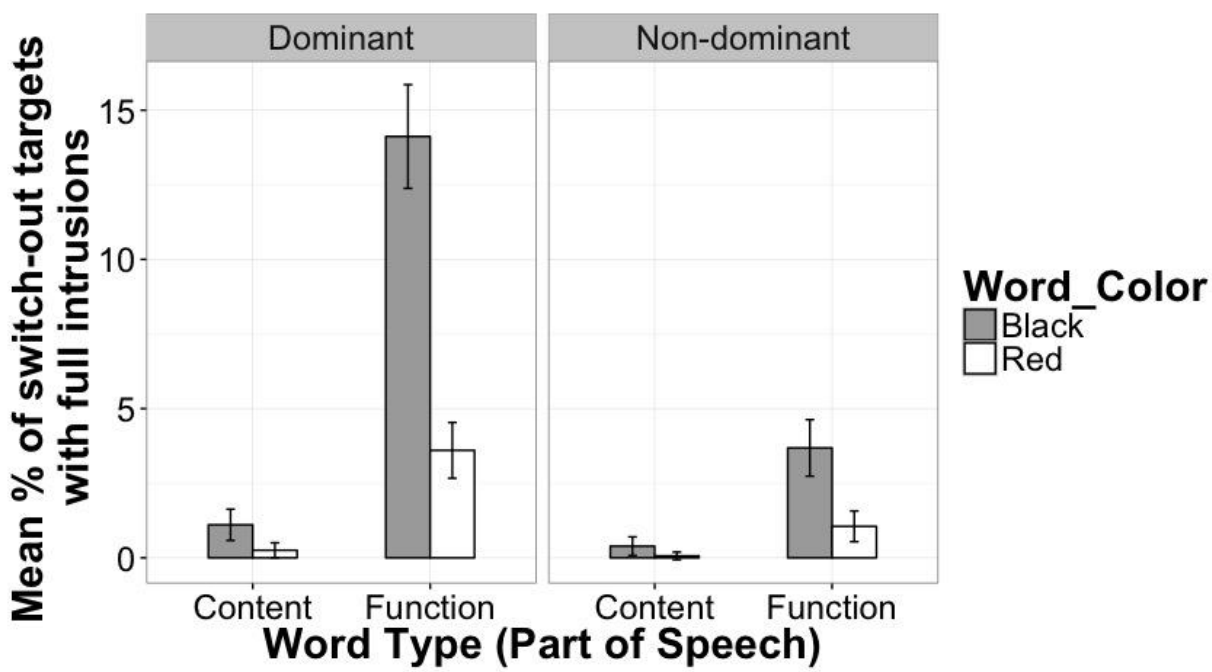

Fig. 1. Mean by participant of percentage of full intrusion errors on switch-out words by part of speech and word color in each target language (dominant vs. nondominant) in Experiment 1. Error bars represent 95% confidence intervals.

We first applied a binomial mixed effects model with full intrusions as the dependent variable, switch-out language (dominant vs. nondominant), part of speech (content vs. function), and color (black vs. red) as contrasted coded fixed effects, and subject and word as random intercepts with related random slopes. In all the analyses we report below, when a maximal model failed to converge, we simplified its random effects structure first by removing random correlations and then by eliminating the components which accounted for the least variance. In the current case, the converged model did not include the correlation between random effects. Significance was assessed via model comparisons (Barr, Levy, Scheepers, & Tily, 2013).

This model yielded significant main effects of all three manipulated factors. There were more intrusions with dominant (English) as compared to nondominant (Spanish) targets (*M* = 4.76% vs. 1.29%; β = -1.31, *SE* β = .38, χ^2 (1) = 13.93, $p < .001$), more intrusions with function compared to content word targets (*M* = 5.65% vs. 0.45%; β = 2.31, *SE* β = .38, χ^2 (1) = 24.11, $p < .001$), and more intrusions with black compared to red targets (*M*= 4.84% vs. 1.23%; β = -1.50, SE β = .34, χ^2 (1) = 45.26, $p < .001$). None of the interactions approached significance (*p*s ≥.60). Finally, an application of the same model on a collapsed data set of both full and partial intrusions (with number of intrusions as dependent variable) yielded the same performance pattern. Namely, similar main effects of dominance, POS and color (*p*s< .001) were shown.

Discussion

The main finding of Experiment 1 was a significant effect of color cue – namely, Spanish-English bilinguals produced fewer intrusions on switch words marked by a color change. Additionally, the current experiment also replicated prior results; Spanish-English bilinguals exhibited reversed dominance effects i.e., produced more intrusions with dominant than with nondominant language switch words, and more intrusions with function than with content words (Gollan & Goldrick, 2016, 2018; Gollan, Schotter, et al., 2014; Gollan et al., 2017).

Returning to our main object of interest, the color cuing effects observed in Experiment 1 indicated that the language control system is sensitive to available novel cues to language switching, even when they are arbitrary to language use. The presentation of switch-out words in red significantly lowered intrusion error rates for Spanish-English bilinguals across all experimental manipulations – for dominant and non-dominant switches, for content and function words. Because Spanish and English do not differ obviously in orthography, the observation of significant color cueing effects might be limited to language pairs that are naturally impoverished in cues to language membership. In Experiment 2 we tested this possibility, by asking if color cues would also reduce intrusion errors in Hebrew-English bilinguals, for whom the written input is naturally rich in visual cues to language membership.

Experiment 2

Method

Participants. Forty eight Hebrew-English bilingual undergraduates at University of Haifa, Israel gave informed consent and were paid 40 NIS an hour for their participation. Table 1 shows self-reported participant characteristics and Multilingual Naming Test scores in both languages (MINT; Gollan et al. 2012). All participants obtained higher picture-naming scores in Hebrew, and were hence classified as Hebrew-dominant.

Materials and procedure. The 32 mixed-language paragraphs used in Experiment 1 were translated and adapted. Once again, each paragraph had 8 versions corresponding to the 8 experimental conditions. Each paragraph contained 8 controlled switches of independent words (i.e. free morphemes) following the same general principles as in Experiment 1, but also making sure no switches occurred at the beginning of a sentence, as capitalization is a property of English (that can therefore be used as a language cue). Note that some function words realized as free morphemes in English (and Spanish) are bound morphemes in Hebrew (e.g., *the*, *and*, *that*). In addition, whereas in English and Spanish possessive pronouns precede the noun they modify, in Hebrew they follow it (e.g., *his toy* translates to *ha-ca'acu'a šelo*; הצעצוע שלו ; the.toy of-him). As mentioned in our report of Experiment 1, switch-out words were always free morphemes inserted only at points where both languages exhibit word order congruency. As a result, Experiment 2 did not include switch-outs from the non-dominant language (English) to dominant language (Hebrew) in the aforementioned cases, which limited potential insertion points as compared to Experiment 1. We return to this point in the cross-experiment analyses below.

Three Hebrew-English bilinguals confirmed that controlled switches met these conditions and did not disrupt the narrative. Appendix B shows an example of how paragraphs were adapted so that the same paragraph was presented in all conditions (between subjects as part of counterbalancing).

Design and procedure was identical to the one reported for Experiment 1, with the exception of the presentation software and the computer used (E-prime 2.0 on a PC computer with a 19 inch color monitor).

Results

One English-Hebrew bilingual transcribed the recordings, coding full and partial language intrusion errors occurring on switch-out, switch-back and non-switch words. The results of this process were double checked by a second Hebrew-English bilingual. Table 3 reports the number of full and partial intrusions on switch-out words, switch back to the default-language words, and words that were non-switches. Figure 2 demonstrates the distribution of full intrusions on switch-out words by target language, color and part of speech.

Table 3

The number of full and partial intrusion errors by target language, part of speech and word color, Experiment 2.

		Full intrusions				Partial intrusions			
Target		Dominant		Nondominant		Dominant		Nondominant	
Type	Color	Function	Content	Function	Content	Function	Content	Function	Content
Switch Out	Black	29	16	6	0	7	20	0	0
	Red	11	6	4	2	5	16	0	4
Switch Back	Black	0	0	0	0	3	3	1	1
Non-switch	Black	0	0	0	0	1	1	0	2

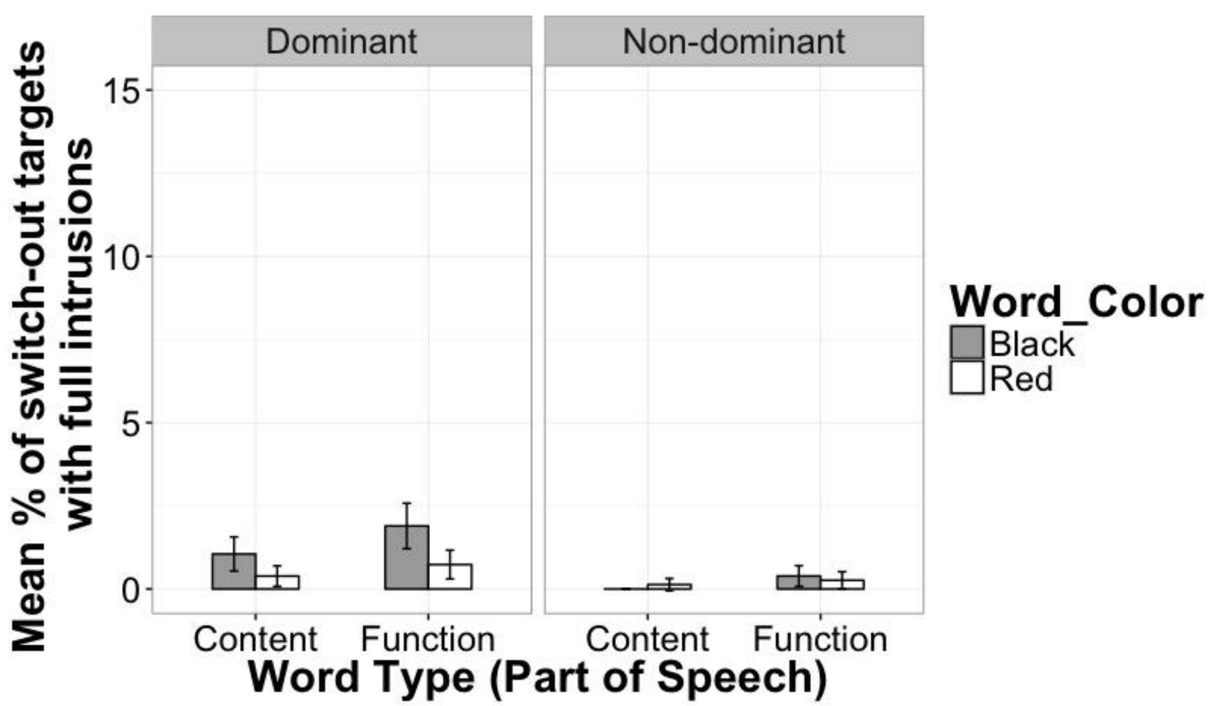

Fig. 2. Mean by participant of percentage of full intrusion errors on controlled switches by part of speech and word color in each target language (dominant vs. nondominant), Experiment 2. Error bars represent 95% confidence intervals.

As in Experiment 1, we first applied a binomial mixed effects model with full intrusions[1] as the dependent variable, switch-out language, part of speech and color as contrasted coded fixed effects, and subject and word as random intercepts with related random slopes. Random slopes were removed due to the failure to converge.

This model yielded a significant main effect of switch-out language such that there were more intrusions with dominant (Hebrew) as compared to the nondominant targets (English) (*M* = 1.02% vs. 0.20%; β = -5.18, *SE* β = 7.31, $\chi^2(1) = 8.03$, $p = .005$). Part of speech effects were in the same direction as in Experiment 1, more intrusions with function as compared to content words (*M* = 0.82% vs. 0.39%), but were only marginally significant in Experiment 2 (β = -4.55, *SE* β = 7.31, $\chi^2(1) = 3.43$, $p = 0.064$). Color effects were not significant (*M* = 0.83% vs. 0.38%; β = -3.34, *SE* β = 7.29, $\chi^2 < 1$). The three-way interaction was marginally significant (β = -15.50, *SE* β = 21.19, $\chi^2 = 2.92$, p= .087); the interaction between color and part of speech was not significant with dominant language targets (β = -.10, *SE* β = .61, $\chi^2 < 1$), but was marginally significant with nondominant targets (β = -15.17, *SE* β = .83.43, χ^2 =3.13, p = .076). All other interaction terms failed to reach significance (ps ≥ .13). The critical results remained similar when the same model was applied to a data-set collapsing full and partial intrusions. Namely, there was a main effect of switch-out language (p< .001) but no significant main effect of color (p = .198). The part of speech main effect was not significant (p = .159), whereas the three-way interaction effect was now significant (β = -17.05, SE β = 18.40, χ^2 = 5.87, p = .015).

The results of Experiment 2 are different in many ways from those reported in Experiment 1. However, in Experiment 2 there were very few intrusions on nondominant targets in general (only 12, 16.2% of all full intrusions), and even fewer on nondominant language content words (only 2, 2.7% of all full intrusions) and no intrusions at all on nondominant content words presented in black. Therefore, caution is needed when interpreting the results of a model applied on data from nondominant targets.

As such, we ran an additional analysis applied on dominant targets only with full intrusions as the dependent variable, part of speech and color as fixed effects and subject and word as random intercepts with related slopes. This model yielded a significant effect of color, such that there were fewer intrusions for targets highlighted in red than for targets written in black (M = 0.56% vs. 1.47%; β = 1.0, SE β = .46, $\chi^2(1)$ = 4.21, p = .042). However, the part of speech effect remained nonsignificant (M = 1.32% for function vs. 0.72% for content; β = -.29, SE β = .56, χ^2 <1), and the interaction between part of speech and color (β = .04, SE β = .92, χ^2 <1) was also not significant.

Discussion

Experiment 2 examined the effects of color cue, language dominance, and part of speech on production of intrusion errors in Hebrew-English bilinguals. Intrusion rates on non-dominant language targets were very low, and thus the most informative analysis focused on full intrusions with dominant language targets, and this analysis revealed a significant effect of color, i.e., bilinguals produced fewer intrusion errors on switch-out words that were marked by a color change. This shows that Hebrew-English bilinguals, for whom mixed-language paragraphs are rich in visual cues to language membership, can nevertheless also be sensitive to new arbitrary cues. This demonstrates the language control system to be highly sensitive to and adaptive to the presence of whatever language cues are available in the input.

Additionally, like in Experiment 1, and in previous studies (Gollan & Goldrick, 2016, 2018; Gollan, et al., 2017; Gollan, Schotter, et al., 2014), we observed reversed dominance effects – i.e., bilinguals produced more intrusions on dominant language targets embedded in paragraphs written primarily in the nondominant language. However, in contrast with Experiment 1, numerous previous studies with the read-aloud task (Kolers, 1966; Gollan & Goldrick, 2016; 2018; Gollan, Schotter, et al., 2014; Schotter et al., in press), and observations of naturally occurring intrusions (Poulisse & Bongaerts, 1994), part of speech was not a significant predictor of intrusion rate in Experiment 2. In addition, the number of intrusion errors observed in Experiment 2 was markedly lower overall than in Experiment 1 (only 74 full intrusions in Experiment 2, compared to 373 in Experiment 1). Before considering if the differences in results between Experiments 1 and 2 should be attributed to difference in orthography, we conducted a direct comparison to determine if there were any significant differences between experiments, and if so to possibly illuminate the source of any differences found.

Cross Experiment analyses

The first set of cross experiment analyses aimed to test whether there is a difference in the effect of the color cue on intrusion reduction between the two groups of bilinguals. To maximize power for this analysis, we started by considering all intrusions errors (i.e., both full and partial) and applied a binomial mixed effects model, with experiment (Experiment 1 vs. 2), color (red vs. black) and switch-out language (dominant vs. nondominant) and all interactions as contrasted coded fixed effects. The converging model included subject and word as random intercepts as well as their related random slopes. . This model yielded a significant main effect of experiment (more intrusions in Experiment 1 than 2, M = 3.84% vs. 1.03%; 472 vs. 126 errors; $\beta = 1.17$, SE $\beta = 0.23$, $\chi^2(1) = 25.19$, $p < .001$) and switch-out language (more intrusions with dominant language switch targets, M= 3.89% vs. 0.99%; 477 vs. 121 errors; $\beta = 1.95$, SE $\beta = .24$, $\chi^2(1) = 73.43$, $p < .001$), but failed to identify a significant main effect of color (M= 3.60% for black words vs. 1.28% for red words; 442 vs. 156 errors; $\beta = .65$, SE $\beta = .45$, $\chi^2(1) = 2.12$, $p = .145$). Nevertheless and importantly, the analysis yielded a significant interaction between color and experiment ($\beta = 1.37$, SE $\beta = .37$, $\chi^2(1) = 13.76$, $p < .001$), such that the color effect was larger in in Experiment 1, namely with the English-Spanish population. An additional significant result was a significant 3-way interaction between experiment, color, and switch-out language ($\beta = 1.49$, SE $\beta = .77$, $\chi^2(1) = 3.87$, $p = .049$), consistent with our finding that the color effect in Experiment 2 was mainly due to performance on dominant language targets, whereas it applied to both languages in Experiment 1 (see results of Experiments 1 & 2). The color-language and experiment-language interactions were also significant (ps < .04).

To reveal the source of the 3-way interaction, our next cross-experiment analysis was applied on a data set consisting only of intrusions on switches to the dominant language**.** We applied a binomial mixed effects model with part of speech (content vs. function), color (black vs. red), experiment and all interactions as contrasted coded fixed effects once on a data set collapsing full and partial intrusions, and once on full intrusions alone. The converging models included subject and word as random intercepts as well as their related random slopes. Both models yielded the same pattern: significant main effects of experiment (more intrusions in Experiment 1 than 2, *M* = 4.76% vs. 1.01%; 292 vs. 62 errors; β = 1.03, SE β = 0.34, $\chi^2(1)$ = 8.966, p = .003), part of speech (more intrusion with function words, *M*= 5.13% vs. 0.70%; 313 vs. 43 errors; β = 1.16, SE β = .34, $\chi^2(1)$ = 11.93, $p < .001$), and color (more intrusions with words in black vs. red *M*= 4.57% vs. 1.24%; 280 vs. 76 errors; β = 1.26, SE β = .24, $\chi^2(1)$ = 24.77, $p < .001$). There was also a significant interaction between experiment and part of speech (β = 2.99, SE β = .68, $\chi^2(1)$ = 19.35, $p < .001$). This interaction can be attributed to the absence of part of speech effect in Experiment 2. All other interactions were not significant (*p*s $\geq$ .23).

In sum then, in accordance with the separate analyses reported above, when considering both dominant and non-dominant switches, color significantly reduced intrusion errors for English-Spanish speakers but not for Hebrew-English speakers. However, when we focused our analysis on dominant language switches, we found that the color cue significantly reduced intrusion rates for both populations and to the same extent.

The second set of cross-experiment analyses was conducted to examine if the lower rate of intrusions in Experiment 2 relative to Experiment 1 might have been caused by factors other than the varying extent of the language intrinsic visual cues between the two experiments. The analyses reported below were applied on a data set that our previous analyses identified as revealing significant patterns in both experiments, namely full intrusions on dominant language targets.

We first asked if the observed experiment effect is due to the existence of more extensive limitations on possible switch points in Experiment 2. We suspected that the word order and morphological differences between Hebrew and English, as described above, may be responsible for the experiment effect, since they limited potential insertion of switches out to the dominant language in Experiment 2 as compared to Experiment 1. Specifically, the three words for which intrusions on switches out to the dominant language (English) were most frequent in Experiment 1 - 'the' (81), 'and' (64) and 'his' (33) - couldn't serve as switch-outs to the dominant language (i.e., Hebrew) in Experiment 2. As explained, the prior two are free morphemes in English (and Spanish) but bound ones in Hebrew, and the latter follows the noun it modifies in Hebrew (as opposed to English and Spanish) [2]. (For number of intrusions by word in both experiments see Appendix C). Accordingly, we asked if intrusions were still more frequent in Experiment 1 than 2, when analyzing only a subset of dominant-language switch-out targets that included translation equivalent words that appeared as switch targets in both experiments. This model still showed significant main effects of experiment (M = 6.51% vs. 1.31%; 93 vs. 23 errors; β = 2.48, SE β = .66, $\chi^2(1)$ = 14.51, p < .001) and color (M = 5.64% vs. 1.64%; 90 vs. 26 errors; β = 1.99, SE β = .69, $\chi^2(1)$ = 9.78, p= .001). The interaction between experiment and color was not significant (β = - .78, SE β = .89, χ^2<1).

Next, since the lengths of words in the dominant language in Experiment 1 (Hebrew, M = 3.79) were generally shorter than the lengths of words in the dominant language in Experiment 2 (English, M = 4.59), we ran an additional cross-experiment model with centered word length as

---

[2] This restriction was partially circumvented by replacing 'and' with the Hebrew equivalent of 'and-also' (*vegam* וגם). However, a simple comparison between the number of times 'and' appeared as a switch out to the dominant language in Experiment 1 (348) and the times וגם appeared as a switch-out in Experiment 2 (139) (across all sessions) demonstrates that possible insertion sites for the latter were limited as compared to and/y ('y' is the Spanish equivalent of 'and').

an additional predictor. Overall, there were more intrusions with shorter words (β = - .47, *SE* β = .13, $\chi^2(1) = 14.82$, $p < .001$). There was also a significant interaction between experiment and word length (β = .53, *SE* β = .27, $\chi^2(1)$ = , $p < .001$), as the main effect of length was significant in Experiment 2 (more intrusions with short targets; β = - .64, *SE* β = .16 , χ^2=23.54, $p < .001$), but was not in Experiment 1 (β = -.25, SE β = .25 , χ^2=1.01, p= .31). Like the analyses we report above, this model also yielded significant main effects of color (β = 1.28, SE β = .31, χ2 = 18.06, $p < .001$), POS (β = - .85, *SE* β = .37 , χ2 = 4.82, $p = .028$) and experiment (β = .93, *SE* β = .38, χ2 = 6.34, $p = .012$) and a significant POS and experiment interaction (β = 1.73, *SE* β = .75, $\chi^2(1) = 5.69$, $p = .017$). All other interactions failed to reach significance. Finally, we conducted a number of additional analyses to consider if cross-experiment differences could be attributed to participant characteristics shown in Table 1 including MINT scores and socioeconomic status (SES, estimated based on parent education level). These measures were chosen because bilinguals in Experiment 1 had lower nondominant language MINT scores, and lower SES than bilinguals in Experiment 2. First, in order to assess if cross experiment differences are due to differences in language proficiency (known to correlate with both the tendency and ability to code-switch, e.g., Toribio, 2001; see Bullock & Toribio, 2009 for a review), we ran a model focused on nondominant language proficiency by adding centered nondominant language MINT scores percentage[3] as a fixed effect to the cross experiment model. This model yielded the same pattern reported above. MINT scores were not significant predictors, and did not interact with any other

[3] This analysis examined percentages rather than raw scores since some participants in Experiment 1 completed every other item on the MINT instead of the full 68 items.

factors in the model (ps $\geq .33$). A similar model with relative MINT score, as an indicator of language balance, (non-dominant MINT raw score / dominant raw MINT score) yielded similar results, no main effect of language balance or interactions with any other factors (ps $\geq .17$). In a third model we added centered SES scores as a predictor to the cross-experiment model, and again found all the effects revealed by the cross-experiment model, while SES was not significant as a main effect and did not interact with any other factors (ps $\geq .29$).

Discussion

The cross experiment analysis found a significant effect of Experiment, indicating that Spanish-English bilinguals produced significantly more intrusion errors than Hebrew- English bilinguals. As indicated by the follow-up analyses, this difference cannot be attributed to demographic differences between the two populations, word length differences between Hebrew and English, or the limitations on switch point in Experiment 2.

When analyzing the effect of color, we found that it facilitated performance for both populations on switch-outs to the dominant language, in a similar manner. However, as reflected by the significant color by experiment interaction observed when both switches to the dominant and to the non-dominant language were analyzed, Spanish-English bilinguals also benefited from the color cue when required to switch to the non-dominant language, whereas Hebrew-English bilinguals showed similar performance in this condition for colored and non-colored words. Thus, the color effect for Hebrew-English bilinguals was modulated by dominance: the color cue was recruited only to avoid intrusions on switches to the dominant language (not on switches to the nondominant language). As mentioned above, this specific type of language control is known to be particularly challenging (Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Christoffels, Firk, & Schiller, 2007; Declerck, Thoma, Koch, & Philipp, 2015; Gollan &

Ferreira, 2009; Gollan & Goldrick, 2016, 2018; Gollan et al., 2017; Gollan, Kleinman, et al, 2014; Gollan, Schotter, et al., 2014).

To summarize, Spanish-English bilinguals, for whom mixed-paragraphs are less rich in language-intrinsic cues, benefitted from a language-external cue, color, to avoid language intrusions on both dominant and non-dominant targets, whereas Hebrew-English speakers, for whom mixed-language paragraphs are rich in cues intrinsic to language, benefitted from the external cue only when required to complete the more challenging task – avoid intrusions with target words in their dominant language. Given that Hebrew-English bilinguals were overall better at avoiding intrusions than Spanish-English bilinguals, a reasonable account of this contrast is that Hebrew-English bilinguals did not utilize the language external color cue in the less cognitively demanding cases, simply because the language intrinsic cue system was sufficient for achieving language control in these cases. This suggests that the sensitivity of the language control to novel cues might depend on the difficulty of the task, which could be partially determined by the availability of other language choice cues. We discuss this possibility in relation to the adaptive control model in the General discussion. Nevertheless, note that any analysis based on these data should be considered with caution, given the very low rate of intrusions. Namely, the lack of a color effect on non-dominant targets in Experiment 2 may reflect the very low number of intrusions on non-dominant targets in Experiment 2.

General discussion

This study investigated the sensitivity of two bilingual populations, Spanish-English (Experiment 1) and Hebrew-English (Experiment 2), to an arbitrary visual language switch cue in a mixed-language paragraph read aloud task. Results indicate that both populations were able to employ color change in written script to avoid language intrusion errors on switch-out words. This finding suggests that cue detection processes in language control are highly attuned to available language-choice cues, even when such cues are arbitrary to language use.

Moreover, recall that these two populations were selected to investigate whether the habitual use of visual cues more intrinsic to language use, such as alphabet and writing direction, might influence the adaptability of the language control system in detecting new ones. Written English and Spanish are quite similar, whereas written Hebrew and English exhibit very pronounced differences. Hence, we hypothesized that language control adaptability to novel cues in a given domain might be modulated by the availability of naturally occurring ones. In the Introduction section we noted that there were two possibilities to how this hypothesized connection might manifest. It was possible that due to more experience with prominent visual cues, Hebrew-English bilinguals would demonstrate higher sensitivity to the novel cue than Spanish-English bilinguals would. Alternatively, it was also possible that Spanish-English bilinguals would be more sensitive to the novel cue since other differences between English and Spanish are not as prominent.

The results demonstrate that both bilingual populations benefited equally from color cues in the most difficult experimental condition with the strongest language-control demands, namely when reading paragraphs written primarily in the nondominant language with switches into the dominant language. However, whereas Spanish-English bilinguals benefitted from the color cues under easier task demands as well, namely with non-dominant switch targets, Hebrew-English bilinguals did not. This pattern suggests that language control processes are opportunistic and highly adaptable, in the sense that any available cue in the input can be recruited to enhance performance, regardless of the habitual presence of such cues for a given bilingual. Further, the cue offered to bilinguals in the present experiment was novel, arbitrary and not in a domain that usually covaries with language, such as orthography (Casaponsa et al., 2014, 2015; Kestere, et al., 2012; Degani et al., 2017; Prior et al., 2017; Vaid & Frenck-Mestre, 2002), or personal cultural associations (Hartsuiker & Declerck, 2009; Li et al., 2013; Woumans et al., 2015; Zhang et al., 2013). Nonetheless, the color information was utilized by bilinguals of both populations.

Further, the current results show that most likely due to the language intrinsic visual cues available to the Hebrew-English bilinguals, they utilized the novel external cue to minimize intrusion errors only in the most demanding experimental condition. In the easier condition, apparently, the language intrinsic orthographic cues were sufficient to allow for error-free performance, and the color cues did not improve performance further. This pattern of results supports the flexibility of language control based on external circumstances as suggested by the adaptive control model (Green & Abutalebi, 2013). Our findings are hence consistent with the view proposed by this model, stating that control processes adapt to the environmental circumstances, namely the availability of cues and the demands placed on the system. Furthermore, they demonstrate that the assessment of these demands is based both on the availability of distinguishing cues in the environment and the difficulty of a given language control task.

Another important finding was that overall, Hebrew-English bilinguals had strikingly lower intrusion rates than Spanish-English bilinguals, again supporting the notion of flexibility and adaptability of the language control system. Independently of participant demographics, and differences in possible switch-out targets across language pairs, Hebrew-English bilinguals made significantly fewer intrusion errors than Spanish-English bilinguals. We attribute this difference to the fact that Hebrew-English mixed-language paragraphs already provide strong orthographic cues (including both visually distinct alphabets and a difference in reading direction), and together with color cues, this led to an overall lower rate of intrusions.[4] This, and other differential patterns of performance across bilingual groups support recent responses (Jiang, 2018; Mishra, 2018; van Heuven & Wen, 2018) to models of bilingual reading that are mostly based on the reality of same-script bilinguals such as Multilink(Dijksatra et al., 2018) and *BIA*/*BIA*+ (Dijkstra & Van Heuven, 1998, 2002; Van Heuven, Dijkstra & Grainger, 1998). These models view non-selectivity in lexical activation from print (i.e., activation of representations from multiple languages) as at least partially driven by cross-language orthographic overlap, leading to simultaneous activation of multiple languages As pointed out in the responses we cite above, these models were evaluated in relation to data from a same-script bilingual population, but a more general description of bilingual lexical access from print should also take into account different-script bilinguals. The current findings support the rationale behind this critique and potentially broaden it beyond the domain of word reading to apply more generally to bilingual language control. We therefore wish to stress the importance of investigating different-script bilinguals, and direct comparisons of same- and different-script bilinguals in the domain of language control as well, in order to achieve a fuller understanding of the inherent flexibility and adaptability of control mechanisms.

[4] It might seem differences between experiments could be attributed to differences in how often Spanish-English versus Hebrew-English bilinguals actually encounter mixed language text in natural settings. We cannot rule out such an explanation because we do not have systematic data that speak to this question. But note that because of the difference in writing direction, the greater syntactic and orthographic restrictions on mixing between Hebrew and English and the geographic distance between Israel and an English speaking country, this seems unlikely. If anything, a quick search of current Israeli newspaper articles and news websites reveals English loan words to be mostly transliterated (spelled in Hebrew letters), as opposed to switched to English letters.

Finally, the finding that part of speech was not a significant predictor of intrusion rate in Experiment 2, further supported by significant interaction between Experiment and part of speech, contrasts with the results of previous studies using the current paradigm (Gollan & Goldrick, 2016, 2018; Gollan, et al., 2017; Gollan, Schotter, et al., 2014) as well as observations in natural speech (Poulisse & Bongaerts, 1994) that intrusion rates are higher with function words. Considering the consistency of these findings and the general trend towards more intrusions on function words observed even in Experiment 2 (overall, 0.82% vs. 0.39%), it seems possible that this failure to reach significance is related to the overall small number of intrusion errors produced by our Hebrew-English bilinguals.

Conclusion

This work extends the findings of previous studies demonstrating the resourcefulness of the language control system in identifying and utilizing language-choice cues. Our results converge with findings of previous studies focused on cues regularly associated with language, such as orthography and facial features, and further generalize these findings to the mixed-paragraph reading aloud paradigm. The current study goes beyond previous research by also demonstrating that the language control system is able to take advantage of an arbitrary and novel cue, color change in written text. Moreover, it shows that this ability might be modulated by how regularly a given population uses other visual (orthography, reading direction) cues to language membership, as well as by the cognitive load posed by a given task. We interpret the results as a clear indication that cue detection processes in the language control system are extremely adaptive and flexible mechanisms, able to identify and utilize perceivable language-choice cues, even when their relation to language use is arbitrary.

Acknowledgments

We thank Aya Meltzer-Asscher for her notes during the development of the Hebrew stimuli, and Rosa Montoya, Mayra Murillo, Razan Silawi and Lynn Cohen for their assistance in data collection and coding.

Declarations of interest: none

Informed consent was obtained for experimentation with human subjects.

Funding: This work was supported by a National Institute on Deafness and Other Communication Disorders grant 011492, a National Science Foundation grant 1457159, Edmond J. Safra Center for the Study of Learning Disabilities (Faculty of Education, University of Haifa), and an Israeli Science Foundation grant 1094/14.

References

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

Bullock, B. E., & Toribio, A. J. (2009). Themes in the study of code-switching. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 1– 17). Cambridge: Cambridge University Press.

Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Frontiers in psychology*, 5, 424.

Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2015). How do bilinguals identify the language of the words they read?. *Brain research*, *1624*, 153-166.

Chan, B. H.-S. (2009) Code-switching between typologically distinct languages. In B. E. Bullock & A. J. Toribio (Eds.), *The Cambridge handbook of linguistic code-switching* (pp. 182–198). Cambridge: Cambridge University Press.

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

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

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283.

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

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

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

Declerck, M., Koch, I., Duñabeitia, J. A., Grainger, J., & Stephan, D. N. (in press). What absent switch costs and mixing costs during bilingual language comprehension can tell us about language control. *Journal of experimental psychology. Human perception and performance*.

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

Declerck, M., Thoma, A. M., Koch, I., & Philipp, A. M. (2015). Highly proficient bilinguals implement inhibition: Evidence from n-2 language repetition costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1911.

Degani, T., Prior, A., & Hajajra, W. (2018). Cross-language semantic influences in different script bilinguals. *Bilingualism: Language and Cognition*, *21*(4), 782-804.

Deuchar, M. (2005). Congruence and Welsh–English code-switching. *Bilingualism: Language and Cognition*, *8*(3), 255-269.

Dijkstra, A., & Van Heuven, W.J.B. (1998). The BIA-model and bilingual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 189–225). Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2018). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition,* 1-23.

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*(3), 640.

Gollan, T. H., & Goldrick, M. (2016). Grammatical constraints on language switching: Language control is not just executive control. *Journal of memory and language*, *90*, 177-199.

Gollan, T. H., & Goldrick, M. (2018). A switch is not a switch: Syntactically-driven bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(1), 143.

Gollan, T. H., Kleinman, D., & Wierenga, C. E. (2014). 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*, *143*(6), 2167.

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*(9), 1155-1164.

Gollan, T.H., Schotter, E.R., Gomez, J., Murillo, M., & Rayner, K. (2014). Multiple levels of bilingual control: Evidence from language intrusions in reading aloud. *Psychological Science*, *25*, 585-595.

Gollan, T. H., Stasenko, A., Li, C., & Salmon, D. P. (2017). Bilingual language intrusions and other speech errors in Alzheimer's disease. *Brain and cognition*, *118*, 27-44.

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

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

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515-530.

Hartsuiker, R. J. (2015). Visual cues for language selection in bilinguals. In F. Huettig (Ed.), *Attention and Vision in Language Processing* (pp. 129–145). India: Springer

Hartsuiker, R. J. & De Clerck, 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.

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

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script?. *Cognition*, *106*(1), 501-511.

Kolers, P. A. (1966). Reading and talking bilingually. *The American Journal of Psychology*, *79*(3), 357-376.

Kroll, J. F., Bobb, S. C., Misra, M., & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta psychologica*, *128*(3), 416-430.

Li, Y., Yang, J., Scherf, K. S., & Li, P. (2013). Two faces, two languages: An fMRI study of bilingual picture naming. *Brain and language*, *127*(3), 452-462.

Muysken, P. (2000). Bilingual speech: A typology of code-mixing. Cambridge: Cambridge University Press.

Poplack, S. (1980). Sometimes i'll start a sentence in spanish y termino en espanol: toward a typology of code-switching1. *Linguistics*, *18*(7-8), 581-618.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied linguistics*, *15*(1), 36-57.

Prior, A., Degani, T., Awawdy, S., Yassin, R., & Korem, N. (2017). Is susceptibility to cross-language interference domain specific?. *Cognition*, *165*, 10-25.

Schotter, E. R., Li, C., & Gollan, T. H. (in press). What reading aloud reveals about speaking: Regressive saccades implicate a failure to monitor, not inattention, in the prevalence of intrusion errors on function words. *Quarterly Journal of Experimental Psychology*.

Toribio, A. J. (2001). On the emergence of bilingual code-switching competence. *Bilingualism: language and cognition*, *4*(3), 203-231.

Vaid, J., & Frenck-Mestre, C. (2002). Do orthographic cues aid language recognition? A laterality study with French–English bilinguals. *Brain and Language, 82*(1), 47-53.Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic bulletin & review*, *9*(4), 780-789.

Van Heuven, W.J.B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*, 458–483.

Van Kesteren, R., Dijkstra, T., & De Smedt, K. (2012). Markedness effects in Norwegian–English bilinguals: Task-dependent use of language-specific letters and bigrams. *The quarterly journal of experimental psychology*, *65*(11), 2129-2154.

Woumans, E., Martin, C. D., Vanden Bulcke, C., Van Assche, E., Costa, A., Hartsuiker, R. J., & Duyck, W. (2015). Can faces prime a language?. *Psychological science*, *26*(9), 1343-1352.

Zhang, S., Morris, M. W., Cheng, C. Y., & Yap, A. J. (2013). Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference. *Proceedings of the National Academy of Sciences*, *110*(28), 11272-11277.

Appendix A. Set example, Experiment 1

Default language: English

Content-Black

Throughout the Land of the Pig River, the nombre Mrs. Peace was very well conocido by everyone. It wasn't so much because of the chisme that traveled from village to village, but due to the stories that circulated declaring her adventures and travesuras. There was something magnetic and encantador about her personality that attracted attention. In fact, there was always someone that had something chistoso to say about Mrs. Peace. The curious thing is that very few gente spoke negatively about her, in spite of her eccentric behavior. The verdad is that almost everyone admired her; even the youngest ones.

Content-Color[5]

Throughout the Land of the Pig River, the **nombre** Mrs. Peace was very well known by everyone. It wasn't so much because of the **chisme** that traveled from village to village, but due to the stories that circulated declaring her adventures and **travesuras**. There was something magnetic and **encantador** about her personality that attracted attention. In fact, there was always someone that had something **chistoso** to say about Mrs. Peace. The curious thing is that very few **gente** spoke negatively about her, in spite of her eccentric behavior. The **verdad** is that almost everyone admired her; even the youngest ones.

Function-Black

Throughout the Land of the Pig River, el name Mrs. Peace was very well known por everyone. It wasn't so much because of the gossip que traveled from village to village, but due to the stories that circulated declaring her adventures y mischief. There was something magnetic and charming de her personality that attracted attention. In fact, there was always someone that had algo funny to say about Mrs. Peace. The curious thing is that very poca people spoke negatively about her, in spite of her eccentric behavior. La truth is that almost everyone admired her; even the youngest ones.

Function-Color

[5] Words that appeared in red in our experiment appear here in boldface

Throughout the Land of the Pig River, **el** name Mrs. Peace was very well known **por** everyone. It wasn't so much because of the gossip **que** traveled from village to village, but due to the stories that circulated declaring her adventures **y** mischief. There was something magnetic and charming **de** her personality that attracted attention. In fact, there was always someone that had **algo** funny to say about Mrs. Peace. The curious thing is that very **poca** people spoke negatively about her, in spite of her eccentric behavior. **La** truth is that almost everyone admired her; even the youngest ones.

Default language: Spanish

Content-Black

Por toda la Tierra del Río Puerco, el name doña Paz era muy bien known por todos. No era tanto por el gossip que corría de pueblito a pueblito, sino las historias que circulaban declarando sus aventuras y mischief. Había algo magnético y charming de su personalidad que llamaba la atención. De hecho, siempre había alguien que tenía algo funny que contar de doña Paz. Lo curioso es que muy poca people hablaba mal de ella, a pesar de su comportamiento excéntrico. La truth es que casi todos la admiraban; hasta los más jóvenes.

Content-Color

Por toda la Tierra del Río Puerco, el **name** doña Paz era muy bien **known** por todos. No era tanto por el **gossip** que corría de pueblito a pueblito, sino las historias que circulaban declarando sus aventuras y **mischief**. Había algo magnético y **charming** de su personalidad que llamaba la atención. De hecho, siempre había alguien que tenía algo **funny** que contar de doña Paz. Lo curioso es que muy poca **people** hablaba mal de ella, a pesar de su comportamiento excéntrico. La **truth** es que casi todos la admiraban; hasta los más jóvenes.

Function-Black

Por toda la Tierra del Río Puerco, the nombre doña Paz era muy bien conocido by todos. No era tanto por el chisme that corría de pueblito a pueblito, sino las historias que circulaban declarando sus aventuras and travesuras. Había algo magnético y encantador about su personalidad que llamaba la atención. De hecho, siempre había alguien que tenía something chistoso que contar de doña Paz. Lo curioso es que muy few gente hablaba mal de ella, a pesar de su comportamiento excéntrico. The verdad es que casi todos la admiraban; hasta los más jóvenes.

Function-Color

Por toda la Tierra del Río Puerco, **the** nombre doña Paz era muy bien conocido **by** todos. No era tanto por el chisme **that** corría de pueblito a pueblito, sino las historias que circulaban declarando sus aventuras **and** travesuras. Había algo magnético y encantador **about** su personalidad que llamaba la atención. De hecho, siempre había alguien que tenía **something** chistoso que contar de doña Paz. Lo curioso es que muy **few** gente hablaba mal de ella, a pesar de su comportamiento excéntrico. **The** verdad es que casi todos la admiraban; hasta los más jóvenes.

Appendix B. Set example, Experiment 2

Default language: Hebrew

Content-Black

ברחבי הארץ, היה שמה של גברת שלום known לכולם. לא רק בגלל רכילות שפשטה בין הכפרים, אלא גם בשל הסיפורים על adventures ומעשי קונדס בהם היא הייתה involved. היה משהו appealing ומקסים באישיות שלה שמשך attention. למעשה, לכל אחד היה סיפור מצחיק לספר על גברת שלום. ובאופן surprising, על אף התנהגותה המוזרה, למעט מאד people היה משהו שלילי לומר עליה. לאמיתו של דבר, כמעט כולם admired אותה; אפילו הצעירים ביותר.

Content-Color

ברחבי הארץ, היה שמה של גברת שלום **known** לכולם. לא רק בגלל רכילות שפשטה בין הכפרים, אלא גם בשל הסיפורים על **adventures** ומעשי קונדס בהם היא הייתה **involved**. היה משהו **appealing** ומקסים באישיות שלה שמשך **attention**. למעשה, לכל אחד היה סיפור מצחיק לספר על גברת שלום. ובאופן **surprising**, על אף התנהגותה המוזרה, למעט מאד **people** היה משהו שלילי לומר עליה. לאמיתו של דבר, כמעט כולם **admired** אותה; אפילו הצעירים ביותר.

Function-Black

ברחבי הארץ, היה שמה של גברת שלום ידוע לכולם. לא just בגלל רכילות שפשטה בין הכפרים, אלא גם בשל הסיפורים of הרפתקאות ומעשי קונדס בהם היא הייתה מעורבת. היהsomething מושך and מקסים באישיות שלה שמשך תשומת לב. למעשה, לכל אחד היה סיפור מצחיק לספר about גברת שלום. ובאופן מפתיע, despite התנהגותה המוזרה, למעט מאד אנשים היה משהו שלילי לומר עליה. לאמיתו של דבר,almost כולם העריצו אותה; even הצעירים ביותר.

Function-Color

ברחבי הארץ, היה שמה של גברת שלום ידוע לכולם. לא **just** בגלל רכילות שפשטה בין הכפרים, אלא גם בשל הסיפורים of הרפתקאות ומעשי קונדס בהם היא הייתה מעורבת. היה **something** מושך **and** מקסים באישיות שלה שמשך תשומת לב. למעשה, לכל אחד היה סיפור מצחיק לספר **about** גברת שלום. ובאופן מפתיע, **despite** התנהגותה המוזרה, למעט מאד אנשים היה משהו שלילי לומר עליה. לאמיתו של דבר, **almost** כולם העריצו אותה; even הצעירים ביותר.

Default language: English

Content-Black

Throughout the Land of the Pig River, the name Mrs. Peace was very well known by everyone. It wasn't so much because of the שמועה that traveled from village to village, but due to the סיפורים that circulated declaring her הרפתקאות and mischief. There was something appealing and מקסים about her אישיות that attracted attention. In fact, there was always someone that had something

מצחיק to say about Mrs. Peace. The curious thing is that very few אנשים spoke negatively about her, in spite of her eccentric behavior. The אמת is that almost everyone admired her; even the youngest ones.

Content-Color

Throughout the Land of the Pig River, the name Mrs. Peace was very well known by everyone. It wasn't so much because of the **שמועה** that traveled from village to village, but due to the **סיפורים** that circulated declaring her **הרפתקאות** and mischief. There was something appealing and **מקסים** about her **אישיות** that attracted attention. In fact, there was always someone that had something **מצחיק** to say about Mrs. Peace. The curious thing is that very few **אנשים** spoke negatively about her, in spite of her eccentric behavior. The **אמת** is that almost everyone admired her; even the youngest ones.

Function-Black

Throughout the Land of the Pig River, the name Mrs. Peace היה very well known by everyone. It wasn't so much בגלל the gossip אשר traveled from village to village, but due to the stories that circulated declaring her adventures and mischief. There was something appealing וגם charming about her personality אשר attracted attention. In fact, there was always someone that had משהו funny to say about Mrs. Peace. The curious thing is that very few people spoke negatively about her, in spite of her eccentric behavior. The truth is that כמעט everyone admired her; אפילו the youngest ones.

Function-Color

Throughout the Land of the Pig River, the name Mrs. Peace **היה** very well known by everyone. It wasn't so much **בגלל** the gossip **אשר** traveled from village to village, but due to the stories that circulated declaring her adventures and mischief. There was something appealing **וגם** charming about her personality **אשר** attracted attention. In fact, there was always someone that had **משהו** funny to say about Mrs. Peace. The curious thing is that very few people spoke negatively about her, in spite of her eccentric behavior. The truth is that **כמעט** everyone admired her; **אפילו** the youngest ones.

Appendix C. Intrusions on switch-outs to dominant language by word, Experiments 1 and 2

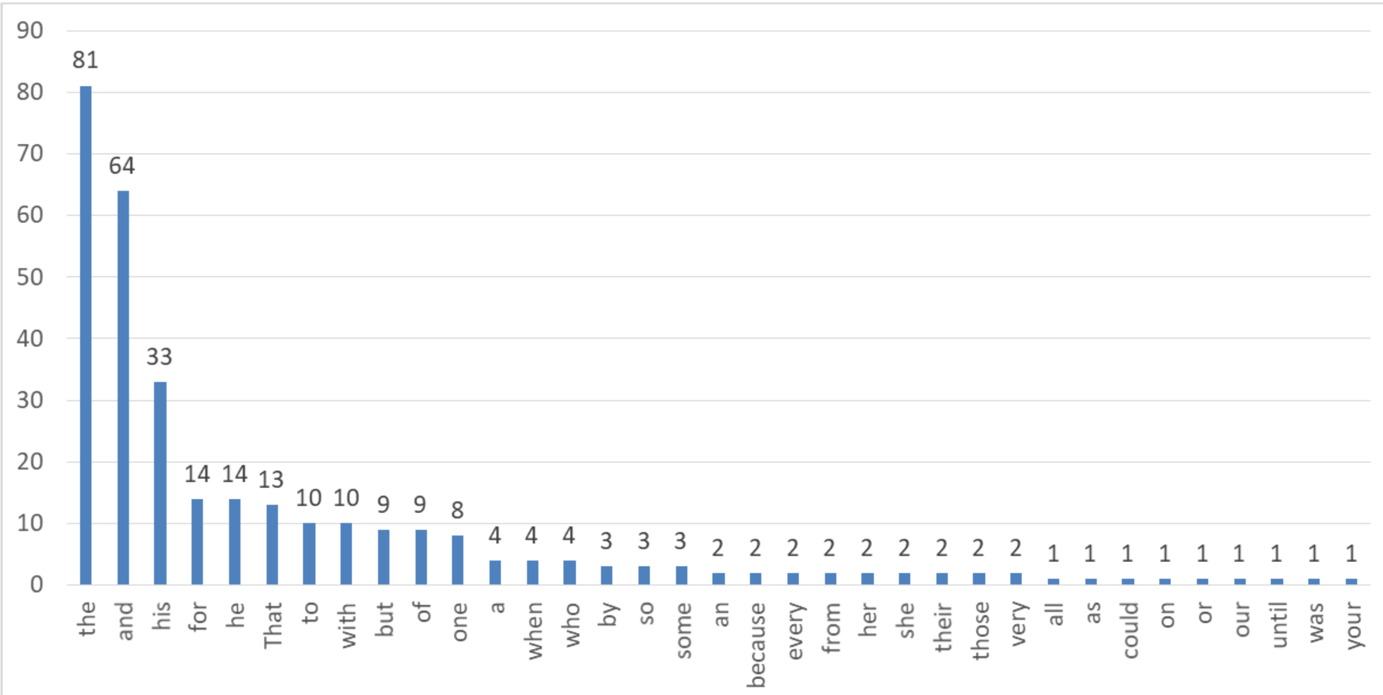

Fig. C. 1. Number of intrusions to dominant language by switch-out word, Experiment 1.

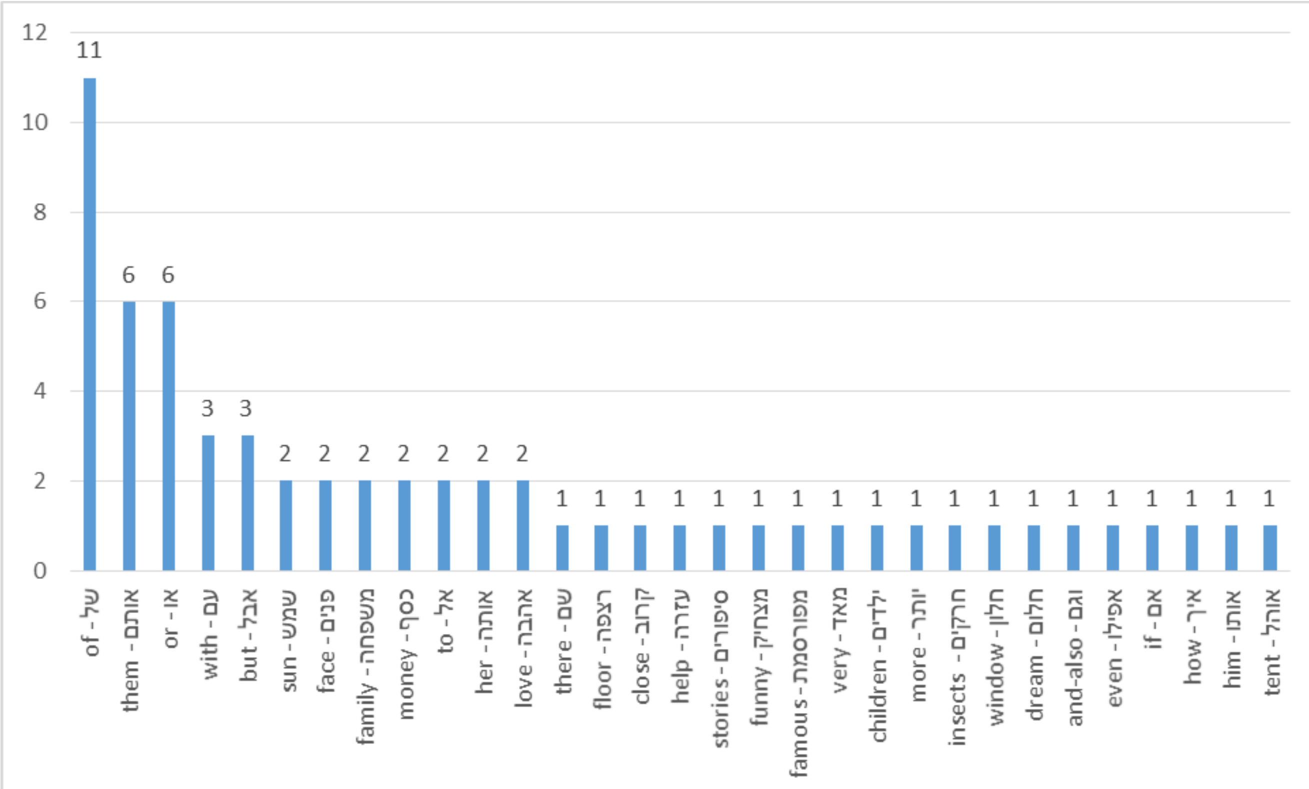

Fig. C. 2. Number of intrusion to the dominant language by switch-out word, Experiment 2.