ABSTRACT

Title of Thesis: THE ROLE OF COGNITIVE CONTROL IN BILINGUAL CODE-SWITCH COMPREHENSION

Lauren Kathleen Salig,
Master of Science, 2021

Thesis Directed By: Dr. Jared Novick, Department of Hearing and Speech Sciences
and
Dr. L. Robert Slevc, Department of Psychology

Bilinguals experience a conflict when comprehending a sentence that code-switches from one language to another. However, why bilinguals experience conflict during code-switch comprehension is unclear. This study asks: Does being in a cognitive state conducive to resolving conflict help bilinguals read code-switches faster? If so, it would indicate that comprehending a code-switch involves conflict at an early lexical/syntactic level because faster resolution of the conflict would facilitate faster code-switch reading. 101 Spanish-English bilinguals completed Flanker-arrow trials to manipulate their engagement of cognitive control—which regulates conflict detection and resolution. Immediately after this cognitive-control manipulation, bilinguals read code-switched or unilingual sentences. Having cognitive control engaged prior to encountering a code-switch did not result in faster reading of code-switches. This finding provides preliminary evidence that reading a code-switch may not involve conflict at a lexical/syntactic level. Future work should further investigate the type of conflict that bilinguals encounter during code-switch comprehension.

THE ROLE OF COGNITIVE CONTROL IN BILINGUAL CODE-SWITCH COMPREHENSION

by

Lauren Kathleen Salig

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2021

Advisory Committee:
Dr. Jared Novick, Chair
Dr. L. Robert Slevc, Co-Chair
Dr. Jorge Valdés Kroff, Special Member

Acknowledgements

Thank you to the co-investigators on this project: Jorge Valdés Kroff, Bob Slevc, and Jared Novick. Your support and guidance—both in this project and in navigating my first two years of graduate school—have been much needed and much appreciated. Thank you to the researchers who shared their materials with us or made them publicly available (Adler et al., 2020; Guzzardo Tamargo et al., 2016; Johns et al., 2019). Thank you to the many researchers who helped us with recruitment by sharing our study with their communities. Thank you to the people who participated in this study. This material is based upon work supported by the National Science Foundation under Grant No.1449815.

Table of Contents

Acknowledgements ii
Table of Contents iii
Introduction 1
  Summary of Relevant Cognitive Control Research 2
  What Type of Conflict do Bilinguals Encounter in Code-switch Comprehension? .. 5
  The Present Study 11
Method 15
  Participants 15
  Design 16
  Materials 18
  Procedure 20
  Data Analysis 21
Results 25
  Flanker Manipulation Check 25
  Effect of Prior Flanker on Sentence Reading 25
  Effect of Switch Type on Sentence Reading 38
  Potential Modulators of Conflict Adaptation 46
Discussion 49
  The Conflict in Code-switch Comprehension 49
  Different Switch Types 53
  General Implications 55
  Conclusion 56
Appendices 57
  Appendix A 57
Bibliography 78

The Role of Cognitive Control in Bilingual Code-Switch Comprehension

When bilinguals interact with each other, they sometimes “code-switch” from one language to another. Although this code-switching behavior is relatively common (Beatty-Martínez & Dussias, 2017; Guzzardo Tamargo et al., 2016; Poplack, 1980), it is not necessarily easy. Code-switching is a complex, rule-governed behavior that requires bilinguals to navigate both languages as they come in close contact (Poplack, 1980), particularly when the switch occurs within a sentence or utterance (intra-sentential code-switching; e.g., Dussias et al., 2014).

When a sentence begins in one language and switches to another, the code-switch may pose a challenge to bilinguals’ language comprehension. Language comprehension is an incremental process involving prediction about upcoming input (e.g., Altmann & Kamide, 1999). When input arrives that is incompatible with the initial interpretation of a sentence or when an unexpected word arrives, comprehenders experience processing difficulty. For instance, comprehending a temporarily ambiguous, or “garden-path,” sentence results in processing delays as compared with comprehending unambiguous sentences (e.g., Frazier & Rayner, 1982). A code-switch may be another such difficulty that slows comprehension processes. Bilinguals experience switch costs when comprehending code-switches, taking longer to process switched input than unilingual input in self-paced reading and eye-tracking studies (Altarriba et al., 1996; Bultena et al., 2015; cf. Johns et al., 2019). Further, EEG studies find a late positive component (LPC) when bilinguals comprehend a code-switch (Kaan et al., 2020; Moreno et al., 2002), similar to the P600 component found for sentence re-analysis (Gouvea et al., 2010). Despite this evidence for measurable switch costs during processing, bilinguals are able to understand code-switches with apparent ease. One way that bilinguals might overcome switch costs is by recruiting cognitive control to resolve the difficulty that arises during code-switch comprehension.

Cognitive control is the executive function that regulates the detection and resolution of conflict when processing information (Botvinick et al., 2001). Comprehending a code-switch may be cognitive control demanding (Adler et al., 2020), so in this study I ask: If bilinguals have cognitive control engaged prior to comprehending a code-switch, are switch costs reduced? That is, can switch costs be reduced when bilinguals are already in a state conducive to resolving conflict? Addressing this question will also provide insight into exactly what type of difficulty or conflict arises when a bilingual comprehends a code-switch. Note that this study does not ask whether bilinguals have "good" or "bad" cognitive control overall, but rather how a bilingual's cognitive state affects their language processing in the moment (Salig et al., 2021).

Below I review the existing literature on the role of cognitive control in code-switch comprehension and further discuss why code-switch comprehension might involve conflict resolution.

Summary of Relevant Cognitive Control Research

Cognitive control regulates behavior to overcome conflict by promoting task-relevant information over task-irrelevant information. For example, consider a garden-path sentence such as "The old man the boat," in which the comprehender needs to inhibit their initial interpretation of "man" as a noun in favor of the correct interpretation of "man" as a verb. Cognitive control promotes the correct but unexpected interpretation over the initially predominant but incompatible interpretation (Hsu & Novick, 2016; Hsu et al., 2021). In a similar way, based on their past language experiences, bilinguals may expect the next word of a sentence to continue in one language and may need to reinterpret some aspect of the sentence when a code-switch occurs (Gosselin & Sabourin, 2021; Guzzardo Tamargo et al., 2016; Moreno et al., 2002). As with garden-path sentences, the conflict between the expected and arriving input during a code-switch may require bilinguals to engage cognitive control to successfully resolve that conflict and comprehend the code-switch (Adler et al., 2020). Given that cognitive control engages flexibly based on processing demands (Botvinick et al., 2001), it is possible to ask if/how bilinguals process code-switches differently under various states of cognitive control engagement.

Once cognitive control is engaged, it makes an immediately subsequent task easier if it also includes conflicting information, a phenomenon called conflict adaptation (Gratton et al., 1992). An established method for manipulating the need for cognitive control is the Flanker task. Participants indicate the direction of a central arrow flanked by congruent (→→→→→) or incongruent (←←→←←) arrows (Eriksen & Eriksen, 1974). Incongruent trials introduce information conflict that engages cognitive control; through conflict adaptation, an incongruent trial is easier when preceded by another incongruent trial. The conflict adaptation pattern suggests that people increase engagement of cognitive control when they encounter an incongruent trial, and cognitive control remains engaged for some time during the processing of a subsequent incongruent trial, helping people respond faster on that trial (e.g., by promoting representations of the central arrow over representations of the distracting flanking arrows).

The conflict adaptation effect is not limited to carryover within the same task: Studies have demonstrated that conflict adaptation can be observed across tasks in different domains (e.g., Kan et al., 2013). For instance, engaging cognitive control on one trial with an incongruent Flanker-like task (i.e., the Stroop task) can help monolinguals reinterpret a garden-path sentence on the next trial (Hsu & Novick 2016; see also Thothathiri et al., 2018). These findings suggest that cognitive control is a domain-general mechanism that operates across linguistic and non-linguistic representations.

Using a cross-task conflict adaptation paradigm, Adler and colleagues (2020) found evidence that bilinguals engage cognitive control when they encounter a code-switch. Bilinguals read code-switched or unilingual sentences at their own pace and then completed a subsequent Flanker trial. After reading a code-switched sentence (compared with a unilingual sentence), bilinguals responded faster to an incongruent Flanker trial. This pattern is consistent with a conflict adaptation effect and suggests that interpreting a code-switch engages domain-general cognitive control and, thus, involves conflict at some level. This interpretation is supported by a study by Wu and Thierry (2013) which found that Welsh-English bilinguals resolved conflict faster on incongruent Flanker trials when they were interleaved with words from both languages rather than only one. Furthermore, although a replication of the Adler et al. study with event related potentials (ERPs) failed to replicate the same behavioral results, the researchers found that prior sentence type (code-switched vs. unilingual) modulated the

P300 component—which was taken to be an index of cognitive interference—during the subsequent Flanker trials (Bosma & Pablos, 2020). Overall, these studies suggest that language context (e.g., code-switched or mixed language contexts vs. unilingual contexts) may alter bilingual processing demands and, in turn, regulate the engagement of cognitive control to meet those demands.

These studies provide evidence that reading a code-switch requires bilinguals to resolve conflict. However, it is unclear if the switch cost induced by a code-switch can be reduced when bilinguals have cognitive control engaged when they encounter the code-switch. The answer to this question likely depends on *why* comprehending a code-switch requires bilinguals to resolve conflict.

What Type of Conflict do Bilinguals Encounter in Code-switch Comprehension?

Conflict at the Lexical and/or Syntactic Level

One possibility is that when bilinguals comprehend a code-switch, they confront conflict between their languages at a lexical level as they attempt to integrate the code-switched content into their existing understanding of the sentence. For example, Altarriba and colleagues (1996) asked bilinguals to read English sentences embedded with a target word that either continued in English or code-switched into Spanish. Bilinguals read high-frequency code-switched words slower in highly constrained sentence contexts than in lower constrained sentence contexts. Altarriba and colleagues proposed that sentence constraints led bilinguals to make language-specific lexical expectations about upcoming content which were violated by a code-switch. That is, bilinguals may experience a switch cost because of a conflict between an expected lexical item in one language and the lexical item that arrives in another language. Under this account, non-habitual code-switchers may experience more conflict than habitual code-switchers by virtue of switches being more unexpected for them. This interpretation is supported by a recent finding that non-habitual code-switchers demonstrated an N400—indexing lexicosemantic integration difficulty—in response to code-switches, while habitual code-switchers did not (although they still demonstrated an LPC in response to code-switches; Gosselin & Sabourin, 2021).

Even in situations where bilinguals are free to switch languages at will, code-switched utterances will be in the minority, occurring between blocks of unilingual utterances (Beatty-Martínez et al., 2020; Fricke & Kootstra, 2016; Piccinini & Arvaniti, 2015). The relative infrequency of code-switches, as compared with unilingual content, could make code-switches unexpected and harder to integrate in comprehension because of conflicting cross-linguistic representations. This explanation would predict that habitual code-switchers may experience more conflict when they comprehend less frequent types of code-switches because the switch is more unexpected.

Similarly (and not mutually exclusively), bilinguals may experience a switch cost because of a conflict at the syntactic level. Code-switching is not simply the existence of two languages next to each other; it can involve grammatical structures unique to code-switching itself. For example, in Spanish a masculine noun nearly always follows a masculine determiner (e.g., "el lago"), but it is acceptable and common in code-switching for an English noun with a feminine Spanish translation to follow a masculine Spanish determiner (e.g., "el milk" instead of "la leche;" Beatty-Martínez & Dussias, 2017; Valdés Kroff, 2016). In those cases, there could be a conflict between the expected syntactic category (e.g., masculine noun in Spanish) and the code-switched syntax that arrives (e.g., English noun with a feminine Spanish translation). Even when the grammar of one language is not directly violated by the code-switch, there may be some conflict from combining two different grammars in one sentence.

If the conflict elicited by a code-switch occurs at a lexical or syntactic level, then engaging cognitive control before reading a code-switch should result in faster reading times by facilitating integration of the code-switch into the reader's linguistic representation. Under this account, the role of cognitive control in code-switch comprehension is a close parallel to its role in comprehension of temporarily ambiguous garden-path sentences. Just as garden-path sentences may require reinterpretation of a sentence's syntax when incoming input conflicts with the comprehender's initial interpretation, code-switched sentences may require the comprehender to resolve conflict at a similar lexical/syntactic level as the code-switched input deviates from their expectations. Past research has established that reading a garden-path sentence engages cognitive control, resulting in faster responses on a subsequent cognitive control task (Kan et al., 2013), and also that engaging cognitive control improves subsequent garden-path sentence processing (Hsu & Novick, 2016). That is, conflict adaptation occurs in both directions (e.g., Stroop to sentence, and sentence to Stroop) for garden-path sentence comprehension where the conflict is at a lexical/syntactic level. If code-switch comprehension involves conflict at a similar level of representation, we would expect to observe conflict adaptation in both the sentence to Flanker direction (Adler et al., 2020) and in the Flanker to sentence direction that this study tests.

Conflict at the Pragmatic Level

The above lexical/syntactic account posits that bilinguals experience conflict during code-switch comprehension immediately upon encountering the switch and attempting to integrate the code-switched input into their interpretation of the sentence. This type of lexical or syntactic conflict would occur early in processing, as soon as bilinguals try to access the code-switched word or insert it into the sentence's structure.

Alternatively, bilinguals may experience late conflict during code-switch comprehension for pragmatic reasons that do not involve early conflict at the lexical or syntactic level. Bilinguals may encounter a conflict not from trying to integrate the code-switched content, but rather from trying to understand the reason for the switch (e.g., the experimental purpose, to signal identity, or for the speaker's ease of production). That is, even if bilinguals seamlessly integrate a code-switch into their interpretation of a sentence, they could experience a conflict later in processing as they assess the pragmatics. A recent ERP study found that when bilinguals read a code-switch out loud to a confederate, the LPC elicited by the switch was reduced if the participant believed the confederate to also be a bilingual. That is, the LPC response to a code-switch was lessened when the switch occurred in a pragmatically supported context (Kaan et al., 2020). If the LPC is taken as an index of conflict—perhaps through reinterpretation of the sentence (Litcofsky & Van Hell, 2017)—then this finding suggests that at least part of the conflict in code-switch comprehension is at the pragmatic level.

Even if bilinguals easily integrate the code-switched input into their linguistic understanding early on, they may still need to engage cognitive control later to resolve conflict arising from pragmatic considerations. The social context of a particular conversation may not pragmatically support code-switching behavior (e.g., Kaan et al., 2020). The physical environment, especially in a university laboratory, may be an unexpected context for code-switching to occur. Or, the sentential context of a code-switch may bring about pragmatic conflict. For instance, bilinguals can use code-switches to predict upcoming low-frequency referents (Tomić & Valdés Kroff, 2021). Under this predictive process, one possibility is that hearing a code-switch may result in a pragmatic conflict between how a bilingual initially expected the sentence to continue (i.e., with a predictable, high-frequency word) and the code-switch's indication that it may continue in a different way (i.e., with a low-frequency word or unexpected topic). In this case, highly constrained sentences that include a code-switch may result in more pragmatic conflict than sentences that do not lead to strong predictions. Another possibility is that a pragmatic conflict occurs when a code-switch is followed by a high-frequency or expected word instead of a low-frequency word, as the bilingual experiences conflict between what a code-switch typically indicates about upcoming content and what arrives. In this case, code-switched sentences that continue in a highly expected way may elicit more pragmatic conflict than code-switched sentences that are followed by low-frequency or surprising words in a naturalistic manner. These are just two possible sources of pragmatic conflict that could emerge from sentential context and bilinguals' associated expectations. The general prediction is that when a code-switch violates bilinguals' expectations about a sentence, they may experience greater pragmatic conflict as they infer the reason for their inaccurate expectation. Based on the above account of pragmatic conflict, I would also expect that pragmatic conflict would be greater in social or physical contexts where code-switches are unexpected than where code-switching is common.

If the conflict occurs at a pragmatic level, then engaging cognitive control beforehand may not impact reading of code-switches. Under the pragmatic conflict account, the role of cognitive control in code-switch comprehension may be considered a parallel to its role in comprehension of irony. Past work has found that interpreting an ironic sentence engages cognitive control and facilitates performance on a subsequent incongruent Stroop trial, indicating that irony comprehension involves conflict at some level. However, engaging cognitive control prior to irony comprehension did not facilitate irony comprehension (Adler et al., 2018). That is, conflict adaptation is only observed in one direction for irony comprehension: sentence to Stroop, but not Stroop to sentence. Adler and colleagues explain this effect as "late irony:" Literal interpretation at the early lexical/syntactic level occurs during initial sentence processing without conflict, and only after this initial processing does conflict arise later at the pragmatic level as the listener attempts to reconcile the literal interpretation with the ironic context.

The conflict involved in code-switch comprehension may occur at a similar pragmatic level; for instance, after bilinguals complete semantic processing of the switch, they may experience pragmatic conflict as they attempt to reconcile the presence of a switch with the reason for why a switch occurred in that particular context or location. If code-switch comprehension involves conflict at a pragmatic level, we would still expect to observe conflict adaptation in the sentence to Flanker direction (Adler et al., 2020); bilinguals would engage cognitive control upon reaching the late pragmatic conflict in code-switch comprehension, and it would remain engaged on a following Flanker trial. However, conflict adaptation should not be expected in the other direction. If the conflict does not occur at the initial sentence processing level, cognitive control engagement should not affect sentence reading since sentence interpretation would not be affected by conflict.

It is also worth noting that if the conflict in code-switch comprehension originates from pragmatic considerations, the cost of interpreting the reason for a switch could be offset by the downstream benefit of, for example, correctly predicting an upcoming low-frequency referent (Tomić & Valdés Kroff, 2021)—offering one explanation for why bilinguals code-switch frequently despite observed switch costs.

The Present Study

Past work suggests that reading a code-switch is cognitive control demanding, requiring bilinguals to resolve a conflict that arises during comprehension (Adler et al., 2020). However, as outlined above, it remains unclear *why* bilinguals experience a conflict during code-switch comprehension—the level of representation (e.g., lexical, syntactic, and/or pragmatic) at which the conflict emerges is not yet clear.

As a next step in advancing our understanding of code-switch comprehension, this study asks whether bilinguals read code-switches faster when they have cognitive control engaged. When bilinguals are in a cognitive state conducive to resolving conflict, are switch costs reduced? If yes and conflict adaptation is observed from cognitive control task to code-switch reading, then this would provide evidence that bilinguals experience conflict at the lexical/syntactic level during code-switch comprehension. If no and conflict adaptation is not observed from cognitive control task to code-switch reading, then this would serve as an initial indication that bilinguals may not experience a lexical/syntactic conflict during code-switch comprehension, opening the door to investigating other possibilities about when and at what level bilinguals experience conflict (including further consideration of conflict at the pragmatic level).

To test this question, Spanish-English bilinguals completed Flanker trials, which modulated cognitive control engagement, followed immediately by self-paced reading of sentences that were either code-switched or unilingual.

Primary Prediction: Cognitive Control Engagement as a Modulator of Code-switch Reading

I predicted that upregulating cognitive control with an incongruent Flanker trial would result in bilinguals reading code-switched content faster; I expected to observe conflict adaptation. Such a finding would suggest that reading a code-switch engages cognitive control to resolve conflict that arises at the lexical/syntactic level. A null result would serve as an initial indication that reading a code-switch does not involve a lexical/syntactic conflict, warranting further consideration of the possibility that comprehending a code-switch involves conflict at a pragmatic level. My primary predication was:

(1) Upregulating cognitive control with an incongruent (vs. congruent) Flanker trial will assist processing of an immediately subsequent code-switch.

Secondary Predictions

Reading of Different Code-switch Types. As my primary purpose was to determine if switch costs could be reduced, I also wanted to consider additional factors beyond cognitive control engagement that might alter the magnitude of switch costs in the first place. Some evidence suggests that more frequent types of code-switches in production are processed faster in comprehension (Beatty-Martínez & Dussias, 2017; Guzzardo Tamargo et al., 2016). Expecting to replicate past findings, I made the following hypotheses about which types of code-switches would be processed faster by bilinguals:

(2a) Masculine determiner-noun switches will be easier to process (have lower switch costs) than feminine determiner-noun switches.

(2b) Perfect tense switches at the auxiliary location (e.g., "ellos have cleaned") will be easier to process than perfect tense switches at the participle location (e.g., "ellos han cleaned").

(2c) Progressive tense switches will be easier to process than perfect tense switches at both the participle and the auxiliary location.

(2d) Mixed noun switches (switch at the noun) will be easier to process than mixed verb switches (switch at the participle).

Different Code-switch Types as a Modulator of Conflict Adaptation. Given the expectation that certain code-switches would be more difficult to process than others, I also expected that the effect of our cognitive control manipulation would vary based on code-switch type. I made the exploratory prediction that:

(3) Harder code-switches will show a greater processing benefit from upregulation of cognitive control than easier code-switches.

Individual Language History as a Modulator of Conflict Adaptation. Much of the recent work on bilingualism highlights the idea that bilingualism is not a singular category but rather a diverse spectrum. We should expect that different bilingual experiences will lead to different cognitive outcomes (e.g., DeLuca et al., 2020; Green & Abutalebi, 2013). For example, cognitive control engagement may modulate code-switch reading differently for bilinguals with extensive code-switching experience versus those with little code-switching experience. As I did not have specific predictions about how exactly individual language background differences would affect our manipulations, I made the exploratory prediction that:

> (4) Bilinguals' individual language history will affect the processing benefit afforded to them from cognitive control engagement.

Method

Participants

Prior to data collection, I used the R package simr (Green & MacLeod, 2016) to simulate a version of the Adler et al. dataset with 500 participants and ran a series of power curve analyses using z-scores to determine how many participants would be required to have at least 80% power to detect an interaction effect of the same size (-0.0306) as Adler et al., 2020. A power analysis with 1,000 simulations indicated that 100 participants would give us 83% power to detect an effect the size of that found by Adler and colleagues.

152 participants who self-identified as Spanish-English bilinguals participated in the study online using their own devices. Participants were recruited through online measures (email listservs and social media posts) and through a university undergraduate research participation system. Participants were offered class credit or entry into a gift card drawing as compensation. 45 participants were excluded from data analysis based on the following pre-registered exclusion criteria: comprehension question accuracy under 80%, Flanker accuracy under 50%, Spanish grammar assessment score under 30%, and/or English grammar assessment score under 30%. Cut-off values for Flanker accuracy and grammar assessment scores were set based on requiring participants to perform approximately at or above chance levels. An additional 6 participants were excluded due to an error that resulted in their responses not being recorded, leaving 101 participants who were included in data analysis. Of the 101 participants, 86 participants identified as female, and 15 identified as male.

59.41% of participants learned Spanish before English, but participants reported being exposed to English more frequently in their daily lives (means of 65.74% English exposure, 31.50% Spanish exposure, 2.75% exposure to other languages) and reported that when given the choice, they would choose to speak English more often (means of 62.50% English, 33.91% Spanish, and 3.59% another language). Thirty participants were born outside of the United States and had lived in the United States for an average of 9.73 years (SD = 7.37 years). Additional participant characteristics are shown in Table 1. Overall, questionnaire responses suggest that the sample consists of relatively balanced bilinguals who may be slightly English dominant.

Table 1

Participants' Language History

	Mean
Age (range in parentheses)	22.8 (18-55)
AoA English (range in parentheses)	3.10 (0-17)
AoA Spanish (range in parentheses)	2.30 (0-19)
MELICET: English grammar score (out of 20)	16.00 (2.76)
DELE: Spanish grammar score (out of 20)	13.72 (3.53)
Self-Rated English Ability (max = 10)	9.59 (0.69)
Self-Rated Spanish Ability (max = 10)	8.45 (1.31)
Code-switching Experience (max = 5)	2.97 (0.83)

Note. Standard deviation in parentheses unless otherwise noted.
AoA = Age of Acquisition

Design

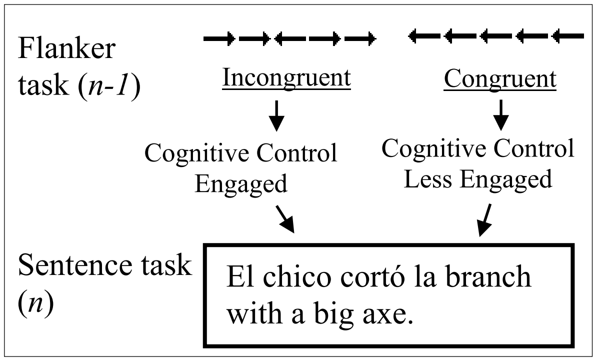
The study's design was pre-registered to OSF prior to the start of data collection (see pre-registration at https://osf.io/8a2z7). I used a 2x2 design that interleaved Flanker trial type (incongruent or congruent) with sentence trial type (code-switched or monolingual): Flanker trials preceded sentence trials (see Figure 1). Participants completed 96 of these Flanker-sentence trial pairs, seeing 24 trial pairs in each of the four Flanker type x sentence type conditions. After each such trial pair, participants answered a comprehension question to gauge attention and offline comprehension.

The design was within-subjects. All participants completed the same fixed order list of materials and were exposed to all four conditions. The 96 trial pairs were split into two experimental blocks and their order randomized within each block. Participants were assigned to one of eight lists which determined which of the four critical conditions they saw for each of the 96 trial pairs. Of the eight stimuli lists, four presented monolingual sentences and comprehension questions in English (base language of English), and four presented monolingual sentences and comprehension questions in Spanish (base language of Spanish). All lists presented code-switched sentences that switched from Spanish into English.

To disguise the manipulation, 92 filler trial pairs were dispersed among the 96 critical trial pairs. While critical trial pairs were always a Flanker-sentence sequence, filler trial pairs could be Flanker-Flanker, sentence-sentence, sentence-Flanker, or Flanker-sentence sequences to ensure that participants could not reliably predict the next task. 36 of the 92 filler trial pairs (and all 96 critical trial pairs) were followed by comprehension questions to ensure participants were on-task. Of the filler sentences, eight were code-switched, and the rest were monolingual to ensure that only about 33% of the sentence stimuli were code-switched to approximate natural proportions of code-switching (Beatty-Martínez et al., 2020; Piccinini & Arvaniti, 2015).

Figure 1

Study Design



Materials

Participants completed the experiment on pcIBEX (Zehr & Schwarz, 2018) followed by the language history questionnaire and grammar assessments on Qualtrics. Within the experiment, each item (Flanker, sentence, or comprehension question) was preceded by a 500ms fixation cross, and there was an inter-stimulus interval of 100ms between each item.

Self-Paced Reading Sentence Task

The language of monolingual sentences was determined by the list base language, but all code-switches were from Spanish to English to align with typical code-switching practices in the United States (Beatty-Martínez & Dussias, 2017; Blokzijl et al., 2017). Of the 96 critical sentences, 48 involved a determiner-noun code-switch or no-switch equivalent and 48 involved a verb code-switch or no-switch equivalent.

Determiner-noun stimuli were a selection of stimuli from Adler et al., 2020 and from Johns et al., 2019. Of the 48 determiner-noun sentences, half involved a noun that is feminine in Spanish and half a noun that is masculine. Verb stimuli were a selection of stimuli from Adler et al., 2020 and from Guzzardo Tamargo et al., 2016. Of the 48 verb sentences, 24 involved perfect tense verbs (half switched at the auxiliary and half switched at the participle) and 24 involved progressive tense verbs (half switched at the auxiliary and half switched at the participle; see Table 2). Filler sentences were taken from Adler et al., 2020 stimuli.

The sentence task was created using the pcIBEX DashedSentence Controller using the default moving window display so that participants saw one word at a time with the rest of the words masked with dashes. After the 500-ms fixation cross, participants advanced through each word of the sentence at their own pace by pressing the space bar.

Table 2

Types of Code-switches in Critical Sentences

	96 Critical Sentences					
	48 Determiner-Noun		48 Verb			
	24 Masculine	24 Feminine	24 Perfect Tense		24 Progressive Tense	
			12 switched at participle	12 switched at auxiliary	12 switched at participle	12 switched at auxiliary
Ex.:	*el lake*	*la milk*	*los modelos han signed*	*los editores have approved*	*los banqueros están preparing*	*los estudiantes are checking*

Flanker Task

The Flanker stimuli were the same images as used in Adler et al., 2020—a central arrow in the middle of the screen flanked by two arrows on each side pointing either the same direction or the opposite direction. Participants were instructed to press the "J" key if the central arrow was pointing right and the "F" key if the central arrow was pointing left. Although there was an equal number of congruent and incongruent Flankers in the critical Flanker-sentence pairs, throughout the entire experiment about 64% of all Flanker items were congruent.

Comprehension Question Task

Yes/no comprehension questions were adapted from comprehension questions used in the studies the sentence stimuli were drawn from or were generated by me. Participants were instructed to press the “J” key to answer yes and the “F” key to answer no. The words “Yes” and “No” were displayed on the bottom right and left sides of the screen respectively during each comprehension question as a reminder of the key response instructions.

Procedure

Participation was completely online and took approximately one hour. Upon opening the pcIBEX study link, participants were reminded that the study was only for Spanish-English bilinguals who were at least 18 years old and were instructed to exit out of the study if they did not fit the requirements. Participants were also instructed to complete the study on a laptop or desktop computer, not on a mobile device. Participants then read the consent form and indicated their consent to participate.

At the beginning of the study, participants were instructed on how to respond to sentence, Flanker, and comprehension question tasks and shown examples of each. Participants were instructed to keep their dominant thumb on the space bar and their index fingers on the “F” and “J” keys throughout the experiment. Instructions for the experiment included code-switches.

Participants completed a 15-item practice block after which they were reminded of the response keys for each task. They then completed the experimental items in two blocks. After the first block, participants were instructed to take a 2-3-minute break. After completing the second block, participants were directed to Qualtrics to complete the questionnaires.

On Qualtrics, participants first completed a demographic and language history questionnaire followed by two grammar assessments that were presented in random order. The grammar assessments were shortened 20-question versions of those used in Adler et al., 2020. The Spanish grammar assessment was adapted from the highest level of the Diploma de español como lengua extranjera [Diploma of Spanish as a Foreign Language] (Ministry of Education, Culture, and Sport of Spain, 2006). The English grammar assessment was adapted from the Michigan English Language Institute College English Test (English Language Institute, 2001).

Data Analysis

Data Cleaning

The regression models used to analyze the data were pre-registered prior to data collection; models that were pre-registered as exploratory or that were not pre-registered will be noted as exploratory in this paper. After applying participant exclusion criteria, I also applied individual trial exclusion criteria before running our models. As pre-registered, trial pairs were excluded from analysis if: the Flanker response time was more than 2.5 standard deviations away from that participant's mean, the reading time in the region of interest was more than 2.5 standard deviations away from the participant's mean, and/or the Flanker response was inaccurate (except for the model that specifically evaluates Flanker accuracy). The number of experimental trials removed varied based on the region of interest, but at least 85% of trials were included in each analysis after trial exclusion criteria were applied. Although bar plots will show raw reaction time data to improve readability, all models evaluating reading/reaction time were run with the log of reading/reaction time.

Regression Models

To address the primary question about the effect of prior Flanker on current sentence reading, I conducted four simple 2x2 (Flanker congruency x sentence type) linear mixed-effects models using the lmerTest package in R (Kuznetsova et al., 2017). These models evaluated if there was an effect of condition on the log of reading times at the critical word, the word after it, the word two words after it, and/or the summed region from the critical word through two words after it. Conducting multiple models allowed me to pinpoint where (if at all) in the sentence the reader benefits from prior cognitive control engagement. In self-paced reading paradigms, effects are often observed in spillover regions following the critical word (e.g., Bultena et al., 2015), so it is important to consider more than the critical word itself.

I also conducted four additional complex models to determine if any effects and interactions found in the initial models were modulated by experimental half or by list base language (i.e., four of the eight stimuli lists had unilingual sentences and comprehension questions in Spanish and four lists had them in English). To control for multiple comparisons in these eight models, I applied a Bonferroni correction to the simple model and complex model pairs that looked at the same region of interest. For example, the simple model and complex model (which added experimental block and base language along with a word length covariate) that looked at the critical word region were evaluated for effects and interactions that met a $p<0.025$ threshold.

Additional pre-registered models were run on the four sentence regions of interest (critical word, first word after the critical word, second word after the critical word, and summed region from critical word to two after) to evaluate if certain types of code-switches were read slower than others. These models included: Models investigating the effect of the determiner's grammatical gender (feminine or masculine) in noun switches, models investigating the effects of verb tense (perfective or progressive) and switch location (at the auxiliary or at the participle) in verb switches, and models investigating reading differences between noun and verb switches.

Other models were also conducted in exploratory analyses and will be indicated as exploratory when discussed.

All models included crossed random effects for participant and item number. If a model with the fully specified random effects structure did not converge, I first tried different optimizers. If a model still failed to converge, I removed all correlations between random effects, and if needed, then also removed the random effect terms for higher-order interaction terms.

Any two-level factors (prior Flanker congruency, current sentence type, base language, experimental block, grammatical gender, verb tense, noun vs. verb switch type) were contrast coded as -0.5 and +0.5 in the models. When word length (in characters) was included as a covariate in models, it was centered. Marginal R^2 values were extracted with the MuMIn R package (Barton, 2020), and effect sizes were extracted with the sjstats R package (Lüdecke, 2020).

Language History Measures

In exploratory analyses, I investigated the role of individuals' language history on code-switch reading and its interaction with prior Flanker congruency. The two variables of interest were code-switch frequency and diversity of language exposure. The measure of code-switch frequency was obtained by averaging participants' responses to four questions about their code-switching behavior that were on a 1-5 Likert scale; a higher score indicated more frequent code-switching use/exposure. The measure of language exposure was converted into language entropy with the languageEntropy R package (Gullifer & Titone, 2020) by inserting the fraction of time that participants reported being exposed to English, Spanish, and if applicable, another language. A higher language entropy value indicated more balanced exposure; an individual who was exposed to English 50% of the time and to Spanish 50% of the time would receive an entropy score of 1.

Results

Flanker Manipulation Check

To ensure the quality of my results, I first looked at Flanker responses as the dependent variable to determine if participants demonstrated the classic Flanker effect of longer reaction times and lower accuracy for incongruent Flanker trials as compared with congruent Flanker trials. In the model looking at the log of Flanker reaction time, there was a significant main effect of Flanker congruency on Flanker reaction time (RT) such that participants responded slower to incongruent trials (M= 647ms, SD=135ms) than to congruent trials (M= 542ms, SD= 124ms ; β=0.188, η_p^2=0.90, $p<0.0001$). Participants also responded faster to Flanker trials in the second block than in the first block (β=-0.05, η_p^2=0.50, $p<0.001$), but there was no interaction between Flanker congruency and block (η_p^2=0.004, p=0.07; model marginal R^2=0.16).

In the logistic regression model looking at Flanker accuracy, there was a significant main effect of Flanker congruency on Flanker accuracy: Participants responded less accurately to incongruent trials (M=87.92%, SD=32.60%) than to congruent trials (M=99.28%, SD=8.48%; β=-5.31, $p<0.0001$). There was no main effect of block (p=0.32) or interaction between block and Flanker congruency when predicting Flanker accuracy (p=0.41).

Effect of Prior Flanker on Sentence Reading

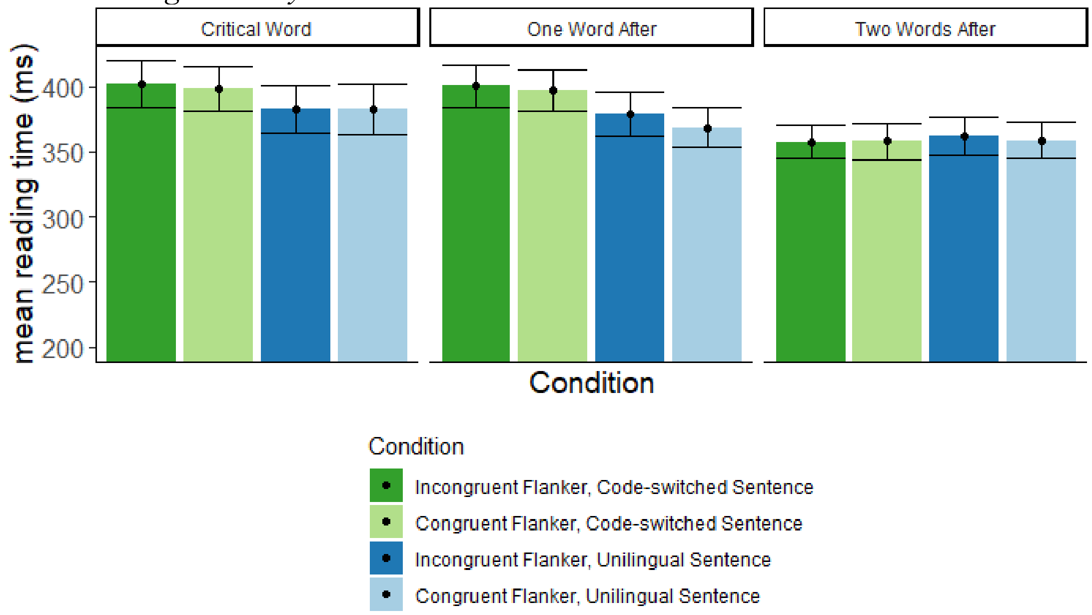
For each of the four regions of interest, I ran two models: (1) A simple model with the predictors Prior Flanker Congruency (congruent or incongruent), Current Sentence Type (unilingual or code-switched), and their interaction, and (2) A complex model that included additional predictors of base language (English or Spanish) and experimental block as well as a covariate for the region's character length. Full regression results for the complex models are available in Appendix A (Tables S1-S4). Effects under a p-value threshold of 0.025 were considered significant in these eight models based on a Bonferroni correction within each region of interest.

Critical Word

As can be seen in the first panel of Figure 2, code-switched words (in green) appear to be read slower than words that continue in the sentence's initial language (in blue). The simple model confirmed this main effect of sentence type: Code-switched words were read slower ($M=400ms$, $SD=178ms$) than their unilingual equivalents ($M=383ms$, $SD=189ms$), demonstrating a classic switch cost effect ($\beta=-0.05$, $\eta_p^2=0.13$, $p=0.0001$). There was no main effect of prior Flanker congruency ($\eta_p^2<0.001$, $p=0.94$), and the predicted interaction between prior Flanker congruency and sentence type did not emerge ($\eta_p^2= 0.001$, $p=0.60$; model marginal $R^2= 0.004$). That is, reading times did not seem to differ when participants had just completed an incongruent Flanker trial (shown in dark colors in Figure 2) as compared with when they had just completed a congruent Flanker trial (shown in light colors in Figure 2).

Figure 2

Word Reading Times by Condition



Note. Bars represent standard error.

The complex model found a three-way interaction between sentence type, base language, and block (β=0.09, η_p^2=0.06, p=0.002) as well as a two-way interaction between sentence type and base language (β=0.14, η_p^2=0.27, p<0.0001; model marginal R^2=0.11). I used the emmeans R package (Lenth, 2021) to do an exploratory investigation of these interactions based on the model's predicted reading times; the same package was used for all exploratory follow-up analyses reported in this section. In interpreting these results, it is important to remember that code-switched sentences always started in Spanish and switched to English at the critical word, while the language of unilingual sentences was determined by the base language assigned to each participant (see Table 3).

Table 3

Sample Critical Sentence by Sentence Type and Base Language

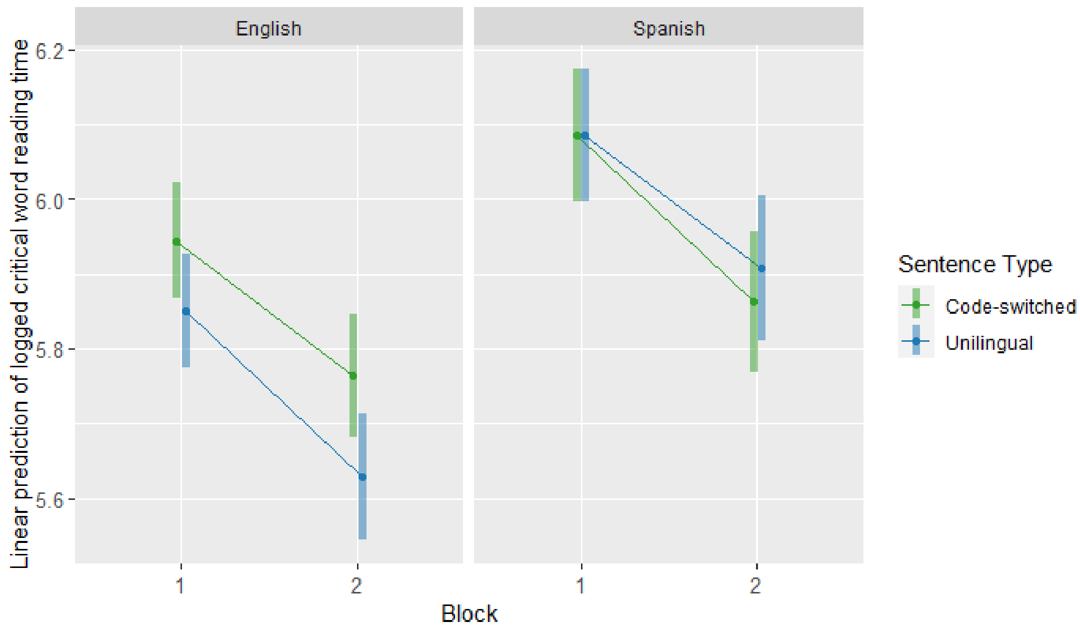
	English base language	**Spanish base language**
Code-switched sentence	El pianista ganó el <u>prize</u> with his original song.	El pianista ganó el <u>prize</u> with his original song.
Unilingual sentence	The pianist won the <u>prize</u> with his original song.	El pianista ganó el <u>premio</u> con su composición original.

Note. The critical word is underlined. Each participant would only be exposed to only version of any particular sentence.

For bilinguals in the English base language, the critical word was always in English. As we can see in the first panel of Figure 3, regardless of block, bilinguals read the critical word faster in a unilingual context (in blue) than when it was code-switched (in green), demonstrating a switch cost that was also reflected in the model's main effect of sentence type (β=-0.05, η_p^2=0.13, p=0.0001). For bilinguals in the Spanish base language, the critical word was in Spanish for the unilingual condition or was code-switched into English for the code-switched condition. In the first block, bilinguals in the Spanish base language list showed no difference in critical word reading time for code-switched critical words versus critical words that continued the sentence in Spanish. This apparent lack of difference between code-switched and unilingual reading may actually reflect a switch cost: Participants were overall faster when reading in English as demonstrated by a main effect of base language (β=0.19, η_p^2=0.10, p=0.002), but they read the code-switched English word as slowly as a Spanish word in this case. In the second block, bilinguals in the Spanish base language began to trend towards reading the critical word faster when it code-switched into the dominant English language than when it continued in Spanish, although this difference was not significant in an emmeans pairwise exploratory analysis. Regardless of base language or sentence type, participants read the critical word faster in the second block, which was supported by a main effect of block in the complex model (β=-0.20, η_p^2=0.41, p<0.0001). The covariate for word length was also significant, indicating that longer words were read slower (β=0.01, η_p^2=0.06, p<0.0001).

Figure 3

Predicted Critical Word Reading Times

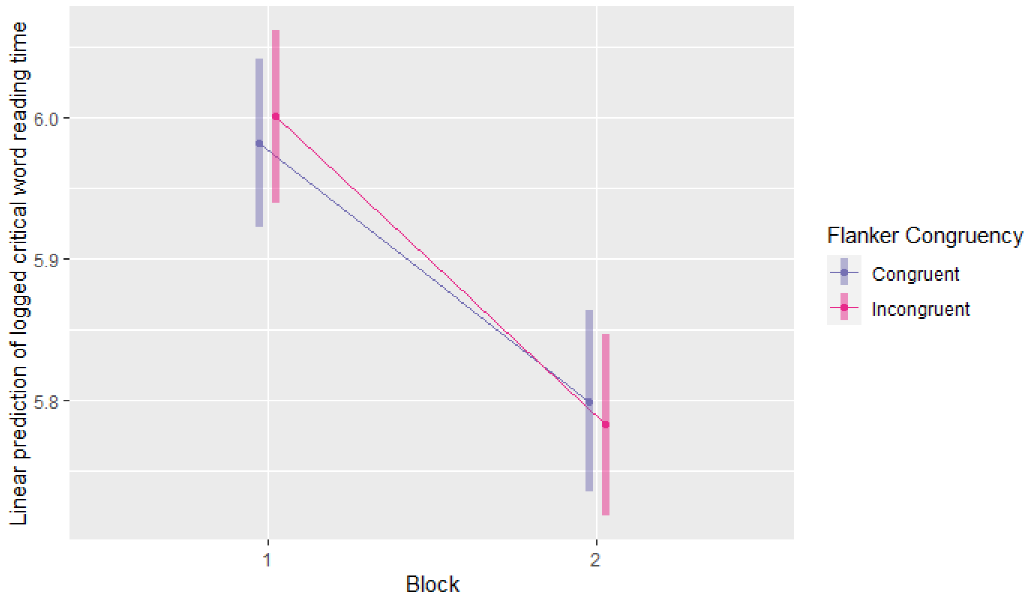


Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

There was also a two-way interaction between prior Flanker congruency and block (β=-0.03, η_p^2=0.05, p=0.012). Exploratory plotting based on the model's predicted reading times (see Figure 4) revealed again that participants read critical words faster overall in the second block. However, in the first block, the critical word was read slower when the prior Flanker trial was incongruent, perhaps reflecting a post-conflict slow down (e.g., Kan et al., 2013; Ullsperger et al., 2005). In the second block, an exploratory pairwise comparison showed that there was no significant difference in critical word reading time based on prior Flanker congruency, although critical word reading times trended towards being faster after a prior incongruent Flanker trial. There was no main effect of prior Flanker congruency in the complex model (η_p^2<0.001, p=0.87).

Figure 4

Predicted Critical Word Reading Times: Prior Flanker x Block Interaction



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

First Word After Critical Word

The simple model to predict the log of reading time for the first word after the critical word revealed a main effect of sentence type such that the word after a code-switch was read slower (M=399ms, SD=161ms) than a unilingual equivalent (M=373ms, SD=163ms), indicating that the switch cost extended beyond the switched word itself (β=-0.07, η_p^2=0.23, p<0.0001). There was no effect of prior Flanker congruency (η_p^2=0.006, p=0.42) and no prior Flanker x sentence type interaction (η_p^2<0.001, p=0.72; model marginal R^2=0.008).

The complex model found a two-way interaction between sentence type and base language (β=0.12, η_p^2=0.25, p<0.0001; model marginal R^2=0.12). Exploratory investigation with the emmeans package showed that this interaction followed the same pattern as it did with the critical word reading times: Participants in the English base language read the word after a code-switch into English slower than when the entire sentence was in English, but participants in the Spanish base language read the word after a code-switch into English at the same speed as the equivalent Spanish word in a fully Spanish sentence. These effects also appear in the complex model as a main effect of sentence type (β=-0.06, η_p^2=0.22, p<0.0001) and a main effect of base language (β=0.18, η_p^2=0.11, p=0.0007).

As with the critical word, participants read the first word after the critical word faster in the second block than in the first block (β=-0.18, η_p^2=0.37, p<0.0001), and read this word slower when it was longer (β=0.02, η_p^2=0.09, p<0.0001). There was no effect of prior Flanker congruency (η_p^2=0.006, p=0.35).

Second Word After Critical Word

The simple model predicting the log of reading time for the second word after the critical word revealed no main effect of sentence type (η_p^2<0.001, p=0.90), no main effect of prior Flanker congruency (η_p^2=0.001, p=0.79), and no interaction between them (η_p^2=0.001, p=0.60; model marginal R^2<0.0001). These results suggest at first glance that any switch costs abated by the time participants reached the second word after a switch, although the complex model provided more detail about this finding.

The complex model also found no main effect of sentence type (η_p^2=0.02, p=0.17), but did reveal a two-way interaction between sentence type and base language (β=0.14, η_p^2=0.40, p<0.0001; model marginal R^2=0.11). Exploratory analysis showed that this interaction was driven by the second word after a code-switch being read slightly slower than a unilingual equivalent in the English base language but the opposite effect in the Spanish base language, again reflecting faster reading times for English words than Spanish words. This effect also appeared in the main effect of base language (β=0.15, η_p^2=0.10, p=0.002).

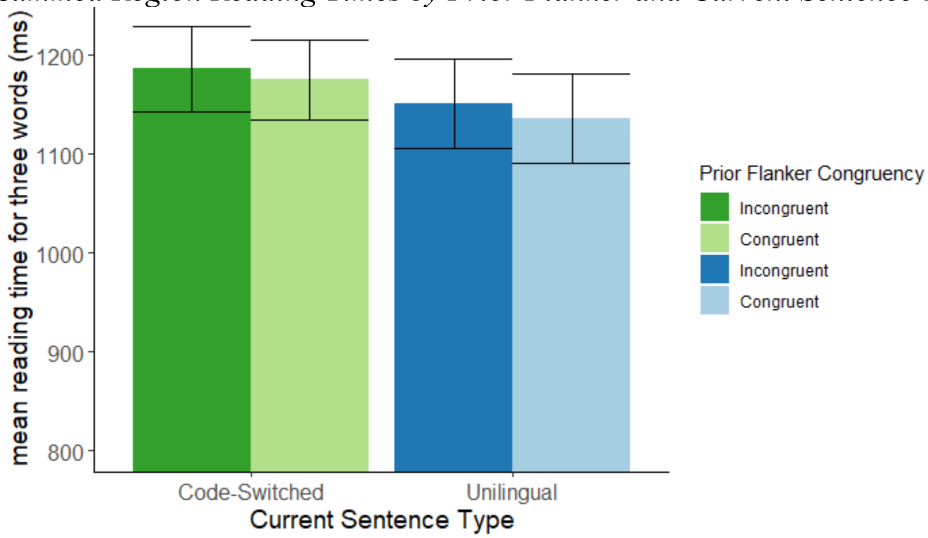
There was no main effect of prior Flanker congruency (η_p^2=0.001, p=0.79) on reading of the second word after the critical word. However, as with the other regions of interest, words were read faster in the second block (β=-0.18, η_p^2=0.34, p<0.0001) and longer words were read slower (β=0.01, η_p^2=0.05, p<0.0001).

Summed Critical Word to Two Words After Region

As seen in Figure 5, when looking at the summed region of the critical word through two words after it, there appears to be a switch cost but no effect of prior Flanker congruency. The simple model predicting logged reading time for the summed region confirmed this by revealing a main effect of sentence type (β=-0.04, η_p^2=0.10, p=0.0008), with slower reading times in this three-word region for code-switches (M=1181ms, SD=419ms) than for single-language sentences (M=1143ms, SD=456ms). There was no main effect of prior Flanker congruency (η_p^2=0.003, p=0.60) or interaction between prior Flanker congruency and sentence type (η_p^2=0.001, p=0.66; model marginal R^2=0.003).

Figure 5

Summed Region Reading Times by Prior Flanker and Current Sentence Type



Note. Bars represent standard error.

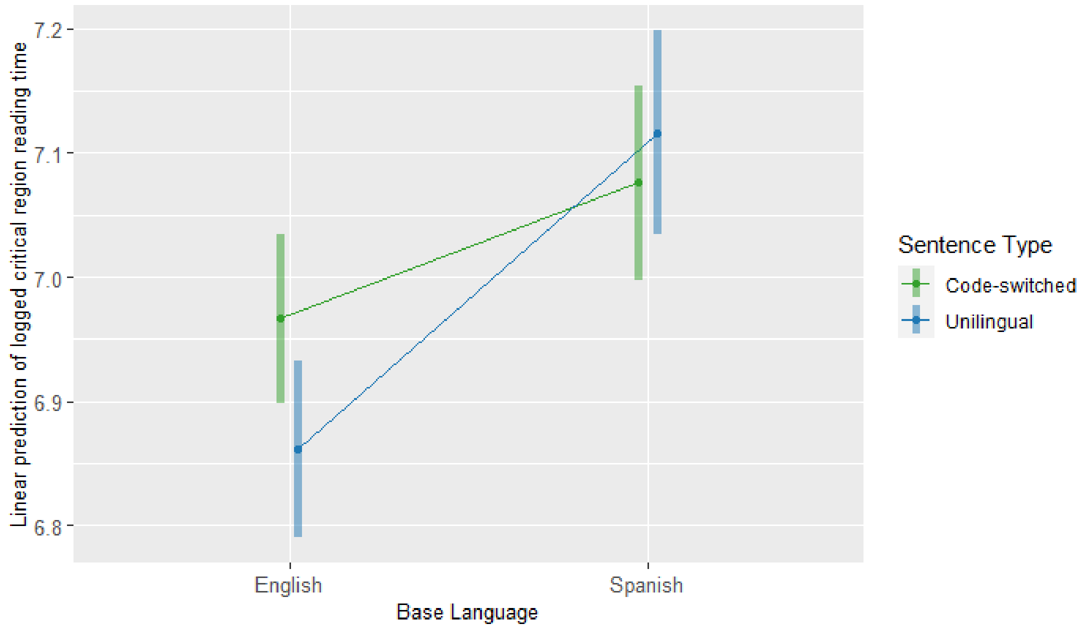
The complex model revealed a two-way interaction between sentence type and base language (β=0.15, η_p^2=0.36, p<0.0001; model marginal R^2=0.14). Exploratory analysis (see Figure 6) showed that in the English base language, the critical region was read slower in a code-switched context than in a unilingual context, which was also reflected in a main effect of sentence type (β=-0.03, η_p^2=0.08, p=0.002). In the Spanish base language, this region was read slower when the sentence continued in Spanish as compared with when it switched to the dominant language, English. Overall, people read slower in the Spanish base language, which was substantiated by a main effect of base language (β=0.18, η_p^2=0.11, p=0.0007). A three-way interaction between sentence type, base language, and block approached significance but did not survive the Bonferroni correction (η_p^2=0.04, p=0.029).

The critical regions of sentences were read faster in the second block (β=-0.21, η_p^2=0.41, p<0.0001) and were read slower if the region had a longer character length

(β=0.01, η_p^2=0.12, p<0.0001). There was no effect of prior Flanker congruency on reading of the summed critical region (η_p^2=0.003, p=0.52).

Figure 6

Predicted Summed Critical Region Reading Times



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

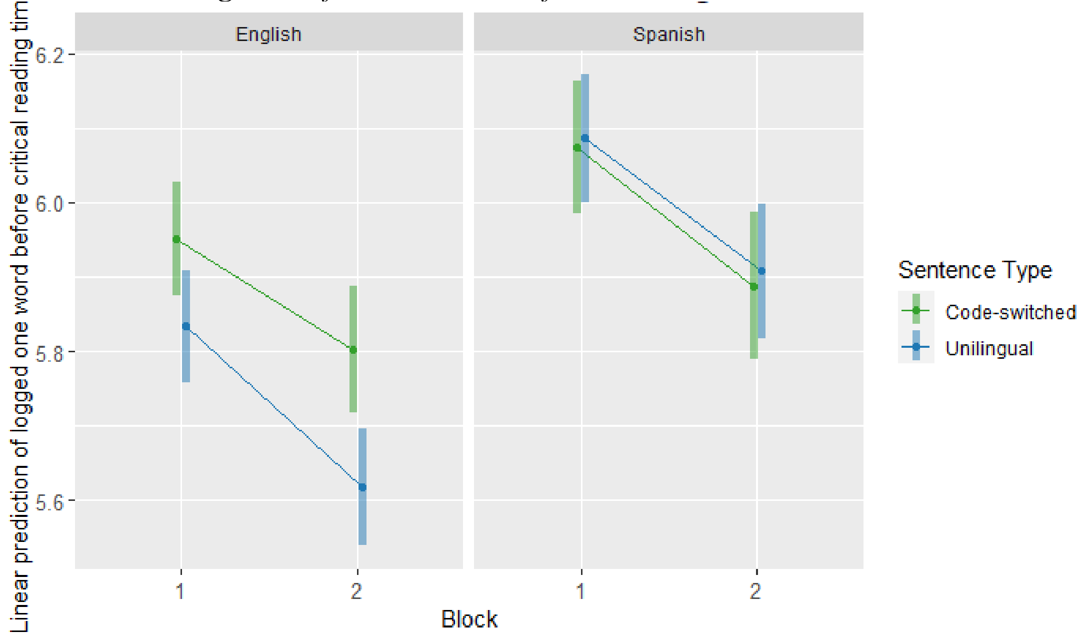
Words Before the Critical Word

Pre-registered analyses on the two words before the critical word were conducted to ensure that there was no unexpected effect of the manipulated variables of interest before the code-switched portion of the sentence. A separate model for each of the two words before the critical word was run. These models were akin to the complex models above, including block and base language as predictors and word length as a covariate; however, a p-value significance threshold of 0.05 was used for these models. In interpreting these results, note that these two words were in Spanish for all code-switched sentences and for unilingual sentences in the Spanish base language while they were in English only for unilingual sentences in the English base language (refer back to Table 3). Full regression results for these two models can be found in Appendix A (Tables S5 and S6).

The complex model predicting logged reading time of the word before the critical word found a three-way interaction between sentence type, block, and base language (β=0.07, η_p^2=0.03, p=0.008); a two-way interaction between sentence type and base language (β=0.16, η_p^2=0.34, p<0.001); and a two-way interaction between sentence type and block (β=-0.03, η_p^2=0.04, p=0.045; model marginal R^2=0.12). As we can see in Figure 7, exploratory analysis showed that in the English base language, this word was read slower in the code-switched condition (when it appeared in Spanish) than in the unilingual condition (when it appeared in English), particularly in the second block. This effect appeared to drive the main effect of sentence type (β=-0.07, η_p^2=0.25, p<0.0001). In the Spanish base language, there was no difference in reading time for this word based on sentence type, as the word was in Spanish for both conditions. Overall, this region was read faster in the second block than in the first (β=-0.18, η_p^2=0.35, p<0.0001), was read faster in the English base language than in the Spanish base language (β=0.19, η_p^2=0.10, p=0.001), and was read faster when the word was shorter (β=0.01, η_p^2=0.12, p<0.0001). There was no effect of prior Flanker congruency on reading time of this region (β=0.002, η_p^2<0.001, p=0.81).

Figure 7

Predicted Reading Times for One Word Before Critical Word



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

The complex model predicting logged reading time of the word two words before the critical word revealed a three-way interaction between sentence type, block, and base language (β=0.10, η_p^2=0.07, p=0.005) and a two-way interaction between sentence type and base language (β=0.15, η_p^2=0.16, p<0.0001; model marginal R^2=0.11). Exploratory analysis showed a similar pattern as was found with the word before the code-switch: no difference in reading times by sentence type for the Spanish base language, but slower reading for the code-switch condition for the English base language which reflected slower reading in Spanish. Again, there was a main effect of sentence type (β=-0.07, η_p^2=0.12, p=0.0002), main effect of block (β=-0.23, η_p^2=0.35, p<0.0001), a main effect of base language (β=0.20, η_p^2=0.09, p=0.002), and a significant covariate for character length (β=0.01, η_p^2=0.07, p<0.0001). There was no effect of prior Flanker congruency (β=-0.01, η_p^2=0.006, p=0.22).

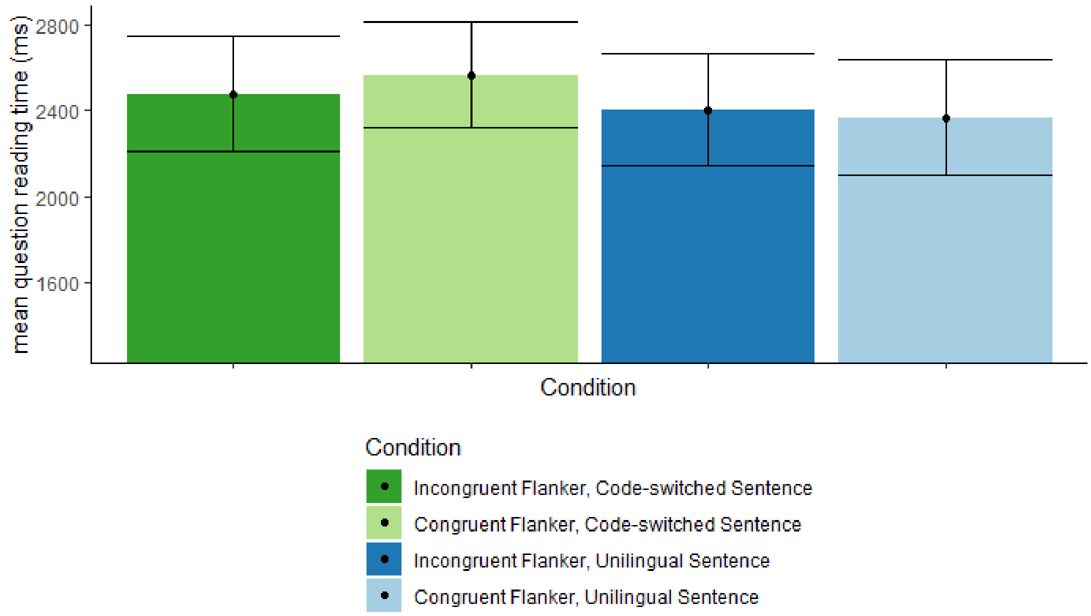
Off-line Comprehension

My primary analyses investigated whether prior Flanker congruency affected on-line comprehension of code-switched and unilingual sentences as measured by reading time. As an exploratory analysis, I also investigated whether prior Flanker congruency affected off-line comprehension as measured by accuracy and reading time of comprehension questions that followed the sentences. These two models include the fixed effects of prior Flanker congruency, sentence type, and base language. Base language was included as it determined which language the comprehension question was displayed in.

In the model predicting the logged question reading time, there was a main effect of question language (β=0.34, η_p^2=0.22, p<0.0001), reflecting faster reading times in English than in Spanish. There was also a main effect of sentence type (β=-0.04, η_p^2=0.03, p<0.0001; model marginal R^2=0.10), reflecting slower comprehension question reading after a code-switched vs. unilingual sentence, which can also be observed in Figure 8 in which the green bars are higher than the blue bars. Since all code-switches were from Spanish to English, in the English base language this may reflect a lingering switch cost, while in the Spanish base language this may reflect a new switch cost as participants switched from a code-switched sentence ending in English to a question in Spanish. No effect or interaction of prior Flanker congruency was significant.

Figure 8

Comprehension Question Reading Times by Condition



Note. Bars represent standard error.

In the logistic regression predicting comprehension question accuracy, the only significant predictor was question language (β=-0.47, p=0.009), reflecting higher accuracy for questions in the dominant English language as compared with Spanish.

Effect of Switch Type on Sentence Reading

To address Hypotheses 2a-2d, I ran analyses to determine if different code-switch types were read differently. In interpreting these results, it is important to remember that my stimuli only allow for direct same-item comparisons for code-switched vs. unilingual sentences. That is, each item had a code-switched and a unilingual version (although each participant only saw one). However, each determiner-noun sentence had *either* a feminine determiner or a masculine determiner, and each verb sentence had *either* a progressive or perfective tense and *either* an auxiliary switch location or a participle switch location (see Table 2). Therefore, sentences that were assigned to have an auxiliary switch are likely qualitatively different (in terms of content, word length, word frequency etc.) than sentences that were assigned to a participle switch, for example.

Determiner-Noun Switches

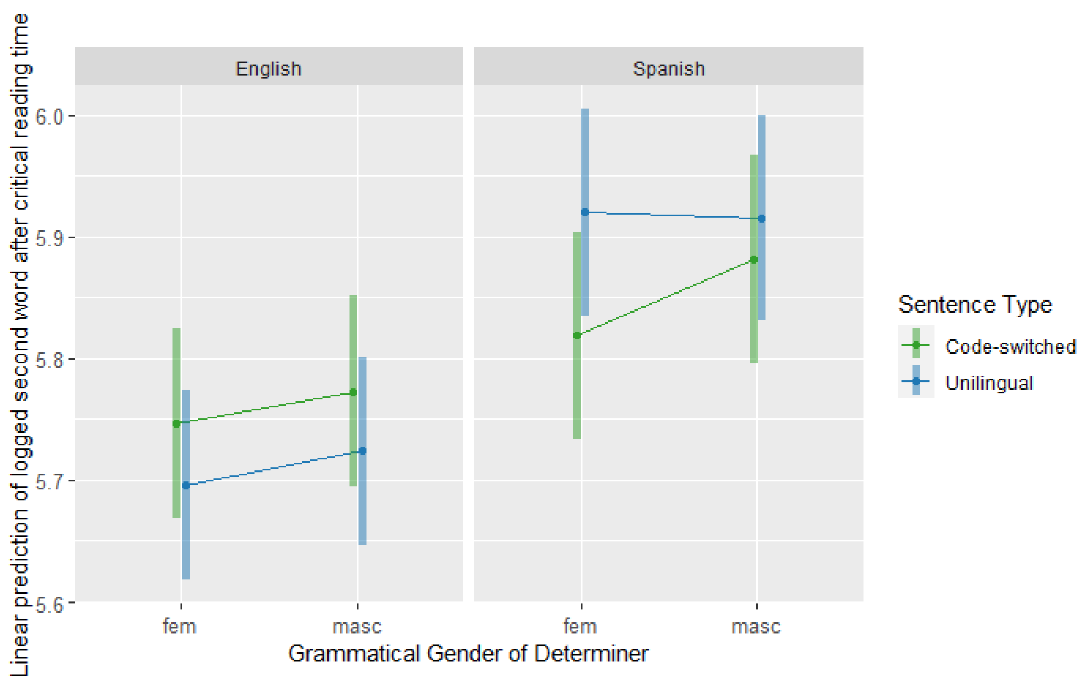
I ran four models on the data generated from participants reading the 48 determiner-noun code-switch sentences to determine if the determiner's grammatical gender in Spanish affected reading of the code-switched (or unilingual equivalent) noun. These models' fixed effects included sentence type, grammatical gender, and base language, along with (centered) region length in characters as a covariate. Full regression results for these four models are in Appendix A (Tables S7-S10).

The models looking at the switch word, one word after it, and the summed region (the switch word through two words after) had similar patterns of results. These models found no main effect of the determiner's grammatical gender and no interactions involving grammatical gender ($ps>0.20$, $\eta_p^2<0.02$). The three models all revealed a two-way interaction between base language and sentence type ($\beta\geq0.10$, $\eta_p^2\geq0.14$, $ps<0.0001$) and main effects that closely paralleled the findings outlined above (models' marginal $R^2\geq0.06$).

However, the model looking at the second word after the noun code-switch included interactions involving the grammatical gender of the pre-switch determiner. Although this model did not show a main effect of grammatical gender ($\eta_p^2=0.01$, $p=0.47$), there was a three-way interaction between sentence type, grammatical gender, and base language ($\beta=-0.07$, $\eta_p^2=0.01$, $p=0.02$) as well as a two-way interaction between sentence type and grammatical gender ($\beta=-0.03$, $\eta_p^2=0.09$, $p=0.04$; model marginal R^2=0.06). Exploratory analysis indicated that in the Spanish base language only, participants read the second word after the critical word faster in a code-switched context than a unilingual context when following a feminine determiner but not a masculine determiner as shown in Figure 9. Participants tended to read faster in English than in Spanish, so it seems that after the code-switch from Spanish to English and its associated “cost,” participants began returning to their faster baseline English reading speed in stimuli that included a feminine determiner but not a masculine determiner. Given that feminine determiner sentences and masculine determiner sentences were not matched, it is likely this finding is due to different sentence features (e.g., syntactic categories) at this region two words after a switch. Besides these additional interactions, the results of this model paralleled those of the other regions of interest.

Figure 9

Determiner-Noun Stimuli: Predicted Reading Times for Second Word After Critical



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

Verb Switches

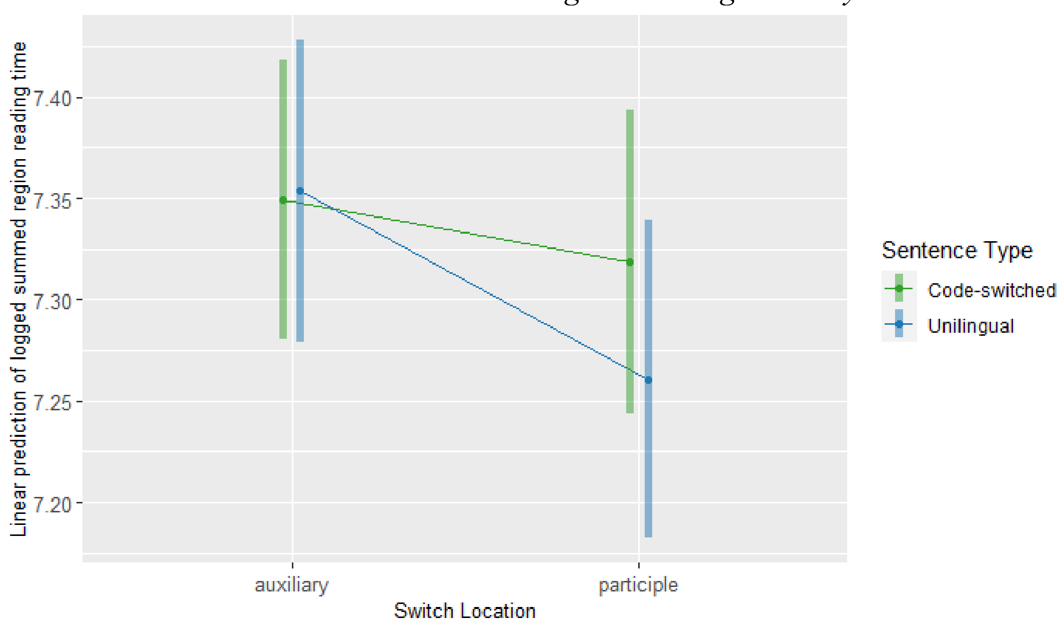
I ran four models (one for each region of interest) on the data generated from participants reading the 48 verb code-switch sentences to determine if verb tense (perfective or progressive) and switch location (at the auxiliary or at the participle) affected sentence reading. These models' fixed effects included sentence type, verb tense, switch location, and base language, along with (centered) region length in characters as a covariate. For these models, the critical region was considered to be the combined two-word auxiliary and participle region (e.g., "han signed," "have signed," "están signing," or "are signing") instead of the critical word alone. This critical region adjustment was done to allow for equivalent comparisons of the same regions for switches that occurred at different locations. Full regression results for these four models are in Appendix A (Tables S11-S14).

None of the models revealed an interaction between verb tense and switch location in predicted logged reading times ($\eta_p^2 \leq 0.07$, $ps > 0.07$), so each factor will be discussed separately.

Switch Location. At the critical region and the one word after region, there was a two-way interaction between switch location and sentence type ($\beta \geq -0.06$, $\eta_p^2 \geq 0.07$, $ps < 0.001$). Follow-up analysis on predicted reading times indicated that there were no switch costs for auxiliary switches (e.g., "ellos have signed" or "ellos are signing") but there were switch costs for participle switches (e.g., "ellos han signed" or "ellos están signing"). This interaction also appeared in the summed region from the critical verb phrase through two words after ($\beta = -0.06$, $\eta_p^2 = 0.12$, $p < 0.0001$) as can be seen in Figure 10. This interaction did not emerge in the model evaluating the second word after the critical verb phrase, although there was a main effect of switch location at that word (β=-0.09, η_p^2=0.12, p=0.02). Follow-up analysis indicated that the second word after the verb phrase was read faster when the sentence was marked as an "auxiliary switch." However, as this did not interact with sentence type (i.e., with whether or not a switch actually occurred), this may simply reflect a difference in the types of words (e.g., syntactic category) that followed stimuli assigned to be auxiliary rather than participle switches.

Figure 10

Verb Stimuli: Predicted Summed Critical Region Reading Times by Switch Location



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

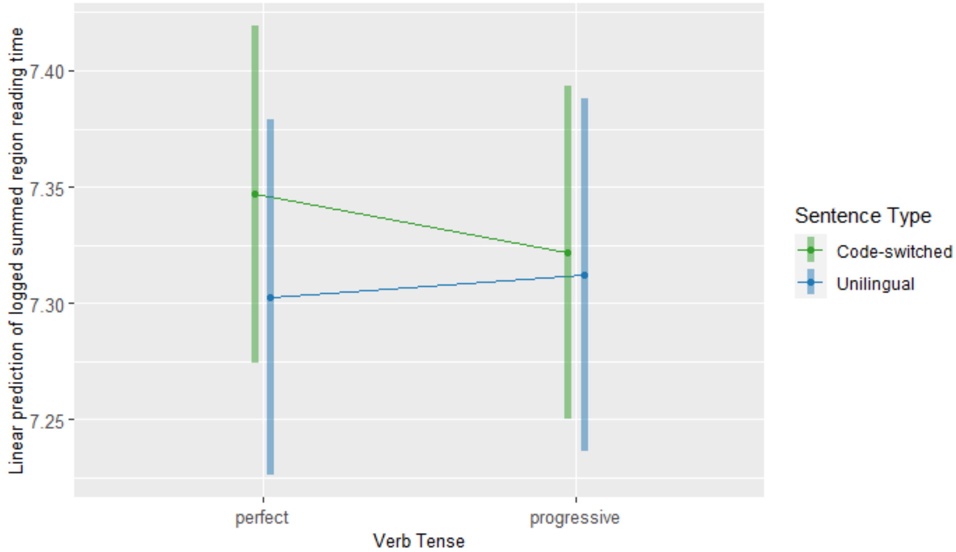
Verb Tense. At the critical two-word auxiliary-participle region, there was a two-way interaction between verb tense and base language (β=0.05, η_p^2=0.008, p=0.009). Follow-up analysis revealed that participants read faster in the English base language than the Spanish base language and that effect was somewhat larger for progressive verb phrases than perfective verb phrases. Because sentence type (code-switched vs. unilingual) was not included in this interaction, it may again indicate a difference in stimuli with progressive vs. perfective verbs rather than an indication that one verb tense was more difficult to read than the other.

At the word after the critical verb phrase, there was a two-way interaction between sentence type and verb tense (β=0.04, η_p^2=0.007, p=0.011). Follow-up analysis indicated that there was a switch cost at this word when the switch occurred in a perfective verb phrase but not in a progressive verb phrase. This interaction was also significant in the summed region from the critical verb phrase through two words after (β=0.03, η_p^2=0.03, p=0.012; see Figure 11) and although not significant, the data trended in this direction for the critical verb phrase as well (p=0.051).

Overall, the models showed additional results consistent with the patterns described above (i.e., two-way sentence type x base language interaction, main effect of base language, main effect of sentence type, and for some models a significant covariate for character length; models' marginal $R^2 \geq 0.06$).

Figure 11

Verb Stimuli: Predicted Summed Critical Region Reading Times by Verb Tense



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

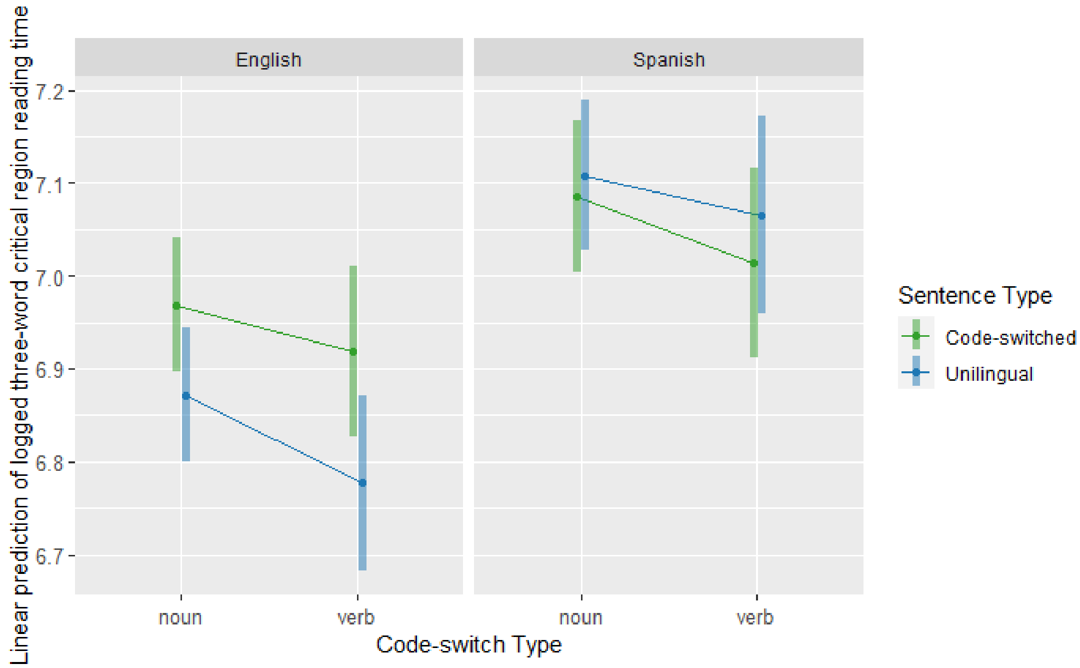
Noun vs. Verb Switches

Having looked at code-switch reading for noun switches and verb switches separately, I also wanted to compare the two directly. To do so, I ran four models (one on each region of interest) to compare reading of determiner-noun switches and verb switches. For verb switches, only participle location items were included so that the comparison would be between content words only—that is, the noun for determiner-noun switches and the participle for verb switches. These models' fixed effects included sentence type, base language, and code-switch type along with (centered) region length in characters as a covariate. Full regression results for these four models are in Appendix A (Tables S15-S18).

All models demonstrated a sentence type by base language two-way interaction ($\beta \geq 0.11$, $\eta_p^2 \geq 0.16$, $ps < 0.0001$; models' marginal $R^2 \geq 0.05$) in line with the previously discussed results. At the critical word, two words after it, and the summed region from the critical word through two words after, there was a three-way sentence type by base language by code-switch type (noun vs. verb) interaction ($\beta \geq 0.06$, $\eta_p^2 \geq 0.01$, $ps<0.03$). Exploratory analysis (visualized in Figure 12) indicated that in the English base language, these regions were read slower in the code-switched vs. unilingual condition, and that switch cost effect appeared larger for verb switches than noun switches. In the Spanish base language, there was no switch cost or noun vs. verb reading difference for the critical word or the summed region. In the Spanish base language at two words after the critical word, there appeared to be a switch benefit rather than no difference in code-switched vs. unilingual conditions as the sentence continued in the dominant English language. This benefit appeared slightly larger after a verb switch than a noun switch. The three-way interaction did not emerge in the model for one word after the critical word ($\eta_p^2=0.002$, $p=0.55$). At the critical word, there was also a significant two-way interaction between base language and code-switch type ($\beta=0.05$, $\eta_p^2=0.05$, $p=0.02$); exploratory analysis indicated this interaction was due to nouns being read slower than verbs (regardless of if they were code-switched), particularly in the English base language.

Figure 12

Predicted Summed Critical Region Reading Times by Noun vs. Verb Stimuli



Note. Bars represent 95% confidence intervals generated by the emmip function in the emmeans R package.

Potential Modulators of Conflict Adaptation

Although the primary analyses did not reveal any prior Flanker congruency effects or interactions indicative of conflict adaptation, I still conducted exploratory analyses to determine if conflict adaptation may emerge for a subset of stimuli or individuals to address Hypotheses 3 and 4. Full regression results for these models are in Appendix A (Tables S19-S21).

Code-switch Type

I ran a model to determine if code-switch type might affect the presence of conflict adaptation from a prior Flanker trial to a subsequent sentence reading trial. In this case, type of code-switch was not contrast coded but rather treated as a 3-level factor—including determiner-noun switches, progressive tense verb switches, and perfective tense verb switches—which was dummy coded using the default in R. The model was a simple model that did not include block, base language, or character length. The model predicted logged reading times for the summed region from the critical word through two words after it to get an overall sense for if there was an effect in that region. There were no significant interactions or main effects including prior Flanker congruency ($\eta_p^2 \leq 0.003$, ps>0.32; model marginal $R^2=0.007$).

Language History

I ran one model on the summed region from the critical word through two after to determine if individuals' code-switching experience or language entropy would affect the (lack of) sentence type by prior Flanker congruency interaction. The model's fixed effects included prior Flanker congruency, sentence type, code-switching experience score, and language entropy. There were no significant interactions or main effects including prior Flanker congruency in the model ($\eta_p^2 \leq 0.013$, ps>0.28; model marginal $R^2=0.03$).

Switch Location

In an additional exploratory analysis, I ran one model on the summed region from the critical word through two after to determine if the critical word's location in the sentence affected the presence of a conflict adaptation effect. The model's fixed effects included prior Flanker congruency, sentence type, and critical word location. The critical word's location was a numeric variable indicating what number word it was in the sentence. There was a main effect of critical word location ($\beta=0.02$, $\eta_p^2=0.04$, $p=0.01$, model marginal $R^2=0.008$) with participants reading critical words that appeared later in the sentence somewhat slower than critical words that appeared earlier. However, critical word location did not interact with sentence type or prior Flanker congruency, ($\eta_p^2 \leq 0.001$, $ps > 0.66$), indicating this effect did not alter the absence of conflict adaptation and likely reflected a characteristic of reading in general rather than code-switch reading in particular.

Discussion

In this conflict adaptation study, Spanish-English bilinguals demonstrated switch costs, but against my prediction, those costs were not attenuated when bilinguals had previously responded to an incongruent Flanker trial intended to increase their cognitive control engagement. I did observe expected differences for reading of different types of code-switches: There were smaller switch costs (or even no switch costs) for perfect tense switches than progressive tense switches, for auxiliary location switches than participle location switches, and for noun switches than for verb switches, although the predicted difference in masculine vs. feminine noun switches did not emerge. However, neither code-switch type nor individual language experience factors altered whether or not conflict adaptation was observed.

In the remainder of this section, I return to the issue of why bilinguals experience conflict during code-switch comprehension, discuss how different code-switch types altered reading, and briefly touch on the broader implications of this line of investigation.

The Conflict in Code-switch Comprehension

Past work found a conflict adaptation effect from code-switch reading to a subsequent Flanker task, indicating that some aspect of code-switch comprehension involves conflict detection and resolution (Adler et al., 2020). The lack of a conflict adaptation effect in this study warrants additional conversation about what type of conflict bilinguals experience during code-switch comprehension. In the Introduction, I laid out two possibilities: (1) Bilinguals experience an early lexical/syntactic conflict during code-switch comprehension when they attempt to integrate code-switched input into their understanding; under this account, we would have expected a conflict adaptation effect in this study as predicted, or (2) Bilinguals experience a later pragmatic conflict during code-switch comprehension when the context of a code-switch fails to align with their prior expectations; under this account, we would not necessarily expect to observe conflict adaptation in this study.

Given a null result, this study cannot provide evidence for either account. One possibility is that the reading time measure used in this study was not sensitive enough to detect a conflict adaptation effect. Although the study had 83% power to detect a small conflict adaptation effect of the same size that Adler and colleagues found (2020), the effect may be smaller in this direction or may be more difficult to detect in a (remotely conducted) self-paced reading paradigm as compared to in Flanker reaction times. Future studies could test for a lexical/syntactic conflict using more sensitive measures such as eye-tracking while reading which would allow for more naturalistic free-reading of sentences, or a visual world eye tracking paradigm which would allow participants to hear code-switches which may be more natural for some bilinguals.

It is worth noting the one indication in my results that prior Flanker congruency could affect reading of the critical word in the current study: In the first block, participants read the critical word slower after an incongruent Flanker trial than after a congruent one. This effect could indicate general post-conflict slowing in which participants respond slower after an incongruent trial regardless of current trial type, perhaps as they adopt a more cautious approach to the task (e.g., Kan et al., 2013). The effect applied to both code-switched and unilingual stimuli, consistent with post-conflict slowing. This effect provides some evidence that the current paradigm was sensitive enough to detect effects of a prior Flanker trial on current sentence reading, although why a slowing effect appeared at the critical word (regardless of if it was code-switched or not) but not at the two words prior to the critical word (which were more immediately after the Flanker trial) is unclear.

Reconsidering the Pragmatic Conflict Account

Having failed to find support for the lexical/syntactic conflict account, let us consider the pragmatic account again and how it could be tested in the future. The pragmatic account posits that bilinguals experience a late conflict after their initial interpretation of a code-switch as they consider the pragmatic context in which the switch occurred.

One limitation of the current study in testing this hypothesis was that I did not intentionally manipulate pragmatic context. Participants read each isolated sentence within the same experimental/virtual context, and the sentential context (e.g., frequencies for the words following a code-switch) was not manipulated. In the Introduction, I offered a wide variety of reasons that bilinguals may experience pragmatic conflict during code-switch comprehension, each with distinct predictions about what would cause more or less pragmatic conflict. These reasons included a social context that does not support code-switching, a physical context (e.g., a university laboratory) that is an unusual environment for code-switching, or various sentential contexts in which a code-switch or its surrounding content presents a mismatch with the bilingual's pragmatic expectations for the sentence. Regarding the various sentential context reasons for pragmatic conflict, it is not entirely clear which sentential contexts would create the most pragmatic conflict (e.g., is it when a code-switch appears in a highly constrained sentence context or when a code-switch is followed by a high-frequency word?).

Therefore, future studies may gain the most traction on this issue by first focusing on manipulating pragmatic conflict through social or physical context. Past work has already demonstrated that code-switch comprehension differs in the presence of a bilingual vs. a monolingual (Kaan et al., 2020). One potential future avenue could be to replicate Adler and colleagues' (2020) code-switch-reading-to-Flanker conflict adaptation paradigm in distinct social contexts that support code-switching more or less. If pragmatic conflict is reduced in contexts in which code-switching is acceptable or expected, then conflict adaptation from code-switch comprehension to Flanker trial may be reduced or eliminated in those contexts. Alternatively, ERP approaches could be used as an even more sensitive measure of the amount of pragmatic conflict experienced, although it would be important to first characterize the ERP signature of the conflict (e.g., is it best represented by N400 or LPC?).

A lingering question left by the pragmatic conflict account is: If the conflict involved in a code-switch occurs at a late, pragmatic level, then why do we observe switch costs at the site of the code-switch itself? Past work suggests that code-switches elicit a switch cost in processing (e.g., Altarriba et al., 1996; Bultena et al., 2015) and that code-switches involve some sort of conflict (Adler et al., 2020). Under the pragmatic conflict account, these two consequences of a code-switch—the slowing of sentence processing and the conflict—need not be conflated. Switch costs may reflect bilinguals taking more time to integrate the input into their linguistic understanding even if that integration does not involve *conflict*. Meanwhile, pragmatic conflict could still occur later in processing as outlined above. If future work does provide support for code-switch comprehension eliciting a pragmatic conflict and not a lexical/syntactic conflict, then that work should also address why sentence processing slows down at a code-switch if not because of conflicting representations.

Different Switch Types

A secondary purpose of this study was to determine if I could replicate past studies' finding that code-switch types that are more frequent in production are processed faster in comprehension (Beatty-Martínez & Dussias, 2017; Guzzardo Tamargo et al., 2016; see also MacDonald, 2013). In general, my results do support the idea that more frequent code-switches—namely, switches at the auxiliary location, switches in the progressive tense, and mixed noun phrases—evoke smaller (or no) switch costs than less frequent types of code-switches. On its own, this replication lends support to the idea that code-switch experience affects code-switch comprehension and highlights the idea that different linguistic contexts impose distinct processing demands on bilinguals.

Surprisingly, this study did not replicate the production-comprehension parallel for determiner-noun switches: Masculine determiners are more common in corpus analyses (Beatty-Martínez & Dussias, 2017; Beatty-Martínez & Dussias, 2019; Valdés Kroff, 2016) and thus were expected to be read faster in this study. One reason for the lack of a determiner grammatical gender effect could be that the self-paced reading paradigm is not sensitive enough to detect comprehension differences previously detected in ERPs (Beatty-Martínez & Dussias, 2017; Beatty-Martínez & Dussias, 2019). Alternatively, although our participants indicated moderate levels of code-switching experience, research indicates that different bilingual communities code-switch differently, with some using code-switched noun phrases more often than others (Beatty-Martínez & Dussias, 2019). Without directly testing the participants' production tendencies or exposure, I cannot rule out the possibility that their code-switching experience simply does not follow the expected pattern. However, this is unlikely to be a satisfactory explanation, as even non-habitual code-switchers show a preference for masculine determiner code-switches to some degree (Beatty-Martínez & Dussias, 2017).

A third possible explanation is that while masculine determiners are common in code-switched noun phrases, they do not result in faster reading of code-switches due to their un-informativeness about the upcoming noun. Because it is acceptable for a Spanish masculine determiner to occur before a code-switch to an English noun with either a masculine or feminine Spanish translation, masculine determiners are not informative about the upcoming noun in the way that feminine determiners are. Eye-tracking research lends support to this view, with Spanish monolinguals using masculine determiners as a reliable predictive cue but not Spanish-English bilinguals (Valdés Kroff et al., 2017). Thus, while masculine code-switched noun phrases are a more frequent code-switch type than feminine code-switched noun phrases, they may not be read any faster because masculine determiners fail to evoke facilitatory predictive processes that feminine determiners can evoke.

Having replicated many expected patterns of code-switch comprehension difficulty based on code-switch type, these findings could be used to further investigate the possibility that code-switch comprehension involves conflict at the lexical or syntactic level. If code-switches involve a lexical/syntactic conflict because the word that arrives is not expected (at least not expected in the language it arrives in), then more frequent code-switches should be more expected and require less cognitive control engagement to comprehend. Future work could capitalize on this by replicating the current study with self-paced reading, eye-tracking, or ERP techniques using only the less frequent switch types that elicited stronger switch costs in this study. This approach should theoretically heighten the lexical/syntactic conflict and ensure that there is a switch cost present for cognitive control to modulate. Alternatively, future work could replicate Adler and colleagues' study and determine if comprehending less frequent code-switch types results in greater conflict adaptation than more frequent code-switch types by virtue of the code-switches requiring different cognitive control engagement states to comprehend.

General Implications

This study addressed the question of why bilinguals experience conflict during code-switch comprehension. Although the answer to this question is still unclear, it could affect how the field interprets various effects of code-switching. For example, if future work suggests that bilinguals experience a pragmatic conflict during code-switch comprehension, this would warrant consideration of how much we can expect results from past research on code-switching that occurred in laboratory contexts where code-switching is not pragmatically supported to transfer to naturalistic code-switching experiences.

Further, it seems clear from past work that bilinguals' linguistic and cognitive processing can affect one another dynamically and flexibly (Adler et al., 2020; Salig et al., 2021; Wu & Thierry, 2013). The degree to which these effects are observable outside of the laboratory hinges on *why* such effects occur, as investigated here, and likely determines the extent to which these effects can be capitalized on in external domains such as education.

Conclusion

Past work has shown that when bilinguals comprehend a code-switch, they engage cognitive control, indicating that some aspect of code-switch comprehension requires bilinguals to resolve conflict (Adler et al., 2020). If this conflict were at a lexical or syntactic level of representation, we would expect that bilinguals would be better equipped to resolve this conflict and read code-switches faster when they already have cognitive control highly engaged. However, in this study having cognitive control engaged prior to encountering a code-switch did *not* result in bilinguals reading code-switches faster. Future work is needed to determine exactly why bilinguals experience a conflict during code-switch comprehension; regardless, this study provides preliminary support for considering the possibility that bilinguals experience a later pragmatic conflict during code-switch comprehension. Although I have outlined lexical/syntactic conflict and pragmatic conflict accounts separately here, future work should also consider the possibility that conflict may be experienced at multiple and/or different levels of representation during code-switch comprehension, likely in a way that is dependent on context.

Appendices

Appendix A: Supplemental Regression Model Results

Table S1

Complex Model Predicting Logged Critical Word Reading Time

lmer(log(CriticalWordReadingTime) ~ SentenceType* FlankerCongruency * Block * BaseLanguageuage + CenteredCriticalWordLength + (SentenceType*FlankerCongruency*Block|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.89	5.83 – 5.95	**<0.001**
SentenceType1	-0.05	-0.07 – -0.02	**<0.001**
FlankerCongruency1	0.00	-0.01 – 0.01	0.874
Block 1	-0.20	-0.24 – -0.16	**<0.001**
BaseLanguage1	0.19	0.08 – 0.30	**0.001**
CenteredCriticalWordLength	0.01	0.01 – 0.02	**<0.001**
SentenceType1 * FlankerCongruency1	-0.01	-0.03 – 0.02	0.612
SentenceType1 * Block 1	0.00	-0.03 – 0.03	0.918
FlankerCongruency1 * Block 1	-0.04	-0.06 – -0.01	**0.011**
SentenceType1 * BaseLanguage1	0.14	0.09 – 0.18	**<0.001**
FlankerCongruency1 * BaseLanguage1	0.01	-0.02 – 0.03	0.687
Block 1 * BaseLanguage1	0.00	-0.05 – 0.05	0.978
(SentenceType1 * FlankerCongruency1) * Block 1	0.00	-0.05 – 0.05	0.879

SentenceType1 * FlankerCongruency1 * BaseLanguage1	-0.02	-0.07 – 0.03	0.404
SentenceType1 * Block 1 * BaseLanguage1	0.09	0.03 – 0.14	**0.001**
FlankerCongruency1 * Block 1 * BaseLanguage1	-0.03	-0.08 – 0.03	0.336
SentenceType1 * FlankerCongruency1 * Block 1 * BaseLanguage1	-0.00	-0.10 – 0.09	0.952
Observations	8329		

Table S2

Complex Model Predicting Logged Reading Time of First Word After Critical

lmer(log(FirstAfterCriticalReadingTime) ~ SentenceType * FlankerCongruency * Block * BaseLanguage + CenteredFirstAfterWordLength + (SentenceType*FlankerCongruency*Block|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.89	5.84 – 5.94	**<0.001**
SentenceType1	-0.06	-0.08 – -0.04	**<0.001**
FlankerCongruency1	0.01	-0.01 – 0.02	0.350
Block1	-0.18	-0.22 – -0.14	**<0.001**
BaseLanguage1	0.18	0.08 – 0.28	**<0.001**
CenteredFirstAfterWordLength	0.01	0.01 – 0.02	**<0.001**
SentenceType1 * FlankerCongruency1	0.01	-0.02 – 0.03	0.477
SentenceType1 * Block1	-0.00	-0.03 – 0.03	0.890
FlankerCongruency1 * Block1	-0.01	-0.04 – 0.01	0.297

SentenceType1 * BaseLanguage1	0.12	0.08 – 0.16	**<0.001**
FlankerCongruency1 * BaseLanguage1	0.01	-0.02 – 0.04	0.480
Block1 * BaseLanguage1	0.02	-0.03 – 0.07	0.455
(SentenceType1 * FlankerCongruency1) * Block1	-0.01	-0.06 – 0.04	0.664
SentenceType1 * FlankerCongruency1 * BaseLanguage1	0.02	-0.03 – 0.07	0.451
SentenceType1 * Block1 * BaseLanguage1	0.03	-0.02 – 0.08	0.240
FlankerCongruency1 * Block1 * BaseLanguage1	-0.00	-0.05 – 0.05	0.987
SentenceType1 * FlankerCongruency1 * Block1 * BaseLanguage1	-0.01	-0.10 – 0.09	0.919
Observations	8341		

Table S3

Complex Model Predicting Logged Reading Time of Second Word After Critical

lmer(log(SecondAfterCriticalReadingTime) ~ SentenceType * FlankerCongruency * Block * BaseLanguage +CenteredSecondAfterWordLength + (SentenceType*FlankerCongruency*Block|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.83	5.78 – 5.88	**<0.001**
SentenceType1	0.01	-0.01 – 0.03	0.169
FlankerCongruency1	0.00	-0.01 – 0.02	0.788
Block1	-0.18	-0.22 – -0.14	**<0.001**
BaseLanguage1	0.15	0.06 – 0.23	**0.001**
CenteredSecondAfterWordLength	0.01	0.01 – 0.02	**<0.001**

SentenceType1 * FlankerCongruency1	-0.00	-0.03 – 0.02	0.762
SentenceType1 * Block1	-0.01	-0.03 – 0.02	0.545
FlankerCongruency1 * Block1	-0.01	-0.03 – 0.02	0.628
SentenceType1 * BaseLanguage1	0.14	0.11 – 0.17	**<0.001**
FlankerCongruency1 * BaseLanguage1	-0.01	-0.03 – 0.02	0.553
Block1 * BaseLanguage1	-0.00	-0.05 – 0.04	0.899
(SentenceType1 * FlankerCongruency1) * Block1	0.01	-0.03 – 0.06	0.572
SentenceType1 * FlankerCongruency1 * BaseLanguage1	-0.00	-0.05 – 0.05	0.915
SentenceType1 * Block1 * BaseLanguage1	0.03	-0.02 – 0.08	0.216
FlankerCongruency1 * Block1 * BaseLanguage1	0.01	-0.03 – 0.06	0.579
SentenceType1 * FlankerCongruency1 * Block1 * BaseLanguage1	0.00	-0.09 – 0.10	0.932
Observations	8264		

Table S4

Complex Model Predicting Logged Reading Time of Critical Three-Word Region

lmer(log(CriticalRegionReadingTime) ~ SentenceType * FlankerCongruency * Block * BaseLanguage + CenteredCriticalRegionLength + (SentenceType*FlankerCongruency*Block|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.01	6.95 – 7.06	**<0.001**

SentenceType1	-0.03	-0.05 – -0.01	**0.002**
FlankerCongruency1	0.00	-0.01 – 0.01	0.522
Block1	-0.21	-0.24 – -0.17	**<0.001**
BaseLanguage1	0.18	0.08 – 0.28	**<0.001**
CenteredCriticalRegionLength	0.01	0.01 – 0.01	**<0.001**
SentenceType1 * FlankerCongruency1	-0.00	-0.02 – 0.02	0.942
SentenceType1 * Block1	-0.00	-0.03 – 0.02	0.903
FlankerCongruency1 * Block1	-0.02	-0.05 – 0.00	0.056
SentenceType1 * BaseLanguage1	0.15	0.11 – 0.18	**<0.001**
FlankerCongruency1 * BaseLanguage1	0.00	-0.02 – 0.03	0.787
Block1 * BaseLanguage1	-0.01	-0.05 – 0.04	0.827
(SentenceType1 * FlankerCongruency1) * Block1	-0.00	-0.04 – 0.04	0.871
SentenceType1 * FlankerCongruency1 * BaseLanguage1	0.00	-0.04 – 0.04	0.896
SentenceType1 * Block1 * BaseLanguage1	0.05	0.01 – 0.09	**0.027**
FlankerCongruency1 * Block1 * BaseLanguage1	-0.00	-0.05 – 0.04	0.962
SentenceType1 * FlankerCongruency1 * Block1 * BaseLanguage1	0.03	-0.05 – 0.10	0.502
Observations	8290		

Table S5

Complex Model Predicting Logged Reading Time of One Word Before Critical Word

lmer(log(`OneBeforeCriticalReadingTime`) ~ SentenceType * FlankerCongruency * Block * BaseLanguage +CenteredOneBeforeWordLength + (SentenceType*FlankerCongruency*Block|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.89	5.84 – 5.95	**<0.001**
SentenceType1	-0.07	-0.09 – -0.04	**<0.001**
FlankerCongruency1	0.00	-0.01 – 0.01	0.809
Block1	-0.18	-0.22 – -0.14	**<0.001**
BaseLanguage1	0.19	0.08 – 0.30	**0.001**
CenteredOneBeforeWordLength	0.01	0.01 – 0.02	**<0.001**
SentenceType1 * FlankerCongruency1	-0.02	-0.04 – 0.01	0.205
SentenceType1 * Block1	-0.03	-0.06 – -0.00	**0.043**
FlankerCongruency1 * Block1	-0.02	-0.04 – 0.01	0.206
SentenceType1 * BaseLanguage1	0.17	0.12 – 0.21	**<0.001**
FlankerCongruency1 * BaseLanguage1	0.01	-0.02 – 0.03	0.557
Block1 * BaseLanguage1	0.00	-0.05 – 0.05	0.992
(SentenceType1 * FlankerCongruency1) * Block1	-0.01	-0.06 – 0.04	0.706
SentenceType1 * FlankerCongruency1 * BaseLanguage1	0.02	-0.03 – 0.07	0.376
SentenceType1 * Block1 * BaseLanguage1	0.07	0.02 – 0.13	**0.007**
FlankerCongruency1 * Block1 * BaseLanguage1	-0.01	-0.06 – 0.04	0.642

SentenceType1 * FlankerCongruency1 * Block1 * BaseLanguage1	0.05	-0.05 – 0.15	0.339
Observations	8339		

Table S6

Complex Model Predicting Logged Reading Time of Word Two Before Critical Word

lmer(log(`TwoBeforeCriticalReadingTime`) ~ SentenceType * FlankerCongruency * Block * BaseLanguage +CenteredTwoBeforeWordLength + (SentenceType*FlankerCongruency*Block|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.90	5.83 – 5.96	**<0.001**
SentenceType1	-0.07	-0.10 – -0.03	**<0.001**
FlankerCongruency1	-0.01	-0.02 – 0.01	0.219
Block1	-0.23	-0.28 – -0.18	**<0.001**
BaseLanguage1	0.20	0.07 – 0.32	**0.002**
CenteredTwoBeforeWordLength	0.01	0.01 – 0.02	**<0.001**
SentenceType1 * FlankerCongruency1	-0.01	-0.04 – 0.02	0.586
SentenceType1 * Block1	0.02	-0.02 – 0.06	0.340
FlankerCongruency1 * Block1	-0.01	-0.04 – 0.02	0.513
SentenceType1 * BaseLanguage1	0.15	0.08 – 0.22	**<0.001**
FlankerCongruency1 * BaseLanguage1	0.01	-0.02 – 0.04	0.518
Block1 * BaseLanguage1	-0.02	-0.08 – 0.04	0.450
(SentenceType1 * FlankerCongruency1) * Block1	-0.00	-0.06 – 0.06	0.930

SentenceType1 * FlankerCongruency1 * BaseLanguage1	0.02	-0.05 – 0.08	0.635
SentenceType1 * Block1 * BaseLanguage1	0.10	0.03 – 0.16	**0.004**
FlankerCongruency1 * Block1 * BaseLanguage1	0.02	-0.04 – 0.08	0.480
SentenceType1 * FlankerCongruency1 * Block1 * BaseLanguage1	0.01	-0.11 – 0.13	0.877
Observations	8323		

Table S7

Determiner-Noun Stimuli: Predicting Logged Critical Word Reading Time

lmer(log(`CriticalWordReadingTime`) ~ SentenceType * DeterminerGender * BaseLanguage + CenteredCriticalWordLength + (SentenceType*DeterminerGender|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.88	5.82 – 5.95	**<0.001**
SentenceType1	-0.04	-0.07 – -0.01	**0.006**
DeterminerGender1	-0.01	-0.09 – 0.07	0.872
BaseLanguage1	0.19	0.08 – 0.30	**0.001**
CenteredCriticalWordLength	0.01	0.00 – 0.02	**0.023**
SentenceType1 * DeterminerGender1	-0.02	-0.06 – 0.02	0.237
SentenceType1 * BaseLanguage1	0.12	0.07 – 0.17	**<0.001**
DeterminerGender1 * BaseLanguage1	-0.02	-0.05 – 0.02	0.295
SentenceType1 * DeterminerGender1 * BaseLanguage1	0.02	-0.05 – 0.09	0.610
Observations	4186		

Table S8

Determiner-Noun Stimuli: Predicting Logged Reading Time of First Word After Critical

lmer(log(`FirstAfterCriticalReadingTime`) ~ SentenceType * DeterminerGender * BaseLanguage + CenteredFirstAfterWordLength + (SentenceType*DeterminerGender|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.88	5.82 – 5.93	**<0.001**
SentenceType1	-0.07	-0.10 – -0.05	**<0.001**
DeterminerGender1	0.01	-0.06 – 0.08	0.805
BaseLanguage1	0.17	0.07 – 0.26	**<0.001**
CenteredFirstAfterWordLength	0.00	-0.01 – 0.02	0.334
SentenceType1 * DeterminerGender1	-0.00	-0.04 – 0.03	0.814
SentenceType1 * BaseLanguage1	0.10	0.05 – 0.14	**<0.001**
DeterminerGender1 * BaseLanguage1	-0.00	-0.03 – 0.03	0.899
SentenceType1 * DeterminerGender1 * BaseLanguage1	-0.01	-0.08 – 0.05	0.674
Observations	4214		

Table S9

Determiner-Noun Stimuli: Predicting Logged Reading Time of Second Word After Critical

lmer(log(`SecondAfterCriticalReadingTime`) ~ SentenceType * DeterminerGender * BaseLanguage + CenteredSecondAfterWordLength + (SentenceType*DeterminerGender|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.81	5.75 – 5.86	**<0.001**
SentenceType1	0.01	-0.01 – 0.03	0.411

DeterminerGender1	0.03	-0.05 – 0.10	0.464
BaseLanguage1	0.15	0.07 – 0.23	**<0.001**
CenteredSecondAfterWordLength	0.02	0.01 – 0.03	**<0.001**
SentenceType1 * DeterminerGender1	-0.03	-0.06 – -0.00	**0.039**
SentenceType1 * BaseLanguage1	0.12	0.08 – 0.16	**<0.001**
DeterminerGender1 * BaseLanguage1	0.00	-0.03 – 0.03	0.891
SentenceType1 * DeterminerGender1 * BaseLanguage1	-0.07	-0.13 – -0.01	**0.021**
Observations	4142		

Table S10

Determiner-Noun Stimuli: Predicting Logged Reading Time of Critical Three-Word Region

lmer(log(`CriticalRegionReadingTime`) ~ SentenceType * DeterminerGender * BaseLanguage + CenteredCriticalRegionLength + (SentenceType*DeterminerGender|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.99	6.93 – 7.05	**<0.001**
SentenceType1	-0.04	-0.06 – -0.01	**0.002**
DeterminerGender1	0.00	-0.08 – 0.08	0.999
BaseLanguage1	0.18	0.08 – 0.27	**<0.001**
CenteredCriticalRegionLength	0.01	0.00 – 0.01	**<0.001**
SentenceType1 * DeterminerGender1	-0.01	-0.04 – 0.02	0.465
SentenceType1 * BaseLanguage1	0.12	0.08 – 0.16	**<0.001**
DeterminerGender1 * BaseLanguage1	-0.01	-0.03 – 0.02	0.601
SentenceType1 * DeterminerGender1 * BaseLanguage1	-0.01	-0.06 – 0.04	0.767
Observations	4150		

Table S11

Verb Stimuli: Predicting Logged Critical Auxiliary + Participle Region Reading Time

lmer(log(`CriticalAuxiliaryParticipleReadingTime`) ~ SentenceType * VerbTense* SwitchLocation * BaseLanguage + CenteredVerbPhraseLength + (SentenceType*VerbTense*SwitchLocation|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.63	6.56 – 6.70	**<0.001**
SentenceType1	-0.04	-0.07 – -0.02	**0.001**
VerbTense1	0.01	-0.06 – 0.08	0.798
SwitchLocation1	-0.07	-0.13 – -0.00	**0.049**
BaseLanguage1	0.22	0.10 – 0.34	**<0.001**
CenteredVerbPhraseLength	-0.01	-0.02 – 0.01	0.342
SentenceType1 * VerbTense1	0.03	0.00 – 0.06	**0.049**
SentenceType1 * SwitchLocation1	-0.06	-0.10 – -0.02	**0.001**
VerbTense1 * SwitchLocation1	0.12	-0.01 – 0.25	0.073
SentenceType1 * BaseLanguage1	0.21	0.16 – 0.26	**<0.001**
VerbTense1 * BaseLanguage1	0.05	0.01 – 0.08	**0.009**
SwitchLocation1 * BaseLanguage1	0.01	-0.03 – 0.05	0.654
SentenceType1 * VerbTense1 * SwitchLocation1	-0.03	-0.10 – 0.04	0.444
SentenceType1 * VerbTense1 * BaseLanguage1	0.06	-0.02 – 0.13	0.131
SentenceType1 * SwitchLocation1 * BaseLanguage1	0.05	-0.02 – 0.12	0.169
VerbTense1 * SwitchLocation1 * BaseLanguage1	0.03	-0.04 – 0.09	0.442

SentenceType1 * VerbTense1 * SwitchLocation1 * BaseLanguage1	-0.03	-0.15 – 0.09	0.647
Observations	4153		

Table S12

Verb Stimuli: Predicting Logged Reading Time for First Word After Critical Auxiliary + Participle Region

lmer(log(`FirstAfterAuxiliaryParticipleReadingTime`) ~ SentenceType * VerbTense* SwitchLocation * BaseLanguage + CenteredFirstAfterWordLength + (SentenceType*VerbTense*SwitchLocation|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.87	5.81 – 5.92	**<0.001**
SentenceType1	-0.04	-0.06 – -0.02	**0.001**
VerbTense1	0.01	-0.05 – 0.07	0.795
SwitchLocation1	-0.05	-0.11 – 0.01	0.126
BaseLanguage1	0.14	0.05 – 0.24	**0.003**
CenteredFirstAfterWordLength	-0.00	-0.02 – 0.01	0.649
SentenceType1 * VerbTense1	0.04	0.01 – 0.07	**0.011**
SentenceType1 * SwitchLocation1	-0.10	-0.14 – -0.07	**<0.001**
VerbTense1 * SwitchLocation1	0.08	-0.04 – 0.20	0.199
SentenceType1 * BaseLanguage1	0.11	0.07 – 0.16	**<0.001**
VerbTense1 * BaseLanguage1	0.02	-0.01 – 0.06	0.206
SwitchLocation1 * BaseLanguage1	0.03	-0.02 – 0.07	0.231
SentenceType1 * VerbTense1 * SwitchLocation1	0.05	-0.01 – 0.12	0.086
SentenceType1 * VerbTense1 * BaseLanguage1	0.00	-0.06 – 0.07	0.902

SentenceType1 * SwitchLocation1 * BaseLanguage1	-0.01	-0.08 – 0.07	0.822
VerbTense1 * SwitchLocation1 * BaseLanguage1	0.04	-0.03 – 0.11	0.294
SentenceType1 * VerbTense1 * SwitchLocation1 * BaseLanguage1	-0.08	-0.20 – 0.05	0.235
Observations	4130		

Table S13

Verb Stimuli: Predicting Logged Reading Time for Second Word After Critical Auxiliary + Participle Region

lmer(log(`SecondAfterAuxiliaryParticipleReadingTime `) ~ SentenceType * VerbTense* SwitchLocation * BaseLanguage + CenteredSecondAfterWordLength + (SentenceType*VerbTense*SwitchLocation|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.86	5.80 – 5.93	**<0.001**
SentenceType1	0.03	0.00 – 0.05	**0.032**
VerbTense1	-0.01	-0.08 – 0.05	0.700
SwitchLocation1	-0.09	-0.15 – -0.02	**0.014**
BaseLanguage1	0.15	0.04 – 0.26	**0.005**
CenteredSecondAfterWordLength	0.02	0.01 – 0.03	**<0.001**
SentenceType1 * VerbTense1	0.02	-0.02 – 0.05	0.381
SentenceType1 * SwitchLocation1	-0.03	-0.06 – 0.00	0.093
VerbTense1 * SwitchLocation1	0.07	-0.07 – 0.21	0.316
SentenceType1 * BaseLanguage1	0.17	0.12 – 0.22	**<0.001**
VerbTense1 * BaseLanguage1	0.02	-0.02 – 0.05	0.304
SwitchLocation1 * BaseLanguage1	-0.03	-0.07 – 0.01	0.125
SentenceType1 * VerbTense1 * SwitchLocation1	0.04	-0.03 – 0.10	0.261

SentenceType1 * VerbTense1 * BaseLanguage1	0.03	-0.04 – 0.10	0.392
SentenceType1 * SwitchLocation1 * BaseLanguage1	0.06	-0.01 – 0.12	0.093
VerbTense1 * SwitchLocation1 * BaseLanguage1	0.06	-0.01 – 0.13	0.079
SentenceType1 * VerbTense1 * SwitchLocation1 * BaseLanguage1	0.02	-0.11 – 0.15	0.786
Observations	4137		

Table S14

Verb Stimuli: Predicting Logged Reading Time for Critical Auxiliary + Participle Region through Two Words After

lmer(log(`CriticalRegionReadingTime`) ~ SentenceType * VerbTense* SwitchLocation * BaseLanguage + CenteredCriticalRegionLength + (SentenceType*VerbTense*SwitchLocation|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.32	7.26 – 7.39	**<0.001**
SentenceType1	-0.03	-0.05 – -0.01	**0.011**
VerbTense1	-0.01	-0.08 – 0.06	0.822
SwitchLocation1	-0.06	-0.13 – 0.01	0.075
BaseLanguage1	0.19	0.07 – 0.30	**0.001**
CenteredCriticalRegionLength	0.01	0.00 – 0.01	**0.001**
SentenceType1 * VerbTense1	0.03	0.01 – 0.06	**0.011**
SentenceType1 * SwitchLocation1	-0.06	-0.09 – -0.03	**<0.001**
VerbTense1 * SwitchLocation1	0.06	-0.07 – 0.20	0.367
SentenceType1 * BaseLanguage1	0.17	0.13 – 0.22	**<0.001**
VerbTense1 * BaseLanguage1	0.02	-0.01 – 0.04	0.256
SwitchLocation1 * BaseLanguage1	0.00	-0.03 – 0.04	0.857

SentenceType1 * VerbTense1 * SwitchLocation1	0.04	-0.01 – 0.10	0.125
SentenceType1 * VerbTense1 * BaseLanguage1	0.00	-0.05 – 0.06	0.864
SentenceType1 * SwitchLocation1 * BaseLanguage1	0.03	-0.02 – 0.09	0.254
VerbTense1 * SwitchLocation1 * BaseLanguage1	0.04	-0.01 – 0.09	0.131
SentenceType1 * VerbTense1 * SwitchLocation1 * BaseLanguage1	-0.02	-0.13 – 0.09	0.735
Observations	4167		

Table S15

Noun vs. Verb Stimuli: Predicting Logged Critical Word Reading Time

lmer(log(`CriticalWordReadingTime`) ~ SentenceType * NounVerbType * BaseLanguage + CenteredCriticalWordLength + (SentenceType*NounVerbType|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.88	5.81 – 5.94	**<0.001**
SentenceType1	-0.05	-0.08 – -0.02	**0.001**
NounVerbType1	-0.03	-0.10 – 0.03	0.324
BaseLanguage1	0.21	0.09 – 0.33	**0.001**
CenteredCriticalWordLength	0.01	0.00 – 0.02	**0.013**
SentenceType1 * NounVerbType1	-0.02	-0.06 – 0.02	0.285
SentenceType1 * BaseLanguage1	0.17	0.12 – 0.23	**<0.001**
NounVerbType1 * BaseLanguage1	0.05	0.01 – 0.09	**0.021**
SentenceType1 * NounVerbType1 * BaseLanguage1	0.10	0.03 – 0.17	**0.003**
Observations	6267		

Note. Verb switches in this model only include participle location items.

Table S16

Noun vs. Verb Stimuli: Predicting Logged Reading Time for First Word After Critical

lmer(log(`FirstAfterCriticalReadingTime`) ~ SentenceType * NounVerbType * BaseLanguage + CenteredFirstAfterWordLength + (SentenceType*NounVerbType|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.86	5.81 – 5.92	**<0.001**
SentenceType1	-0.08	-0.11 – -0.06	**<0.001**
NounVerbType1	-0.02	-0.08 – 0.03	0.391
BaseLanguage1	0.17	0.07 – 0.27	**0.001**
CenteredFirstAfterWordLength	0.01	-0.00 – 0.01	0.236
SentenceType1 * NounVerbType1	-0.02	-0.06 – 0.01	0.198
SentenceType1 * BaseLanguage1	0.11	0.06 – 0.15	**<0.001**
NounVerbType1 * BaseLanguage1	0.01	-0.03 – 0.04	0.765
SentenceType1 * NounVerbType1 * BaseLanguage1	0.02	-0.04 – 0.08	0.546
Observations	6299		

Note. Verb switches in this model only include participle location items.

Table S17

Noun vs. Verb Stimuli: Predicting Logged Reading Time for Second Word After Critical

lmer(log(`SecondAfterCriticalReadingTime`) ~ SentenceType * NounVerbType * BaseLanguage + CenteredSecondAfterWordLength + (SentenceType*NounVerbType|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	5.80	5.75 – 5.86	**<0.001**
SentenceType1	0.01	-0.01 – 0.03	0.575

NounVerbType1	-0.06	-0.12 – 0.01	0.086
BaseLanguage1	0.14	0.05 – 0.24	**0.003**
CenteredSecondAfterWordLength	0.02	0.01 – 0.02	**<0.001**
SentenceType1 * NounVerbType1	-0.01	-0.03 – 0.02	0.654
SentenceType1 * BaseLanguage1	0.15	0.11 – 0.19	**<0.001**
NounVerbType1 * BaseLanguage1	-0.01	-0.05 – 0.03	0.504
SentenceType1 * NounVerbType1 * BaseLanguage1	0.06	0.01 – 0.11	**0.026**
Observations	6210		

Note. Verb switches in this model only include participle location items.

Table S18

Noun vs. Verb Stimuli: Predicting Logged Reading Time of Critical Three-Word Region

lmer(log(`CriticalRegionReadingTime`) ~ SentenceType * NounVerbType * BaseLanguage + CenteredCriticalRegionLength + (SentenceType*NounVerbType|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.98	6.92 – 7.04	**<0.001**
SentenceType1	-0.04	-0.06 – -0.02	**<0.001**
NounVerbType1	-0.06	-0.13 – 0.00	0.056
BaseLanguage1	0.18	0.08 – 0.29	**0.001**
CenteredCriticalRegionLength	0.01	0.01 – 0.01	**<0.001**
SentenceType1 * NounVerbType1	-0.01	-0.04 – 0.02	0.575
SentenceType1 * BaseLanguage1	0.16	0.11 – 0.20	**<0.001**
NounVerbType1 * BaseLanguage1	0.02	-0.02 – 0.05	0.381
SentenceType1 * NounVerbType1 * BaseLanguage1	0.07	0.02 – 0.12	**0.004**

Observations	6236

Note. Verb switches in this model only include participle location items.

Table S19

Code-switch Type as a Conflict Adaptation Modulator: Predicting Logged Reading Time of Critical Three-Word Region

lmer(log(CriticalRegionReadingTime) ~ SentenceType * CodeSwitchType*FlankerCongruency + (SentenceType+FlankerCongruency+CodeSwitchType||Participant) + (SentenceType||Item)			
Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.97	6.91 – 7.03	**<0.001**
SentenceType1	-0.05	-0.07 – -0.02	**0.001**
CodeSwitchType [perfect]	0.04	-0.03 – 0.10	0.248
CodeSwitchType [progressive]	0.04	-0.02 – 0.10	0.169
FlankerCongruency1	0.00	-0.01 – 0.01	0.940
SentenceType1 * CodeSwitchType [perfect]	-0.02	-0.04 – 0.01	0.162
SentenceType1 * CodeSwitchType [progressive]	0.03	0.00 – 0.05	**0.025**
SentenceType1 * FlankerCongruency1	-0.00	-0.03 – 0.02	0.789
CodeSwitchType [perfect] * FlankerCongruency1	0.01	-0.01 – 0.03	0.319
CodeSwitchType [progressive] * FlankerCongruency1	-0.00	-0.02 – 0.02	0.811
SentenceType1 * CodeSwitchType [perfect] * FlankerCongruency1	0.01	-0.04 – 0.05	0.820
SentenceType1 * CodeSwitchType [progressive] * FlankerCongruency1	-0.01	-0.05 – 0.04	0.809

Observations	8290

Note. CodeSwitchType was dummy coded in this model with Determiner-Noun switches coded as the reference group.

Table S20

Language History as a Conflict Adaptation Modulator: Predicting Logged Reading Time of Critical Three-Word Region

lmer(log(CriticalRegionReadingTime) ~ FlankerCongruency * SentenceType*CodeSwitchExperience*LanguageExposureEntropy+ (SentenceType*FlankerCongruency|Participant) + (SentenceType|Item), control = lmerControl(optimizer = "bobyqa")			
Predictors	*Estimates*	*CI*	*p*
(Intercept)	7.02	6.31 – 7.73	**<0.001**
FlankerCongruency1	0.08	-0.06 – 0.22	0.273
SentenceType1	-0.11	-0.43 – 0.21	0.493
CodeSwitchExperience	0.05	-0.19 – 0.30	0.673
LanguageExposureEntropy	0.04	-0.73 – 0.80	0.919
FlankerCongruency1 * SentenceType1	-0.14	-0.40 – 0.11	0.279
FlankerCongruency1 * CodeSwitchExperience	-0.03	-0.08 – 0.02	0.296
SentenceType1 * CodeSwitchExperience	0.04	-0.07 – 0.15	0.437
FlankerCongruency1 * LanguageExposureEntropy	-0.06	-0.21 – 0.10	0.476
SentenceType1 * LanguageExposureEntropy	0.04	-0.31 – 0.38	0.838
CodeSwitchExperience * LanguageExposureEntropy	-0.08	-0.34 – 0.18	0.534
(FlankerCongruency1 * SentenceType1) * CodeSwitchExperience	0.02	-0.07 – 0.11	0.628

(FlankerCongruency1 * SentenceType1) * LanguageExposureEntropy	0.12	-0.15 – 0.40	0.380
(FlankerCongruency1 * CodeSwitchExperience) * LanguageExposureEntropy	0.02	-0.03 – 0.07	0.484
(SentenceType1 * CodeSwitchExperience) * LanguageExposureEntropy	-0.03	-0.15 – 0.08	0.564
(FlankerCongruency1 * SentenceType1 * CodeSwitchExperience) * LanguageExposureEntropy	-0.01	-0.11 – 0.08	0.775
Observations	8141		

Table S21

Switch Location as a Conflict Adaptation Modulator: Predicting Logged Reading Time of Critical Three-Word Region

lmer(log(CriticalRegionReadingTime) ~ SentenceType * FlankerCongruency*CriticalWordLocation + (SentenceType*FlankerCongruency*CriticalWordLocation|Participant) + (SentenceType|Item)

Predictors	*Estimates*	*CI*	*p*
(Intercept)	6.88	6.78 – 6.98	**<0.001**
SentenceType1	-0.04	-0.09 – 0.02	0.227
FlankerCongruency1	-0.01	-0.05 – 0.03	0.777
CriticalWordLocation	0.02	0.00 – 0.03	**0.014**
SentenceType1 * FlankerCongruency1	0.01	-0.07 – 0.09	0.819
SentenceType1 * CriticalWordLocation	-0.00	-0.01 – 0.01	0.736
FlankerCongruency1 * CriticalWordLocation	0.00	-0.00 – 0.01	0.669

(SentenceType1 * FlankerCongruency1) * CriticalWordLocation	-0.00	-0.01 – 0.01	0.751
Observations	8290		

Bibliography

Adler, R., Novick, J., & Huang, Y. (2018). Context, conflict, and the time course of interpreting irony [Conference abstract]. In: *The 31st Annual CUNY Sentence Processing Conference Program*, Davis, CA, United States. Abstract no 107.

Adler, R. M., Valdés Kroff, J. R., & Novick, J. M. (2020). Does integrating a code-switch during comprehension engage cognitive control?. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(4), 741-759.

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477–492. https://doi.org/10.3758/BF03200936

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*(3), 247-264.

Barton, K. (2020). MuMIn: Multi-Model Inference. R package version 1.43.17.

https://CRAN.R-project.org/package=MuMIn

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language*, *95*, 173–189. https://doi.org/10.1016/j.jml.2017.04.002

Beatty-Martínez, A. L., & Dussias, P. E. (2019). Revisiting masculine and feminine grammatical gender in Spanish: Linguistic, psycholinguistic, and neurolinguistic evidence. *Frontiers in Psychology, 10*, 751.

Beatty-Martínez, A. L., Navarro-Torres, C. A., & Dussias, P. E. (2020). Codeswitching: A bilingual toolkit for opportunistic speech planning. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01699

Blokzijl, J., Deuchar, M., & Parafita Couto, M. C. (2017). Determiner asymmetry in mixed nominal constructions: The role of grammatical factors in data from Miami and Nicaragua. *Languages, 2*(4), 20. https://doi.org/10.3390/languages2040020

Bosma, E., & Pablos, L. (2020). Switching direction modulates the engagement of cognitive control in bilingual reading comprehension: An ERP study. *Journal of Neurolinguistics*, *55*. https://doi.org/10.1016/j.jneuroling.2020.100894

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*(3), 624-652. DOI: 10.1037/0033-295x.108.3.624

Bultena, S., Dijkstra, T., & Van Hell, J. G. (2015). Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition*, *18*(3), 453–469. https://doi.org/10.1017/S1366728914000145

DeLuca, V., Segaert, K., Mazaheri, A., & Krott, A. (2020). Understanding bilingual brain function and structure changes? U Bet! A unified bilingual experience trajectory model. *Journal of Neurolinguistics, 56*, 100930. https://doi.org/10.1016/j.jneuroling.2020.100930

Dussias, P. E., Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Gerfen, C. (2014). Looking into comprehension of Spanish-English code-switched sentences: Evidence from eye movements. *Peaches and plums*, 335-351.

English Language Institute. (2001). *MELICET—GCVR user's manual*. Ann Arbor, MI: English Language Institute, The University of Michigan.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*(2), 178-210.

Fricke, M., & Kootstra, G. J. (2016). Primed codeswitching in spontaneous bilingual dialogue. *Journal of Memory and Language, 91*, 181–201. https://doi.org/10.1016/j.jml.2016.04.003

Gosselin, L., & Sabourin, L. (2021). Lexical-semantic processing costs are not inherent to intra-sentential code-switching: The role of switching habits. *Neuropsychologia*, 107922.

Gouvea, A. C., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes, 25*(2), 149-188.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General, 121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480

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

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution, 7*(4), 493-498.

Gullifer, J. W., & Titone, D. (2020). Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and Cognition, 23*(2), 283-294.

Guzzardo Tamargo, R. E., Valdés Kroff, J. R., & Dussias, P. E. (2016). Examining the relationship between comprehension and production processes in code-switched language. *Journal of Memory and Language*, *89*, 138–161. https://doi.org/10.1016/j.jml.2015.12.002

Hsu, N. S., & Novick, J. M. (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science, 27*(4), 572-582.

Hsu, N. S., Kuchinsky, S. E., & Novick, J. M. (2021). Direct impact of cognitive control on sentence processing and comprehension. *Language, Cognition and Neuroscience, 36*(2), 211-239.

Johns, M. A., Valdés Kroff, J. R., & Dussias, P. E. (2019). Mixing things up: How blocking and mixing affect the processing of codemixed sentences. *International Journal of Bilingualism, 23*(2), 584-611.

Kaan, E., Kheder, S., Kreidler, A., Tomić, A., & Valdés Kroff, J. R. (2020). Processing code-switches in the presence of others: An ERP study. *Frontiers in Psychology, 11,* 1288.

Kan, I. P., Teubner-Rhodes, S., Drummey, A. B., Nutile, L., Krupa, L., & Novick, J. M. (2013). To adapt or not to adapt: The question of domain-general cognitive control. *Cognition*, *129*(3), 637–651. https://doi.org/10.1016/j.cognition.2013.09.001

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(1), 1-26.

Lenth, R., V. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.6.1. https://CRAN.R-project.org/package=emmeans

Litcofsky, K. A., & Van Hell, J. G. (2017). Switching direction affects switching costs: Behavioral, ERP and time-frequency analyses of intra-sentential codeswitching. *Neuropsychologia, 97*, 112-139.

Lüdecke D (2020). sjstats: Statistical Functions for Regression Models (Version 0.18.0). doi:10.5281/zenodo.1284472

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology, 4*, 226.

Ministry of Education, Culture, and Sport of Spain. (2006). *Diploma de español como lengua extranjera examination*. Retrieved from http://dele.org

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language, 80*(2), 188-207.

Piccinini, P., & Arvaniti, A. (2015). Voice onset time in Spanish–English spontaneous code-switching. *Journal of Phonetics, 52*, 121–137. https://doi.org/10.1016/j.wocn.2015.07.004

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics*, *18*(7–8), 581–618. https://doi.org/10.1515/

Salig, L. K., Valdés Kroff, J. R., Slevc, L. R., & Novick, J. M. (2021). Moving from bilingual traits to states: Understanding cognition and language processing through moment-to-moment variation. *Neurobiology of Language*, 1-62.

Thothathiri, M., Asaro, C. T., Hsu, N. S., & Novick, J. M. (2018). Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension. *Cognition*, *178*, 162–177. https://doi.org/10.1016/j.cognition.2018.05.014

Tomić, A., & Valdés Kroff, J. R. (2021). Expecting the unexpected: Code-switching as a facilitatory cue in online sentence processing. *Bilingualism: Language and Cognition*, 1-12. https://doi.org/10.1017/S1366728921000237

Ullsperger, M., Bylsma, L. M., & Botvinick, M. M. (2005). The conflict adaptation effect: It's not just priming. *Cognitive, Affective, & Behavioral Neuroscience*, *5*(4), 467–472. https://doi.org/10.3758/CABN.5.4.467

Valdés Kroff, J. R. (2016). Mixed NPs in Spanish-English bilingual speech: using a corpus-based approach to inform models of sentence processing. In R. E. Guzzardo, C. M. Mazak, & M. C. Parafita (Eds.), *Spanish-English Codeswitching in the Caribbean and the US.* (pp. 281-300) John Benjamins.

Valdés Kroff, J. R., Dussias, P. E., Gerfen, C., Perrotti, L., & Bajo, M. T. (2017). Experience with code-switching modulates the use of grammatical gender during sentence processing. *Linguistic Approaches to Bilingualism, 7*(2), 163-198.

Wu, Y. J., & Thierry, G. (2013). Fast modulation of executive function by language context in bilinguals. *Journal of Neuroscience*, *33*(33), 13533–13537. https://doi.org/10.1523/JNEUROSCI.4760-12.2013

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832