Code-switching aids the prediction of the unexpected

Aleksandra Tomic & Jorge R. Valdés Kroff

## Introduction

Code-switching (CS), or the use of several codes in the same conversation (Gardner-Chloros, 2009) is one of the hallmarks of bilingualism. Decades of sociolinguistic research on the rules governing the use of code-switching, its functions and meanings, found that code-switching serves a variety of sociopragmatic purposes (Gumperz, 1982; Myers-Scotton, 1993). Recently, psycholinguists have shown increasing interest in understanding the neural and cognitive processes that support the integration of code-switched speech in comprehension. This interest has so far mostly focused on the processing costs of integrating code-switches (Litcofsky & Van Hell, 2017; Olson, 2017), as well as potential attenuation of these costs under certain linguistic and extra-linguistic contexts (Fricke, Kroll, & Dussias, 2016; Guzzardo Tamargo, Valdés Kroff, & Dussias, 2016; Valdés Kroff et al., 2017). While this line of research provides important insights into bilingual sentence processing, it does not take into account the sociopragmatic motivations for code-switching (Myers-Scotton & Jake, 2001) and the potential benefits that code-switching may provide in processing subsequent linguistic information. Despite the robust processing costs associated with it, code-switched speech makes up a relatively large portion of bilingual discourse--around 20% (Beatty-Martinez & Dussias, 2017). This ubiquitous bilingual practice (among certain communities, Poplack & Meechan, 1998; Valdés Kroff, Guzzardo Tamargo, & Dussias, 2018) suggests that code-switching affords processing benefits which override these purported processing costs. The study reported here takes this approach and experimentally tests one such processing benefit of CS: alerting to and aiding prediction of lower frequency, unexpected information.

## Background

Some of the most cited sociopragmatic functions of code-switching are identity expression (Velasquez, 2010), situational marking, (re)negotiating social relations (Myers-Scotton, 1993), face-saving (Bentahila, 1983), discourse organization (Auer, 1988), emphasis (Gumperz, 1982), and introducing indirect speech (Albirini, 2011). More recently, Myslín and Levy (2015) proposed that code-switching serves an information-distribution function for organizing discourse. The authors compiled a corpus of Czech-English bilingual speech. They calculated the predictability of words as the percentage of participants who listened to the corpus and correctly guessed which word occupies the gap after a portion of discourse. They found that item predictability predicts the code-switching behavior above and beyond 10 other control factors, such that bilinguals code-switch from their more frequently used language to their less used, or more salient language, at more unpredictable words. Relatedly, Bentahila (1983) compiled a corpus of 7.5 hrs of informal conversation between balanced Arabic-French bilinguals aged 17 to 40. One of the recurrent sociopragmatic motivations he identified is code-switching to the more marked language -- French -- when speaking about emotional, i.e. information-rich, taboo topics. These observations suggest that code-switching may offer online processing benefits by signaling and thus aiding the prediction of highly informative or more unpredictable portions of upcoming speech.

This potential function of code-switching aligns it with other discourse-organizational markers, such as disfluencies. Disfluencies, or irregularities in fluent speech, such as “uh”, “um”, or pauses, have been found to occur when referring to new vs. given information (Arnold et al., 2000; Barr, 2001). Experimental research showed that monolinguals make use of this distribution regularity in on-line language processing to help them predict unexpected, new (Arnold, Fagnano, & Tanenhaus, 2003; Arnold, Tanenhaus, Altmann, & Fagnano, 2004; Arnold, Kam, Hudson, & Tanenhaus, 2007) or low-frequency words (Bosker, Quené, Sanders, & de Jong, 2014). In these visual world studies (eye-tracking while looking at images and listening to instructions), disfluent instructions to select an image caused participants to start looking at the unexpected item faster, shortly prior to or following the onset of the target word. While CS is not equivalent to disfluent speech, we suggest that CS, despite its observed processing costs, may serve a similar important discourse function to the bilingual comprehenders. We propose to test whether CS can serve a facilitative function for upcoming speech by signaling upcoming lower frequency words.

## Current Study

To test whether Spanish-English bilinguals use code-switching as a facilitative cue in sentence processing, we employed the visual world paradigm with eye-tracking (Tanenhaus et al., 1995). Predictability was operationalized as lexical frequency to simplify the experiment design. Spanish-English bilinguals were presented with two-picture panels, one representing a low-frequency word and the other a high-frequency word. Audio instructions asked participants to select a target image via mouse click. The instructions were either in unilingual Spanish or included an alternational code-switch from Spanish into English. Crucially, the code-switch occurred before the naming of the target object. Analyzing total proportion of fixations to images on the visual scene, if bilinguals indeed interpret a code-switch as a signal to upcoming unexpected words, then looks to lower frequency items should be higher on code-switch trials as compared to unilingual trials *before* the onset of target words. The independent variables of experimental interest are Language of the Instructions with 2 levels: Spanish (L1) and Code-switched (Spanish-English, L1-L2), and Frequency of the Fixated Image: Low and High. The study was thus geared towards bilinguals with Spanish as an L1. Only the L1-L2 CS direction was chosen to simplify the design, since L1-L2 is the attested switch direction used to signal more informative portions of speech (Myslín & Levy, 2015), and is more representative of the code-switching practices of the Spanish-English bilingual community in the U.S. (Moreno, Federmeier, & Kutas, 2002; Valdés Kroff et al., 2018).

### Materials and Methods

### Instructions and audio recordings

We constructed two carrier phrases for the instructions, each with a unilingual Spanish and Spanish-English code-switched variant. Code-switches preceded the name of the target object by three words, one content and two function ones, to avoid any immediate effects of switch costs affecting the results. The code-switch was placed within a noun phrase, after an article and at the noun, which was deemed an acceptable place for a code-switch by three bilingual speakers and by prior work (Beatty-Martínez & Dussias, 2017; Valdés Kroff, 2016). The CS and Spanish carrier phrases and image names were recorded with the help of a balanced Puerto Rican Spanish-English speaker who is also a trained audiologist. Two carrier verbs were chosen for the carrier phrases: *elegir* “choose” and *encontrar* “find” to increase variety and approach ecological validity. The carrier phrases are presented in (1) and (2) below.

1. a. Encuentra el dibujo de un/una/Ø \_\_\_\_\_\_\_\_\_\_

b. Encuentra el drawing of a/an/Ø \_\_\_\_\_\_\_\_\_\_

“Find the drawing of a/an/ Ø \_\_\_\_\_\_\_\_\_\_”

2) a. Elige el dibujo de un/una/ Ø \_\_\_\_\_\_\_\_\_\_

b. Elige el drawing of a/an/ Ø \_\_\_\_\_\_\_\_\_\_

“Select the drawing of a/an/ Ø \_\_\_\_\_\_\_\_\_\_

The picture names were recorded in isolation, with the speaker instructed to imagine the noun coming at the end of a sentence and produced with a declarative intonation. The carrier phrases were recorded in combination with a filler noun, and subsequently cut, to ensure that the intonation and article pronunciation were as natural as possible. The onset of the code-switch was briefly delayed compared to the comparable point in the Spanish-only instructions (mean difference = 22 ms). The comparable time frame from the onset of the code-switch to the target word onset was longer in the Spanish-only instructions (Mean = 955 ms) than in the CS condition (Mean = 833 ms), presumably due to an additional syllable in the Spanish equivalent for “drawing” and the additional syllable in the article for female determiners.

All carrier phrases were scaled to an average intensity of 70 dB, and all nouns were scaled to the average intensity of 66 dB using Praat (Broersma & Weenink, 2018), to ensure volume uniformity among carrier phrases and noun, and a natural volume decline at the end of the sentence. The final versions of carriers and nouns were concatenated without a pause.

### Picture panels

Sixty-two images, 32 representing high-frequency words and 32 representing low frequency words were extracted from the International Picture Naming Project database (Szekely et al., 2003, 2004, 2005; Bates et al., 2003), to form 32 experimental panels with pairs of images. The frequency measures used in this database are log natural transformed frequency counts, taken from the CELEX Lexical database (Baayen, Piepenbrock, & Gulikers, 1995). All experimental pictures, picture names, and frequencies can be found in Appendix A.

Experimental images were chosen such that there is the smallest possible discrepancy in the frequency of the Spanish and English picture name. Descriptive statistics for English and Spanish words are presented in Table 1.

|  |  |  |
| --- | --- | --- |
|  | English | Spanish |
| Overall Mean | 3.145 (1.87) | 3.245 (1.908) |
| High Frequency Mean | 4.892 (.713) | 5.028 (.772) |
| Low Frequency Mean | 1.398 (.542) | 1.48 (.552) |

Table 1: Means and standard deviations in parentheses for all experimental words in English and Spanish and split by frequency.

The average absolute difference between the English and Spanish object name counterparts was 0.444 frequency counts (SD = 0.371, range: 0 – 1.61). A two-tailed paired t-test showed no significant difference between English and Spanish picture name counterparts, ensuring no extraneous factors would affect the results of the study (t[63] = -1.527, p = 0.131, mean of differences = -0.109).

Pictures were paired such that there is the largest possible frequency gap between them. Average difference between the high and low frequency members of a pair was 3.54 frequency counts for CS trials, and 3.6 counts for Spanish trials. Paired t-tests for both English and Spanish trials show that there is a significant frequency difference between the high and low frequency members of the experimental pairs (Spanish counterparts: t(33) = 21.778, p-value < 0.001; English counterparts: t(33) = 33.993, p-value < 0.001). The experimental two-picture panels were also matched for the gender of the Spanish picture name/translation equivalent to prevent participants from using gender of the article as a cue (Valdés Kroff et al., 2017). Out of 32 experimental trials, 15 were feminine gender pairs (or translation equivalents). Due to difficulties in attaining a sufficient number of appropriate pictures from the database, 12 pictures representing English-Spanish cognates were included in the experimental trials. These pictures were paired with each other to control for any possible cognate effects, resulting in 6 cognate-pair experimental trials.

Four experimental lists were created, with one experimental item pair appearing in one of the 4 conditions/versions in one of the lists: Spanish-only, Low Frequency target, CS Low Frequency target, Spanish High Frequency target, or CS High Frequency target. This process resulted in 8 trials per condition within a list and ensured that a participant sees a given experimental item in only one of the 4 conditions. The order of experimental and filler trials was pseudorandomized using nested lists to ensure that there was never more than 3 experimental items in a row. There were 64 filler trials in each list. The filler images were drawn from the same database, and item pairs were similar in terms of frequency. The fillers were the same across lists, but their order and image position were randomized in presentation.

The experiment was programed in Experiment Builder (SR Research, 2011). Images were presented two at a time on white background. The pictures could appear anywhere on the sample array in Figure 1. The position of the images was randomized using a prepared Python script. To make sure that there is no overlap between the looks to the target vs. distractor items, the items were never adjacent.

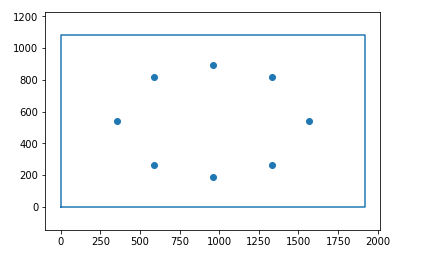


Figure 1. Possible positions for the images in the panel. Two images were not allowed to appear in consecutive positions.

### Participants

Thirty Spanish-English bilingual participants (4 male), age range 18-32 (M = 20.83, SD = 3.53), were recruited at the University of Florida. They were compensated for their participation in course credit or cash. All participants reported having begun learning both English and Spanish before puberty (Spanish age of acquisition [AoA] Mean = 0.67, SD = 2.35, Max = 12; English AoA Mean = 3.67, SD = 2.48, Max = 10).

Participants completed a Language History Questionnaire (LHQ), semantic fluency tests in Spanish and English, and adapted standardized grammar tests from the Michigan English Language Institute College English Test (MELICET) and the Diplomas of Spanish as a Foreign Language (DELE). Proficiency results are summarized in Table X. The tests were administered either prior to or following the main reading experiment. The order of the main experiment and tests was counterbalanced, as well as the language order of the additional tests.

Twenty-eight participants reported code-switching in the LHQ, whereas 2 participants responded with “Not sure”. All participants filled out additional questions on the frequency of use and exposure to code-switching in speaking and writing. The mean response to frequency of code-switching use was 4.1, SD = 0.845, and the mean response to exposure to code-switching in spoken conversation was 3.867, SD = 0.973, with 1 indicating “Never” and 5 “Always”. The results of the other proficiency measures are presented in Table 2. As revealed by the proficiency measures, our sample indicated greater proficiency in English than Spanish, reflecting our participants’ likely status as heritage speakers (REF).

|  |  |  |  |
| --- | --- | --- | --- |
|  | English: M (SD) | Spanish: M (SD) | Difference: Paired 2-tailed T-test |
| LHQ – Speaking | 9.73 (.64) | 8.27 (1.55) | \*\*\* |
| LHQ – Listening | 9.73 (.58) | 9.17 (1.15) | \* |
| LHQ – Writing | 9.53 (1.2) | 7.67 (1.73) | \*\*\* |
| LHQ – Reading | 9.67 (.84) | 8.07 (1.6) | \*\*\* |
| MELICET | DELE | 43.73 (3.36) | 30.97 (6.97) | \*\*\* |
| Semantic fluency |  |  |  |

Table 2: Proficiency profile for participants (n = 30). LHQ values represent self-reported ratings of proficiency from the Language History Questionnaire on a scale of 1 (no proficiency) to 10 (highly proficient). MELICET and DELE scores are calculated out of 50.

### Procedure

The study took place at the Bilingual Sentence Processing Lab at the University of Florida. Participants were greeted, signed an informed consent form, and were given code-switched instructions to listen to audio recordings and click on the correct image. Eye-movements were recorded from the right eye (viewing was binocular) using an SR Research Eyelink 1000 Plus desk-mounted eye-tracker. Participants’ heads were stabilized using a chin rest and were seated approximately 70 cm from a 24-inch LED Benq monitor. Participants completed a 9-point calibration and validation test. Calibration was deemed successful if average error was at or below 0.5 degrees. Participants competed eight practice items before the main experiment. Audio instructions playback was delayed by 200ms after the presentation of the panel. Participants were instructed to use a mouse to click on the correct image.

## Results

### 4.1 Data Analysis

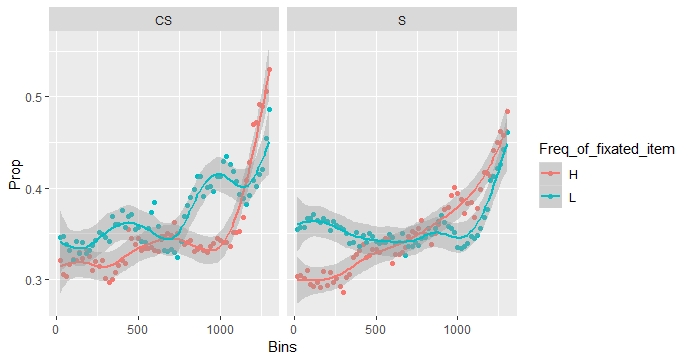
Because of our interest in examining the effects of CS on anticipating upcoming speech, the target time period for eye-movement analysis was 200 ms before and 200 ms after target onset. Planning and launching an eye-movement has been shown to take around 200 ms (Allopenna, Magnuson, & Tanenhaus, 1998; Travis, 1936), so the 200 ms post-target-word region reflects predictive processes which had taken place prior to participants’ ability to process the onset of the target word.

We removed the trials in which participants clicked on the incorrect image prior to analyzing the eye-tracking data. This resulted in 1.35% trial loss, evenly distributed among participants (max. 2 trials per participant), suggesting that participants were performing at ceiling despite relative low frequency of certain items.

The Time variable in the eye-tracking data was binned into 20 ms bins. The dependent variable in the analyses was the Proportion of fixations towards items aggregated by condition and time bins. After calculating the proportions, blinks and in-saccade eye-states were removed from the data. The independent variables were Language Context (Spanish, Code-switched), Frequency of the Fixated Image (High, Low), and Dominance (a continuous variable). The values for the independent variable Frequency of the Fixated Image were assigned to each fixated image regardless of their target/distractor status. Dominance was operationalized as the ratio between the DELE and MELICET scores, with higher ratio meaning more Spanish dominant. Due to an error in procedure, the Dominance data was not available for 1 participant. Their Dominance score was filled with the Dominance average. Dominance and Proportion of Looks were standardized by z-scoring.

### Growth Curve Analysis

We initially performed a Growth Curve Analysis model, with the time variable transformed with orthogonal polynomials of 1st, 2nd, and 3rd order: OT1 – linear term, OT2 – quadratic term, OT3 – cubic term (Mirman, 2014). This analysis would have allowed us to potentially take into account the curvature of the time series data and increasing trends in the dependent variable. Nevertheless, the OT2 and OT3 terms as fixed effects did not improve the model significantly[[1]](#footnote-1), as confirmed by comparing the models with and without OT2 and OT3 terms as fixed effects, using the *anova* function (Fox & Weisberg, 2011). Therefore, we only report the more parsimonious model without these terms as fixed effects. A graphical representation of the data is presented in Figure 1.



Frequency of fixated image

Milliseconds from word onset

Proportion of looks to images

-800

+500

0

0

-800

+500

Figure 1: Y-axis represents the Proportion of looks to High and Low frequency images split by language context (CS = code-switched; S = Spanish) and Frequency of fixated item (H = high, L = Low). X-axis represents the time course of -800 ms (approximate CS onset) to +500 ms from the target word onset (red line). The blue rectangle represents the target time period used in the analysis. (Created using ggplot2 package, Wickham, 2009)

Z-scored proportions of time spent looking at a set of items in a particular condition in each bin was fit to a linear mixed-effects model using the lme4 package (Bates et al., 2015) in R (R Core Team, 2017). The model included Language (contrast coded as Spanish -0.5, Code-switched +0.5), Frequency (contrast coded as High -0.5, Low +0.5), Dominance (continuous), and linear orthogonal polynomial term, OT1, (continuous), as well as their interactions, as predictors. Participant Language, Frequency, and orthogonal polynomial terms OT1, OT2, and OT3 intercepts were included in the model as random effects.[[2]](#footnote-2) Results of the analysis are summarized in the Table 3.

|  |  |  |  |
| --- | --- | --- | --- |
| Fixed effects: | Estimate | Standard Error | t-value |
| (Intercept) | -0.02832 | 0.08300 | -0.341 |
| Language | 0.01940 | 0.05761 | 0.337 |
| Dominance | -0.15906 | 0.08345 | -1.906 |
| Frequency | -0.03351 | 0.12170 | -0.275 |
| ot1 | **0.28034** | **0.09074** | **3.090** |
| Language:Dominance | -0.05961 | 0.05781 | -1.031 |
| Language:Frequency | **0.49200** | **0.04740** | **10.380** |
| Dominance:Frequency | -0.14483 | 0.12224 | -1.185 |
| Language:ot1 | -0.18068 | 0.10538 | -1.714 |
| Dominance:ot1 | -0.09330 | 0.09103 | -1.025 |
| Frequency:ot1 | **0.30105** | **0.10554** | **2.852** |
| Language:Dominance:Frequency | **0.28593** | **0.04722** | **6.055** |
| Language:Dominance:ot1 | 0.12296 | 0.10548 | 1.166 |
| Language:Frequency:ot1 | **0.89128** | **0.21102** | **4.224** |
| Dominance:Frequency:ot1 | 0.01930 | 0.10559 | 0.183 |
| Language:Dominance:Frequency:ot1 | **0.82272** | **0.21112** | **3.897** |

Table 3: Coefficients, standard errors, and t-values for the GCA model reported here. Significant values are bolded. Relevant interactions are underlined.

The main effect of the linear time term main effect was found to be significant, b = 0.28, SE = 0.091,t = 3.09, such that overall looks increased over time. There was a trend towards a main effect for Dominance, b = -0.159, SE = 0.084, t = -1.906, such that more Spanish-dominant participants had overall fewer looks to images. Importantly, several interactions were found to be significant. The Language x Frequency interaction proved significant, b= 0.492, SE = 0.047, t = 10.38, such that Low frequency items were fixated more in the CS condition compared to Spanish condition. We report partial effects means and standard errors for this key interaction in Table 4, as well as the plot of means and SEs in Figure 2. These were produced using the *effects* package (Fox & Weisberg, 2019). Additionally, the interaction of Language, Dominance, and FreqFix was significant, b = 0.285, SE = 0.047, t = 6.055, such that the participants with higher relative Spanish dominance looked at the Low frequency items more in the CS condition compared to Spanish condition. Although dominance was calculated as a continuous measure, we graphically present the results by categorically splitting the participants into two dominance groups (using a median split) for visual presentation (Figure 3).

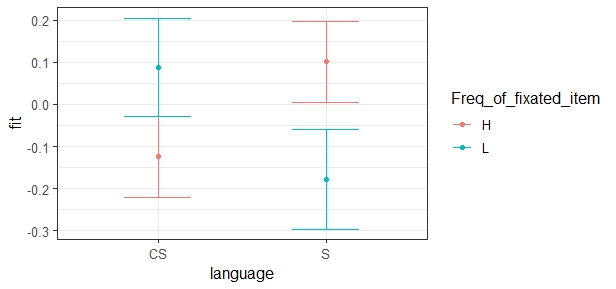
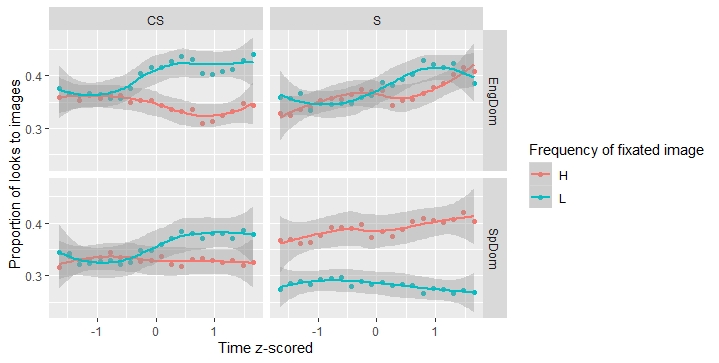


Figure X: Partial effects plot for Language x Frequency of the fixated item interaction, means and standard errors for conditions.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Language | Frequency | fit | se | lower Confidence Limit | upper Confidence Limit |
| CS | H | -0.125 | 0.096 | -0.313 | 0.063 |
| S | H | 0.102 | 0.097 | -0.088 | 0.291 |
| CS | L | 0.087 | 0.117 | -0.142 | 0.316 |
| S | L | -0.178 | 0.119 | -0.411 | 0.054 |

Table X: Partial effects table for the Language x Frequency of the fixated item interaction: condition means, standard errors, lower and upper Confidence Limits.



-200

+200

+200

0

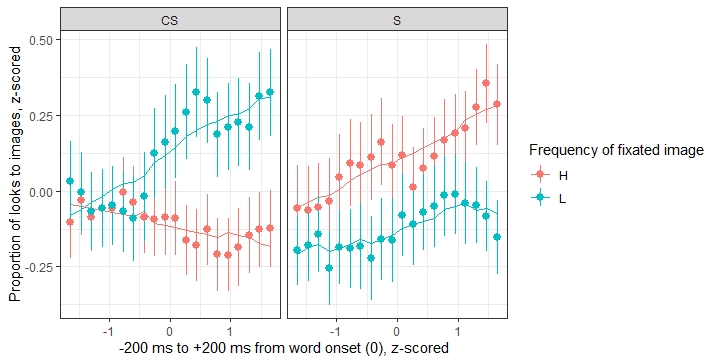
0

-200

Milliseconds from word onset

Figure 2: Looks to High and Low Frequency items in CS and Spanish condition split by Dominance. Dominance was transformed into a categorical variable via a median split. CS = code-switched instructions, S = Spanish instructions; H = High Frequency; L = Low Frequency; EngDom = relatively more English-dominant group; SpDom = relatively more Spanish-dominant

The interaction of Language, Frequency, and the linear time term was also found to be significant, b = 0.891, SE = 0.211, t = 4.224 such that the looks towards the Low frequency item increased over time in the CS condition compared to Spanish condition. The interaction of Language, Dominance, FreqFix, and the linear time term was significant as well, b = 0.822, SE = 0.211, t = 3.897, such that the participants with higher relative Spanish dominance looked more over time at the Low frequency item in the CS condition compared to the Spanish condition. Graph X shows the model fit.



Graph X: Reported GCA model fit, with modeled data represented as lines and observed data as points.

## Discussion

Using the visual world paradigm, we investigated whether bilingual listeners interpret a code-switch as signaling upcoming less-expected content as operationalized by lexical frequency. The results from the Growth Curve Analysis point to a global increase of looks to low frequency items in the CS condition in a 400 ms critical time window prior to when participants process the onset of target items.  This result corroborate our hypothesis that code-switching plays a role to signaling upcoming information in on-line language processing. This result is especially revealing given the robustly documented frequency bias in visual world studies, which is a general tendency to look at higher frequency, more familiar items (e.g. Dahan, Magnuson, & Tanenhaus, 2001). Moreover, the results for the interaction of Language, Frequency, and Time further support our predictions, since there is a continuous increase for the looks to the lower frequency items in the CS context compared to the Spanish context as the target onset approaches.

Our higher order interactions highlight that dominance further plays a role in how bilinguals interpret a visual scene. Nevertheless, the effect of dominance mainly affects whether or not the frequency bias is apparent in Spanish-only instructions. Consequently, the fact that English-dominant and Spanish-dominant bilinguals look more towards the lower frequency item in CS conditions makes all the more revealing and highlights the socio-pragmatic function of code-switching in sentence processing. This study represents a first proof-of-concept test of investigating what is the potential benefit of code-switching to sentence processing and moves the focus away from examining switch costs at the code-switch site itself (e.g., Valdés Kroff et al., 2018). Future studies should include an English-Spanish code-switch direction condition and the main effect of CS Direction and its interaction with Dominance in the analysis. This manipulation will allow us to determine whether it is lower dominance or higher salience of the language being switched into that are driving the predictive process.

## Conclusions

The primary goal of the study was to account for the discrepancy between the processing costs and the ubiquity of code-switching, as well as to bridge the gap between the sociolinguistic and experimental psycholinguistic research on code-switching. The psycholinguistic focus on code-switching costs may be neglecting the fact that code-switching does not occur randomly and can serve various socio-pragmatic functions. It is crucial for any sound psycholinguistic theory of code-switching to account for sociopragmatic functions associated with it (Myers-Scotton, 2006), such as the one experimentally probed here: discourse organization in terms of information distribution.

The results of the study corroborate prior findings on the information distribution within bilingual discourse (Myslín & Levy, 2015). The results also confirm our hypothesis that code-switching provides experimentally detectable processing benefits as a cue to bilinguals to anticipate unexpected information, much like disfluencies in studies on monolinguals (Arnold et al., 2003; 2004; 2007). Interestingly, the production of disfluencies and code-switches are associated with production costs, yet both highlight the potential for benefits to comprehension. That CS in discourse can affect ease of processing in comprehension is also in line with models linking production and comprehension, such as the Production-Distribution-Comprehension model (PDC; MacDonald, 2013). These models posit that production pressures shape the language distribution. Subsequently, emerging distribution patterns affect the ease of language comprehension.

In this study, we began with a simple operationalization of unexpectancy or salience as lexical frequency. However, this function of L1-L2 code-switching could extend to other non-salient/salient information contrasts. Future studies could thus probe the role of code-switching in on-line processing of given vs. new and emotionally neutral vs. taboo information (Tomic & Valdés Kroff, in prep.). We hope that this and similar studies will further open the scientific conversation on the roles of code-switching in language processing and continue to bring psycholinguistic research closer into alignment with sociolinguistic approaches to code-switching.

References:

Albirini, A. (2011). The sociolinguistic functions of codeswitching between Standard Arabic and Dialectal Arabic. *Language in society*, *40*(5), 537-562.

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, *38*(4), 419-439.

Arnold, J. E., Kam, C. L. H., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 914.

Arnold, J. E., Tanenhaus, M. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological science*, *15*(9), 578-582.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of psycholinguistic research*, *32*(1), 25-36.

Arnold, J., Wasow, T., Ginstrom, R., & Losongco, T. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language,* *76*(1), 28–55.

Auer, P. (1988). A conversation analytic approach to code-switching and transfer. *Codeswitching: Anthropological and sociolinguistic perspectives*, *48*, 187-213.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In C. Cave´, I. Guaı¨tella, & S. Santi (Eds.), *Oralite´ et gestualite´: Interactions et comportements multimodaux dans la communication* [Orality and gestuality: Multimodal interaction and behavior in communication] (pp. 597–600). Paris: L’Harmattan.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.<[doi:10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01)>.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., & Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10*(2), 344­380.

Beatty-Martínez, A. L., & Dussias, P. E. (2017). Bilingual experience shapes language processing: Evidence from codeswitching. *Journal of Memory and Language, 95*, 173-189.

Bentahila, A. (1983). Motivations for code-switching among Arabic-French bilinguals in Morocco. *Language & communication*.

Bosker, H. R., Quené, H., Sanders, T., & De Jong, N. H. (2014). Native ‘um’s elicit prediction of low-frequency referents, but non-native ‘um’s do not. *Journal of memory and language*, *75*, 104-116.

Broersma, P., & Weenink, D. (2018) “Praat: Doing phonetics by computer” [computer program], version 6.0.36. Retrieved from from <http://www.praat.org>, 7 June 2018.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive psychology*, *42*(4), 317-367.

Fox, J. (2016) *Applied Regression Analysis and Generalized Linear Models*, Third Edition. Sage.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression, 3rd edition*. Sage, Thousand Oaks CA. <http://tinyurl.com/carbook>.

Fox, J., & Weisberg, S. (2011). Multivariate linear models in R. *An R Companion to Applied Regression. Los Angeles: Thousand Oaks*.

Fricke, M., Kroll, J. F., & Dussias, P. E. (2016). Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. *Journal of memory and language, 89*, 110-137.

Gardner-Chloros, P. (2009). *Code-switching*. Cambridge University Press.

Gumperz, J. J. (1982). *Discourse strategies* (Vol. 1). Cambridge University Press.

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.

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.

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

Mirman, D. (2014). Growth Curve Analysis: A Hands-On Tutorial on Using Multilevel Regression to Analyze Time Course Data. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).

Moreno, E. M., Federmeier, K. D., & Kutas, M. (2002). Switching languages, switching palabras (words): An electrophysiological study of code switching. *Brain and Language, 80,* 188–207. <http://dx.doi.org/10.1006/brln.2001.2588>

Myers-Scotton, C. (2006). Natural codeswitching knocks on the laboratory door. *Bilingualism: Language and Cognition*, *9*(2), 203-212.

Myers-Scotton, C., & Jake, J. L. (2001). Explaining aspects of code-switching and their implications. *One mind, two languages:* *Bilingual language processing*, 84-116.

Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in society, 22*(4), 475-503.

Myslín, M., & Levy, R. (2015). Code-switching and predictability of meaning in discourse. *Language*, *91*(4), 871-905.

Olson, D. J. (2017). Bilingual language switching costs in auditory comprehension. *Language, Cognition and Neuroscience, 32*(4), 494-513.

Poplack, S., & Meechan, M. (1998). Introduction: How languages fit together in codemixing. *International journal of bilingualism, 2*(2), 127-138.

R Core Team (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arévalo, A. L., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex, 41*(1), 7-26.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O., Arevalo, A., Vargha, A., Butler, A. C., Buffington, R., & Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247-250.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods Instruments & Computers, 35*(4), 621-633.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634. <http://dx.doi.org/10.1126/science.7777863>

Tomic, A., & Valdés Kroff, J. R. (in prep.). Effects of code-switching on the emotional processing of taboo words.

Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Dussias, P. E. (2018). Experimental contributions of eye-tracking to the understanding of comprehension processes while hearing and reading code-switches. *Linguistic Approaches to Bilingualism, 8*(1), 98-133.

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.

Valdés Kroff, J. R. (2016). Mixed NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing. *Spanish-English Codeswitching in the Caribbean and the US*, 281-300.

Velasquez, M. C. (2010). *Language and identity: Bilingual code-switching in Spanish-English interviews* (Doctoral dissertation).

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag: New York.

1. Only the FreqFix:OT3 interaction was found to be significant, β = -0.265, SE = 0.105, t = -2.518. [↑](#footnote-ref-1)
2. The model without OT2 and OT3 Participant intercepts failed to converge. It is interesting that the OT2 and OT3 were necessary to explain the variation in looks among participants, while they did not provide a significant improvement as fixed effect. [↑](#footnote-ref-2)