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 large spectrum of sociopragmatic purposes (……………). Nevertheless, experimental psycholinguistic research has so far mostly focused on the processing costs of code-switching (Litcofsky & Van Hell, 2017; Olson 2017), as well as potential attenuation of these costs in certain circumstances (Fricke et al., 2016; Guzzardo Tamargo et al., 2016; Valdés Kroff et al., 2017). While this line of research provides important insights into the cognitive control processes associated with switching languages, it does not take into account the sociopragmatic motivations for code-switching and potential benefits that code-switching could provide in processing subsequent segments. D**espite the processing costs associated with it (at least in cued paradigms, e.g. Meuter & Allport, 1999),** code-switched speech makes up a relatively large portion of bilingual discourse - around 20% (Beatty-Martinez & Dussias, 2017). This ubiquity of code-switching **suggests that code-switching affords benefits to bilinguals which override the purported processing costs. The study reported here experimentally tested and confirmed one such processing benefit of CS, founded in sociopragmatic motivations: alerting to and aiding prediction of lower frequency, unexpected words.**

Some of the most cited functions of code-switching are

A more recently proposed function for code-switching is organizing discourse in terms of information content (Myslin & Levy…………, 2015). The authors analyzed bilingual corpora using statistical modeling, as well as the results of an offline cloze task to find that bilinguals code-switch from their more frequently used language to their less used, or more salient language at more information-rich, unpredictable words. The offline task, Shannon guessing game, showed that bilinguals expect harder-to-predict words to be the L1-L2 code-switch sites, suggesting that bilinguals use code-switches as a discourse marker for information content. Another sociolinguistically attested motivation for switching to the more marked language is speaking about emotional, predominantly taboo topics (Bentahila, 1983). CS to the more salient code could thus alert to increased emotionality as well, which is another type of information content. These observations suggest that code-switching could offer online-processing benefits by alerting to and thus aiding the prediction of highly informative portions of speech.

This potential function of code-switching puts it in the line with other discourse-organizational markers, such as disfluencies. Disfluencies, or irregularities in fluent speech, such as “uh”, “um”, pauses have been found to occur prior to ………….

Experimental research showed that bilinguals 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 paradigm 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.

Given the parallels in the distribution of CS and disfluencies, as well as the role of disfluencies in on-line processing, we propose to test whether CS can serve a similar function, i.e. whether code-switchers are able to make use of this discourse marker during on-line processing to help them predict more unexpected, lower frequency words coming up.

Stimuli:

Start with the same letter?

Cutting and intensity scaling?

Participants:

Age 18-32 (M = 20.83, SD = 3.53)

Gender (26 female, 4 male)

Age of acquisition of Spanish and English

All participants reported having begun learning both English and Spanish before puberty (Spanish AoA Mean = 0.67, SD = 2.35, Max = 12; English AoA Mean = 3.67, SD = 2.48, Max = 10).

**Proficiency – dominance**

**Graphical dominance but continuous variable**

## **Introduction and Background**

Nevertheless, studies have found that there is a cost associated with its processing (Altarriba et al., 1996; Meuter & Allport, 1999), albeit reduced with the increase in experiments’ ecological validity (Blanco-Elorrieta & Pylkkänen, 2017). This seeming contradiction between the ubiquity of code-switching and the associated processing costs has not found a clear explanation in the current psycholinguistic research on the online processing of CS.

Nevertheless, offline psycholinguistic studies and sociolinguistic observations offer converging socio-pragmatic and information-organizational motivations for CS. These motivations might outweigh any purported costs of processing for bilinguals and solidify the use of CS as a discourse marker.

## **Research Design**

The participants were presented with two-picture panels, one representing a low-frequency word, and the other a high-frequency word. The participants listened to the instructions on which object to click on. The instructions were either in monolingual Spanish, or took the form of Spanish-English code-switched sentences. The code-switch preceded the name of the object by 3 words, one content and two function ones, to avoid any immediate effects of the code-switch processing 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 3 bilingual speakers. Examples………………..

The independent variables are Language of the Instructions with 2 levels: Spanish (L2) and Spanish-English (L1-L2 code-switch), and Frequency of the Fixated Image: Low and High. The study was thus geared towards Spanish-dominant or Spanish-first bilinguals. Only Spanish-English code-switched instructions were chosen to simplify the design and since L1-L2 is the attested switch direction used to signal more informative portions of speech to come (Myslín & Levy, 2015). Further studies will include English-Spanish and English monolingual conditions to more precisely determine the factors affecting frequency-level prediction in code-switching.

Sound recordings and preprocessing steps

Instructions were recorded with the help of a balanced Spanish-English speaker of Puerto Rican descent. Two carrier verbs were chosen for the carrier phrases: eligir - ‘choose’, and encontrar - ‘find’, to increase variety and approach ecological validity. The carrier phrases were as follows:

‘Encuentra el dibujo de un/una/Ø \_\_\_\_\_.’

‘Elige el dibujo de un/una/Ø \_\_\_\_\_.’

‘Encuentra el drawing of a/an/Ø \_\_\_\_\_.’

‘Elige el drawing of a/an/Ø \_\_\_\_\_.’

The names of the pictures were recorded in isolation, with the speaker being given the instructions to imagine the noun coming at the end of a sentence. The carrier phrases were recorded in combination with a mock noun, and subsequently cut, to ensure that the intonation and article pronunciation were as natural as possible. The carrier phrases containing code-switches were somewhat longer in the initial Spanish portion and contained cues for the upcoming code-switch. This was not deemed a problem, since it mimics the way code-switches are uttered in real-world conversations and it could potentially alter the comprehenders to the code-switch and make the code-switch processing “easier”, i.e. less of a surprise. This would ultimately allow the comprehenders to smoothly process the code-switch, focusing only on its informative value.

All carrier phrases were scaled to the average intensity of 70 decibels, and all nouns were scaled to the average intensity of 66 decibels, using sound manipulation software Praat (Broersma & Weenink, 2018), to ensure volume uniformity among carrier phrases and noun, and a natural volume decline at the end of a sentence. The final versions of carriers and nouns were concatenated without a pause, again for the sake of greater ecological validity as there was no pause in the full carrier phrase + noun recordings.

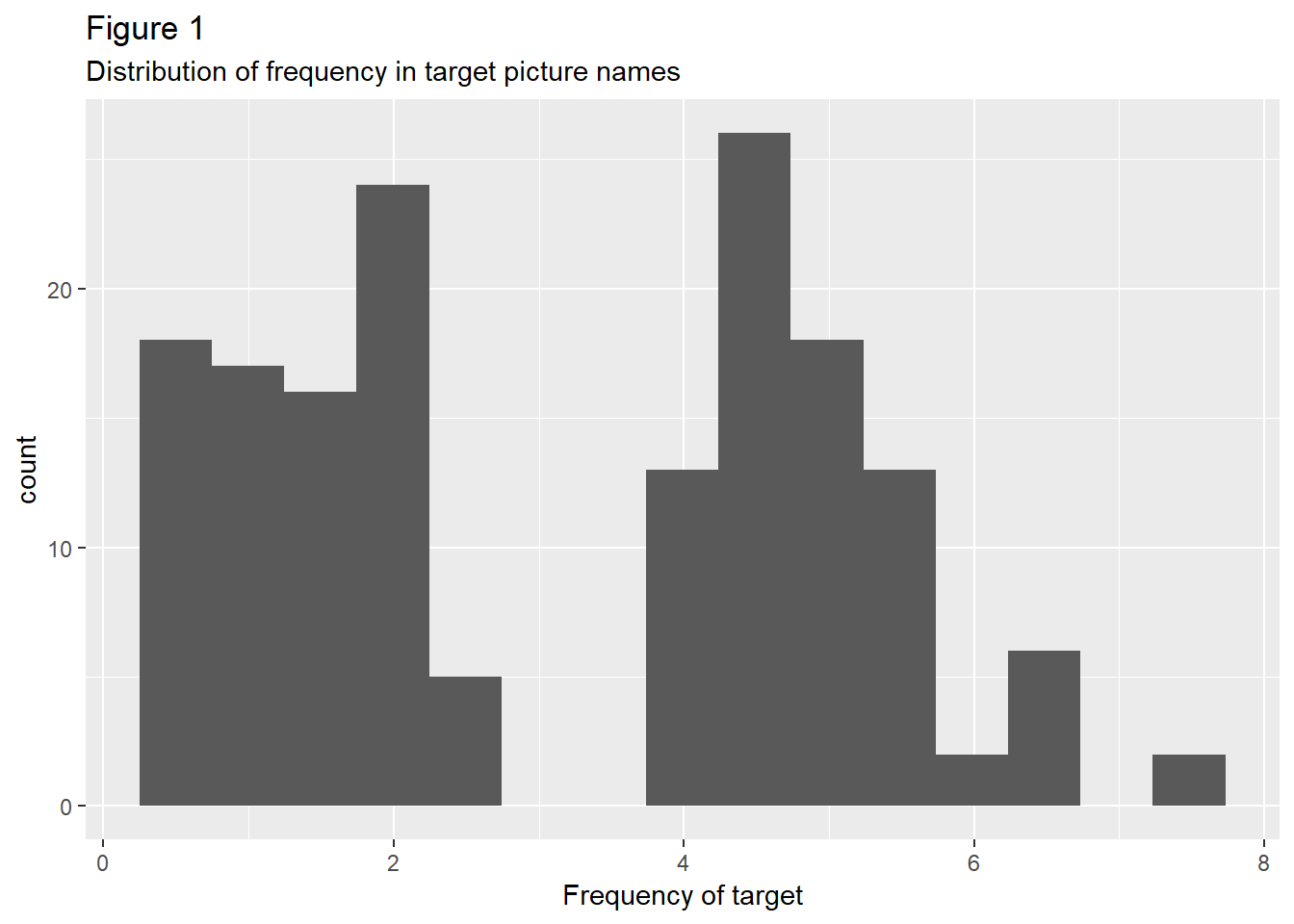
## **Materials**

### 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., 2004, 2005), 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).

Experimental images were chosen such that there is the smallest possible discrepancy in the frequency of the Spanish and English picture name. The frequency for English words ranges from 0.69 to 6.08 log natural transformed frequency counts, whereas the Spanish word frequency ranges from 0.6 to 7.58 counts. Average absolute difference between the English and Spanish object name counterparts was 0.444 frequency counts (range: 0 – 1.61). A paired t-test showed no significant difference between English and Spanish picture name counterparts, ensuring that no extraneous factors affect the results of the study (t[63] = -1.5286, p-value = 0.1314, mean of differences -0.1090411). (English counterpart low frequency group range: 0.693-2.303; Spanish counterpart low frequency group range: 0.602-2.303; English counterpart high frequency group range: 3.989-7.396; Spanish counterpart high frequency range: 4.007-7.584.)

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 3.535894 frequency counts for CS trials, and 3.595042 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 < 0.0001; English counterparts: t(33) = 33.993, p-value < 0.0001). Figure 1……………. below is a histogram representing the distribution of the target picture name frequencies in all lists. There is a marked gap between the frequencies of the two groups, ensuring that the research design includes the desired manipulation.



In addition to being paired for the maximum frequency difference between low and high frequency member pairs, as well as for the maximum frequency similarity between Spanish and English picture name counterparts, the experimental two-picture panels were also matched for the gender of the Spanish picture name/translation equivalent. Out of 32 experimental trials, there were 15 where both pictures had a feminine gender Spanish name.

Due to difficulties in attaining a sufficient number of 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 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, Low Frequency target;

CS Low Frequency Target;

Spanish High Frequency target;

CS High Frequency target.

This process resulted in 8 items per condition in a list and ensured that a participant sees a given experimental pair 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 sixty-four filler trials in each list, twice as many as there were experimental trials. The fillers were mostly neutral and similar in terms of frequency (………….). The fillers were the same across lists, but their order and position was randomized in presentation.

Presentation

The experiment was programed in Experiment Builder (………………….). Images were presented two at a time in the panels. The pictures could appear anywhere on the circular pattern shown in graph x………….. The position of the images was randomized prior to feeding the stimuli list into Experiment Builder, by 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.

Participants were greeted and given instructions in the form of a code-switched written and spoken passage to listen to instructions and click on the correct image. Six …………… practice items followed the calibration process. Audio instructions playback was delayed by 200ms after the presentation of the panel. The participants were instructed to use a mouse to click on the correct image.

Participants

Thirty Spanish-English bilingual participants who self-report code-switching were recruited on the UF campus. They were compensated for their participation in course credit or cash.

Participants completed Language History Questionnaire (………………), semantic fluency tests in Spanish and English, adapted standardized grammar tests: Michigan English Language Institute College English Test (MELICET) and Diplomas of Spanish as a Foreign Language (DELE), and the cognitive control task AX-CPT. 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. Upon the completion of the main experiment, the participants rated the words which appeared in the experiment on the frequency of use and amount of exposure. These ratings would have been used in case there was a large incorrect trial percentage, or a large discrepancy between the frequency counts of experimental words taken from CELEX compared to participants own exposure to and use of the words.

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 the question of the frequency of code-switching use was 4.1, SD = 0.845, with 1 meaning standing for “Never” and 5 for “Always”. The mean response to the question on the exposure to code-switching in spoken conversation was 3.867, SD = 0.973, with the same response coding.

Two different proficiency measures were collected in addition to self-reports: standardized grammar tests and semantic fluency tests. Proficiency scores were transformed into dominance by taking the ratio between the Spanish- and English-language standardized tests DELE and MELICET, and the ration between Data on semantic fluency was also collected and

|  |  |  |  |
| --- | --- | --- | --- |
|  | 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 |  |  |  |

Data Analysis

Linear Mixed Effects Model

The target time period for eye-movement analysis was 200 ms before the onset of the target and 200 ms after the onset of the target. Planning and executing an eye-movement has been shown to take around 200 ms, so the 200 ms post-target-word region still reflects processes which occurred prior to the presentation of the target.

(13 out of 960)

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. Each participant saw 8 items in the same condition, amounting to 160 ms total time which one participant could spend looking at the set of items in a particular condition within one time-bin. The dependent variable in the analyses was the Proportion of milliseconds within a time-bin that each participant spent looking towards a particular image, regardless of their target/distractor status, in the Spanish and the Code-switched language context. The independent variable Frequency of the Fixated item …………

The proportion of time spent looking at a set of images in a particular condition was normalized by z-scoring. ………………. (………After calculating the proportions, blinks and in-saccade eye-states were removed from the data.

Dominance (DELE/MELICET). All continuous predictors have been normalized by z-scoring. Bins/Time were excluded from the analysis due to autocorrelation………………..

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 lme4 package in R (Bates et al., 2011) with the independent variables Language Context (contrast coded as Spanish -0.5, Code-switched +0.5), Frequency of the Fixated Item (FreqFix; contrast coded as High -0.5, Low +0.5), and Language Dominance (continuous), as well as their interactions as predictors. Participant intercept and Language Context and FreqFix slopes were included in the model as random effects. None of the main effects was found to significantly affect the dependent variable. Nevertheless, the Language Context X FreqFix interaction was significant, β = -0.481, t(4753) = -2.026, such that participants were looking at the low frequency item in the CS condition more than in the Spanish condition (Table X……). Graph X ……………. shows the data predicted by the model (line) vs. observed data (points).

**Growth Curve Analysis**

**We report the results of a model with the time variable transformed with orthogonal polynomials from 1st to 3rd order, i.e. Growth Curve Analysis (………………………). ……………. Nadji negde gde se pominje**

**The time transformed using orthogonal polynomials were included in the model along Language, FreqFix, and Dominance, as well as their interactions with each variable. The interactions among the three transformed time variables were not included in the model. All continuous variables have been standardized as described above. Participant** **Language Context, and FreqFix intercepts, as well as Language and FreqFix slopes for each transformed time variable, were included in the model as random effects.**

**Results of the analysis are summarized in the Table X ……………**

**Only Language main effect was found to be significant, …………….., such that there were fewer looks overall in the CS condition compared to Spanish. Importantly , several interactions were found to be significant. The** Language x FreqFix **interaction proved significant,** **β** **= 0.492, SE = 0.047, , t(4750)= 10.374, such that Low frequency items were fixated more in the CS condition compared to Spanish condition. Also, the interaction of** Language, Dominance, and FreqFix **was significant, β** **= 0.287, SE = 0.047, t(4750) = 6.066, such that** **the participants with higher relative Spanish dominance looked at the Low frequency items more in the CS condition compared to Spanish condition. The interaction of** Language, FreqFix, and the 1st order polynomial **transformed time variable was also found to be significant, β = 0.891, SE = 0.211, t(****4750) = 4.218 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 1st order polynomial transformed time variable was significant as well, β = 0**.82025, SE = 0.21126,** t(4750) = **3.883, 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, with modeled data represented as lines and observed as points.**

**Discussion**

**The results from the Growth Curve Analysis point to a global increase of looks to the low frequency item in the CS condition in the period of 200 ms prior to the onset of the target word and 200 ms following the onset of the target word. The chosen interval following the target onset still reflects the processing prior to the onset, since it takes about 200 ms to plan and execute a saccade (…………). This result provides support to our hypothesis that code-switching from L1 to L2 plays a role in on-line language processing: it changes bilinguals’ expectations in terms of the information load, i.e. frequency, of the following items. This result is especially significant given the Frequency Bias, i.e. 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, FreqFix, and OT1 further support our predictions, since there is a continuous increasing trend for the looks to the lower frequency items in the CS context compared to the Spanish context, as the target onset approaches.**

**The significant positive interactions Language x Dominance x FreqFix and** Language x Dominance x FreqFix x 1st order polynomial transformed time variable **suggests that the bilinguals with a higher proficiency in Spanish are more privy to the cues provided by the L1-L2 code-switch, or can simply make use of it in real time more efficiently than relatively lower Spanish proficiency bilinguals. This was expected, since these bilinguals presumably (check this………….) use Spanish more in daily life and code-switch more as well.**

**Conclusion**

**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.** It is crucial for any sound psycholinguistic theory of of code-switching to account for sociopragmatic factors associated with it (Myers-Scotton, 2016), such as the one tested ………..here: discourse organization. **The results of the study corroborate the findings of the off-line cloze study and corpus modeling of the information distribution within bilingual discourse (Myslin & Levy, 2015………………..).**

**These results confirm our hypothesis that code-switching provides processing benefits to bilinguals which can be probed experimentally, much like disfluencies in studies on monolinguals (…………….). Here, we began with a simple operationalization of unexpectancy as lexical frequency. However, in our view, this could extend to other areas, e.g., given/new; neutral/taboo…………….**

**These findings are also in line with the distribution based models, such as Production-Distribution-Comprehension (PDC, MacDonald, 2013; ………..). These models posit that production pressures shape the language distribution, with patterns in the distribution subsequently affecting the ease of language comprehension.**

**We hope that this and similar studies will further open the scientific conversation on the roles of code-switching.**

**Future studies**

* **Include an English-Spanish CS condition:**
  + **to see how prediction is affected by the direction of CS in combination with dominance**
  + **This will also help us separate whether the effects are driven by dominance v salience**

**Only target looks for basic model with participants random and participants+items random**

**And ortho poly**

**Items as thingies random effects – nope, the total looks are then just 20 ms, cause each item in a set of items in the same condition has a different pair number, so the proportion is like 15 looks to target out of 20, the props are then showing per each individual item, per each individual bin**

**Both as random effects?**

**Do the items as random thingies in CS-aids-prediction**

**CS as a freaking variable**

How did I do it right the first time?

N-k-1 = df

4757 = N

k = number of variables in the model

Reml was true here, but what if it should be false?

I’ve done it once like a pro, I’ll do it again.

One advantage of mixed models compared to traditional repeated-measures ANOVA is that we’re explicitly calcuating a regression model. As such, we have an individual estimate for each interaction as part of our model summary. We do not have to resolve interactions explictly, we can instead simply look at the individual interaction levels in the coefficients.