Investigating the costs of intra-sentential code-switching for lexical processes

Anthony Yacovone, Emily Moya, & Jesse Snedeker (Harvard University)
*anthony_yacovone@g.harvard.edu

When bilinguals speak to one another, they often engage in *code-switching* or switching between languages within a single utterance. Prior studies have found that code-switched words are harder to process[1] resulting in longer reaction times in naming tasks[2] and larger N400 responses in lexical decision and reading tasks.[3,4] However, the majority of these code-switched EEG tasks have used single words or single sentence contexts. We reasoned that these code-switching difficulties might be facilitated by top-down expectations developed over the course of a spoken discourse. In this study, we take advantage of new naturalistic EEG techniques to see how code-switches are perceived in rich contexts. We also compared the N400 responses for code-switched words with those for other within-language lexical switches to see if there are any savings to retrieving code-switched words that strongly fit with the context.

The N400 response is an index of the ease of lexical processing,[5-8] in particular *two* sub-processes: lexical access[9-10] and lexical integration,[11] which occur at different times within the N400 window.[12] Lexical access occurs in the earlier interval of 250-350ms when the N400 is sensitive to contextual expectations (e.g. cloze probabilities) and forms of lexical pre-activation.[13] Lexical integration occurs later (350-450ms) when lexico-semantic processes begin and integration difficulties are indexed by robust N400 modulations.[14] Again, the EEG studies showing reliable N400 modulation for code-switched words used primarily isolated words and sentences.[1] Recent accounts have argued that N400 modulation in these isolated word tasks index lexical access difficulties—as bilinguals inhibit their native language (L1) to better retrieve L2 words.[1] Alternatively, N400 modulation in tasks with more context are argued to reflect lexical integration difficulties—as bilinguals must adjust activation of both languages to help build up context.[1] Thus, we can interpret N400 effects to reflect difficulties with accessing a less-expected lexical form[13] and/or integrating meaning across multiple languages.[14]

In this study, we investigated how enhanced discourse contexts influence the ease of these lexical processes during code-switching. Spanish-English bilinguals passively listened to oral narratives told by a bilingual speaker while EEG signals were recorded. Our participants largely reported hearing code-switching in their everyday lives (**Table 1**). Passive listening tasks have been used in recent EEG studies; however, these approaches have only been correlational.[15-16] To circumvent this limitation, we successfully grafted experimental controls onto this paradigm, preserving its ecological-validity while reaping the benefits of having factorially-manipulated variables. We manipulated the *contextual fit* of the words (**strong**, *weak*) and the *language* (English, Spanish). The story-original English words and their translations had strong contextual fits, whereas the English alternatives and their translations had weaker fits. We quantified *contextual fit* as cloze probabilities (**strong**: 97.5-12.5%; *weak*: 0%) and naturalness ratings out of 7 (**strong**: 6.38; *weak*: 3.28). We then embedded our selected targets by identity-splicing and cross-splicing the material into the discourse—this process resulted in sentences like "The wig is so hot and heavy on my (**head** | **cabeza** | *cranium* | *craneo*)." We reasoned enhanced top-down expectations could facilitate lexical access (early differentiations between Spanish conditions at ~300ms) and lexical integration (late differentiation at ~450ms) for code-switched words.

Our analyses (N=32) replicate the previous findings that contextually-weak English words and code-switched words elicit N400 effects within 300-500ms (**Figure 1**). Furthermore, we find an interaction such that contextual fit only significantly affects English targets early on (i.e. Spanish conditions show similar N400 amplitudes). These findings suggest that the initial access of both Spanish conditions show similar degrees of difficulty to each other and to the contextually-weak English condition. However, despite these initial difficulties, there is rapid facilitation for contextually-strong Spanish words (~400ms), which is consistent with evidence that code-switches can be less costly to process than within-language switches in certain contextually-rich contexts.[17] Taken together, these data support a two-factor model of the N400 response with early form-based and late context-based effects, and that enriched top-down expectations in spoken discourses can facilitate code-switching when the word makes sense in the given context.

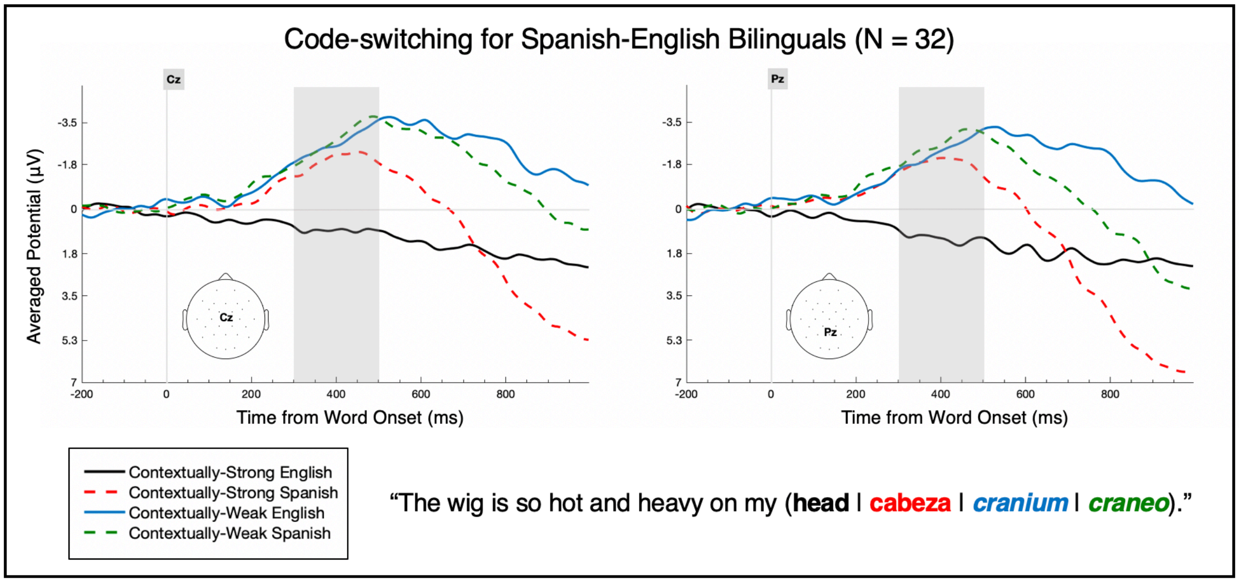

Figure 1: Grand averaged N400 responses to both within-language lexical switches (*cranium*) and Spanish translations for both strong (*cabeza*) and weak (craneo) English targets at midline electrodes Cz and Pz.

Table 1: Participant responses to how often they hear code-switching in their everyday lives.

Often	22 participants
Sometimes	9 participants
Never	1 participant

Model outputs from linear mixed effects model and pairwise comparisons (lme4):

Main Model:
Predictability (β = -3.37, SE = 0.68, t = -4.86, p < .001),
Language (β = -2.62, SE = 0.58, t = -3.84, p < .001),
*Predictability*Language* (β = 2.58, SE = 0.85, t = 3.00, p < .01)

Pairwise 1:
Within English: High vs. Low (β = -4.19, SE = 1.06, t = -3.96, p < .001);
Within Spanish: High vs. Low (β = -0.78, SE = 0.66, t = -1.19, p = .55)

Pairwise 2:
Within High: English vs. Spanish (β = -6.84, SE = 1.43, t = -4.79, p < .001);
Within Low: English vs. Spanish (β = -0.04, SE = 0.50, t = -0.07, p = 1.00)

References:

[1]Van Hell, Litcofsky, & Ting, 2015; [2]Bobb & Wodniecka, 2013; [3]Van der Meij et al., 2011; [4]Proverbio et al., 2004; [5]Van Hell & Witteman, 2009; [6]Petten, 1993; [7]Kutas & Federmeier, 2009; [8]Kutas & Federmeier, 2011; [9]Kutas & Federmeier, 2000; [10]Lau, Almeida, Hines, & Poeppel, 2009; [11]Brown & Hagoort, 1993; [12]Grainger & Holocomb, 2009; [13]Lau, Holcomb, & Kuperberg, 2013; [14]Molinaro & Carreiras, 2010; [15]Brennan, Cantor, Eby, & Hale, 2016; [16]Levari & Snedeker, 2018; [17] Moreno, Federmeier, & Kutas, 2002.