Supplementary Analyses:

Probability Unit Analysis (PROBIT)

Participants' perceptual boundary (i.e., 50% crossover) along a VOT range used to distinguish only Spanish voiced and voiceless stop consonants in both Spanish and English contexts was quantified using PROBIT analysis.

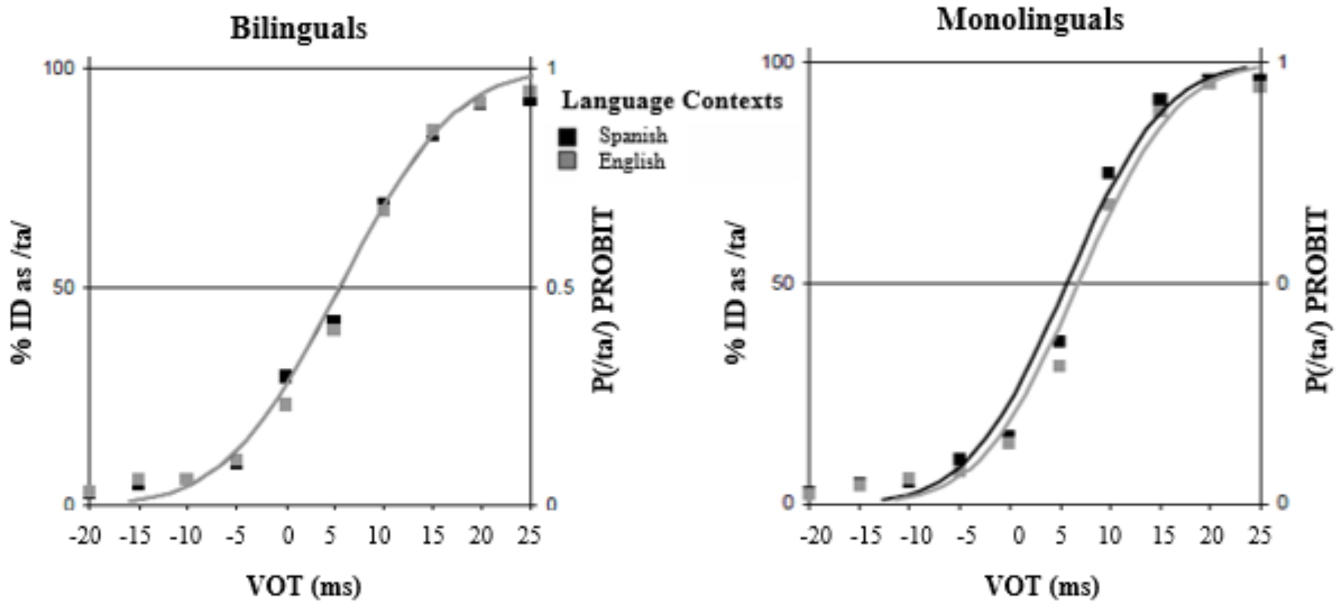
PROBIT analysis associates a probability value to each stimulus based on the number of times each stimulus was presented, and the number of times each stimulus was judged /ta/. Thus, PROBIT analysis can easily account for the unequal delivery of stimuli within an oddball paradigm (i.e., 80% standards/20% deviants), which was needed to concurrently collect electrophysiological data in the current study. Further, logistic regressions typically used by studies that present each stimulus equally (Casillas & Simonet, 2018; Flege & Eefting, 1987a; García-Sierra & Champlin, 2003; García-Sierra, Diehl, & Champlin, 2009; Gonzales & Lotto, 2013; Gonzales, Byers-Heinlein, & Lotto, 2019), and PROBIT analyses have been shown to consistently produce the same results (Long, 1997).

Accordingly, each participant's /ta/ cumulative distribution was transformed into a cumulative probability distribution by means of a PROBIT function. Standard sounds (i.e., -20 ms to 0 ms VOT) that did not precede deviant sounds (i.e., 5 ms to 25 ms VOT) were not included in the final average to provide the same number of deviant and standard responses in the statistical analyses.

Figure S1 shows how monolinguals and bilinguals judged the stimuli as a function of VOT. The figure shows the percent (squares) and the probability (line) that each stimulus was perceived as /ta/.

<Insert Figure S1 about here>

Figure S1. Bilinguals' and monolinguals' phonemic boundaries (50% point) in two language contexts derived from curves fitted via PROBIT analysis. Note: percent values (left y-axis) are indicated with squares and probability values (right y-axis) are indicated with lines.



The VOT values at the 50% crossover point were compared in a 2 x 2 (group x language context) analysis of variance (ANOVA) with repeated measures (language context). The ANOVA results revealed no significant main effects for language or group (F (1, 52) = 1.739, *p* = .193, F (1, 52) = .295, *p* = .589), nor interaction (F (1, 52) = 2.105, *p* = .153). This suggests that both groups' perceived boundary across both language contexts did not significantly differ.

Although upon visual inspection of Figure S1, stimuli 0, 5, and 10 ms of VOT display the highest variability across contexts in both groups. Thus, three independent 2 x 2 (group x language context) ANOVA with repeated measures (language context) were performed for stimuli 0, 5, and 10 ms of VOT.

Stimulus 0 ms VOT revealed a significant main effect for group (F (1, 52) = 8.20, *p* = .006) and an effect approaching significance for language context (F (1, 52) = 3.70, *p* = .061), but no significant interaction (F (1, 52) = 1.01, *p* = .319). Bilinguals' /ta/ identification mean was larger in the Spanish context (Mean = .32, SD = .21) than the English context (Mean = .25, SD = .18, *t* (52) = 1.81, *p* = .08), although this difference was statistically insignificant. No differences across language contexts for monolinguals were revealed for 0 ms of VOT (Mean in Spanish context = .15, SD = .16, and Mean in English context = .13, SD = .19, *t* (52) = .77, *p* = .45). This suggests that only bilinguals showed a tendency to identify stimulus 0 ms VOT in alignment with the phonetic rules of the immediate language context, as expected.

Stimuli 5 and 10 ms VOT revealed no significant main effects for group (F (1, 52) = 1.88, *p* = .175, and F (1, 52) = .03, *p* = .86; respectively), language (F (1, 52) = 2.37, *p* = .130 and F (1, 52) = 2.54, *p* = .12; respectively), nor interaction (F (1, 52) = .61, *p* = .44, and (F (1,

52) = 2.72, p = .10; respectively). Accordingly, no further pair-wise comparisons were done for these stimuli.

Discussion

These results reveal that both bilinguals and monolinguals similarly perceived a VOT continuum, whose phonetic ranges provided only a Spanish phonemic contrast, in both Spanish and English language contexts. Specifically, all participants divided the range of stimuli in two equal halves (i.e., 50% crossover at or near stimulus 0 ms VOT). This can be explained in more than one way.

First, it is possible that range effects played a major role. In other words, participants placed the phonetic boundary as a function of the range of stimuli tested, as opposed to linguistic experience. Since the current study presented the same range of stimuli in both contexts, range effects would explain our observed lack of a boundary shift across contexts. Range effects would further explain why monolinguals identified a phonetic boundary that better corresponds to the phonemic categories of Spanish (i.e., at the midpoint of continuum; 0 ms VOT) rather than those of English (i.e., towards the voiceless end of the continuum; 25 ms VOT) (Brady & Darwin, 1978). On that note, since the range of perceptual cues did not provide monolinguals the opportunity to perceive the ideal English phonemic contrast, they instead used the range of stimuli to detect the ideal acoustic contrast provided at the midpoint of the continuum (i.e., range effects).

Although, it is possible that Spanish-English bilinguals' boundary placement at the midpoint of the VOT continuum across language contexts was influenced by linguistic processes. Explicitly, the phonetic range of perceptual cues provided an ideal Spanish phonemic contrast between voiced and voiceless stop consonants at the midpoint of the continuum (i.e., 0 ms VOT). In accordance with research that suggests bilinguals are increasingly sensitive to perceptual cues, including VOT (Ju & Luce, 2004; Lagrou et al., 2011; Llanos et al., 2013), Spanish-English bilinguals in this study may have recognized that the VOT range tokens best suited Spanish phonemic contrast, and thus placed the boundary with respect to Spanish phonemic categories regardless of language context.

Therefore, we chose to analyze participants' overall sensitivity to the VOT continuum across contexts using Signal Detection Theory (i.e., cumulative *d'*). If bilinguals' boundary placement was indeed influenced by linguistic processes, we would expect bilinguals to show sensitivity differences to stimuli across contexts. Simply, sensitivity differences would suggest that the language context, and its linguistic properties, influenced bilinguals' perception of the respective stimuli. However, since monolinguals have not been observed to have the same level of linguistic perceptual sensitivity as bilinguals, we do not expect monolinguals to reveal sensitivity differences to stimuli across contexts. Cumulative *d'* analyses, results, and discussion are included in the manuscript (see sections *Data analyses: Behavioral data, Results: Behavioral responses, and Discussion: Behavioral responses*, respectively).

References

Brady, S. A., & Darwin, C. J. (1978). A range effect in the perception of voicing. *Journal of the Acoustical Society of America, 63*, 1556-1558.

Casillas, J. V., & Simonet, M. (2018). Perceptual categorization and bilingual language modes: Assessing the double phonemic boundary in early and late bilinguals. *Journal of Phonetics, 71*, 51-64.

Flege, J. E., & Eefting, W. (1987a). Cross-Language Switching in Stop Consonant Perception and Production by Dutch Speakers of English. *Speech Communication, 6*, 185-202.

García-Sierra, A., & Champlin, C. A. (2003). Shift in the perceived voicing boundary of bilingual listeners. *Journal of the Acoustical Society of America, 114*, 2336.

García-Sierra, A., Diehl, R. L., & Champlin, C. (2009). Testing the double phonemic boundary in bilinguals. *Speech Communication, 51*, 369-378.

Gonzales, K., Byers-Heinlein, K., & Lotto, A. J. (2019). How bilinguals perceive speech depends on which language they think they're hearing. *Cognition, 182*, 318-330.

Gonzales, K., & Lotto, A. J. (2013). A bafri, un pafri: Bilinguals' Pseudoword Identifications Support Language-Specific Phonetic Systems. *Psychological Science, 24*, 2135-2142.

Long, J. Scott (1997). Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage Publications.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide* (Second ed.). Mahwah, NJ: Lawrence Erlbaum Associates.