

# Effects of Social Factors on Phonetic Convergence in Language Alternation

Ernesto R. Gutiérrez Topete

*University of California, Berkeley (USA)*

Recent studies have reported on the dynamic nature of bilingual speech (i.e., two languages produced in a single discourse event) compared to monolingual speech (i.e., one language produced in a single discourse event) by multilingual speakers. For example, in monolingual speech, multilingual speakers “maintain language-specific phonetic categories”; meanwhile, in bilingual speech, the speakers displayed “phonetic convergence” [1, pg. 1280-1]. Seemingly, the active usage of multiple languages increases phonetic convergence in the production of multilingual speakers. Other studies on the production of voice onset time (VOT) during language alternation (i.e., language switching and code switching) corroborate the convergence effect observed in bilingual speech [e.g., 2, 3, 4, 5]. However, some inconsistencies have been pointed out [6, 7]. Namely, while some studies find bidirectional convergence, others only find effects in one language. And among those studies with a unidirectional effect, some researchers claim the effect occurs only in the speakers’ L2 (as opposed to the L1) while others claim the dominant language is more susceptible to convergence. The wide array of research tasks used to study this phenomenon makes it all the more difficult to pinpoint the cause of the aforementioned inconsistent directionality of the convergence effect. For example, some studies relied on word list reading tasks [6] or phrase/passage reading tasks [2, 3], and others used speech spontaneously produced during sociolinguistic interviews [4], inter-subject group conversations [4, 5], or puzzle tasks [5].

In order to shine light on the possible effect of task type on the directionality of convergence effects in bilingualism research, the present study analyzes acoustic productions across four of the most popular research tasks (i.e., word list reading, passage reading, puzzle—spot the difference—and casual interview tasks) from a single group of Spanish-English bilingual speakers to obtain VOT measurements for word-initial voiceless stops /p t k/. A total of 60 Spanish-English bilingual subjects participated in the four tasks. Data collection took place in a sound booth at the Berkeley PhonLab. The audio for the word list and passage reading tasks were annotated by hand. The transcriptions for the puzzle and interview tasks were automated with OpenAI’s Whisper automatic speech recognition model. All data were then processed with the Montreal Forced Aligner [8], and VOT measurements for voiceless stop-initial words were obtained in both languages using AutoVOT [9].

A LASSO regression model with a conservative L1 regularization penalty was performed on the VOT values of the training and validation data sets. The trained model was later applied to unseen test data. While the study covers linguistic and social predictors, this presentation focuses exclusively on the social factors. The analysis included all answers and dominance scores from the BLP survey [10] as independent variables. While the data set included 52 variables, LASSO employed variable selection to identify only the most informative predictors. The results show an influence of task on English, with passage reading experiencing the highest level of convergence, followed by the interview (see [figure 1](#)). Spanish showed little variability across tasks. Among the most salient results, the statistical model revealed that those who learned Spanish later in life (i.e., L2 Spanish learners) showed lower VOT productions, suggesting a higher propensity for convergence (see [figure 2](#)). Moreover, older subjects showed higher VOT values with limited convergence in their English speech,

compared to younger speakers. Finally, speakers with more schooling in Spanish and usage of Spanish at home have lower VOT values in Spanish but also showed higher convergence in their English productions.

All in all, this study provides (1) a comparative analysis of the research tasks typically used in code-switching studies to uncover task effects in production studies, (2) a glimpse of the social factors influencing phonetic category productions, and (3) a better understanding of the language processing mechanisms that are engaged during bilingual speech.

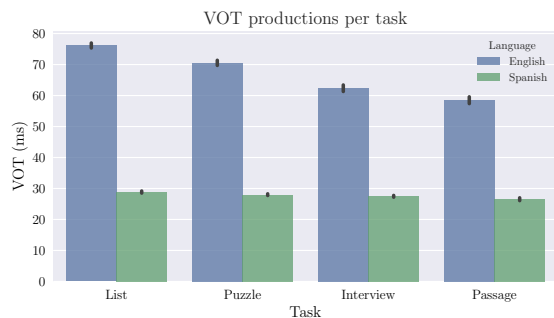


Figure 1: VOT distributions

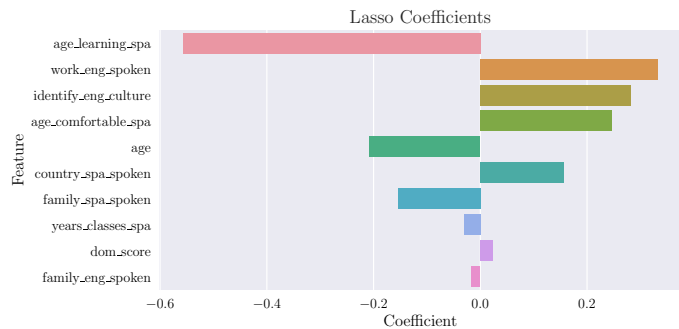


Figure 2: Top 10 social factors

## References

- [1] M. Amengual. “The acoustic realization of language-specific phonological categories despite dynamic cross-linguistic influence in bilingual and trilingual speech.” In: *The Journal of the Acoustical Society of America* 149.2 (2021), pp. 1271–1284.
- [2] A. J. Toribio et al. “Perseverative phonetic effects in bilingual code-switching.” In: *Theoretical and experimental approaches to Romance linguistics*. Ed. by R. Gess and E. Rubin. Philadelphia/Amsterdam: John Benjamins Publishing Company, 2005, pp. 291–306.
- [3] B. E. Bullock et al. “Language dominance and performance outcomes in bilingual pronunciation.” In: *Proceedings of the 8th generative approaches to second language acquisition conference*. Ed. by M. Grantham O’Brien, C. Shea, and J. Archibald. Cascadia Proceedings Project Somerville, MA. 2006, pp. 9–16.
- [4] C. Balukas and C. Koops. “Spanish-English bilingual voice onset time in spontaneous code-switching.” In: *International Journal of Bilingualism* 19.4 (2015), pp. 423–443.
- [5] P. Piccinini and A. Arvaniti. “Voice onset time in Spanish–English spontaneous code-switching.” In: *Journal of Phonetics* 52 (2015), pp. 121–137.
- [6] D. J. Olson. “Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production.” In: *Journal of Phonetics* 41.6 (2013), pp. 407–420.
- [7] M. Fricke, J. F. Kroll, and P. E. Dussias. “Phonetic variation in bilingual speech: A lens for studying the production–comprehension link.” In: *Journal of memory and language* 89 (2016), pp. 110–137.
- [8] M. McAuliffe et al. “Montreal forced aligner: Trainable text-speech alignment using kaldi.” In: *Inter-speech*. Vol. 2017. 2017, pp. 498–502.
- [9] J. Keshet, M. Sonderegger, and T. Knowles. *AutoVOT: A tool for automatic measurement of voice onset time using discriminative structured prediction [Computer program]*. Version 0.94, retrieved June 2020 from <https://github.com/mlml/autovot/>. 2014.
- [10] D. Birdsong, L. M. Gertken, and M. Amengual. *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. Retrieved from <https://sites.la.utexas.edu/bilingual/>. 2012.