

The full cycle of the sociolinguistic enterprise: Corpus building, student engagement, and critical language pedagogy

The importance of sharing sociolinguistic data has been the subject of workshops (LSA 2012, 2016), included in publications (Mallinson, 2013) and encouraged by funding agencies (NSF 2016). In line with these initiatives, is the creation of student-based corpora of U.S. Spanish (e.g., Corpus of Mexican Spanish in Salinas, Spanish in Texas, Unnamed Corpus), in which students take part in building sociolinguistic corpora. Student-based corpora in US Spanish not only provide important data for variationist research including the opportunity to analyze changes in progress (Torres Cacoullós & Berry, forthcoming), but also provide students with training in sociolinguistic methods. This hands-on approach in turn raises sociolinguistic awareness among students, especially among minority language speakers whose native dialects are often seen as mixed hybrids. In this poster, we illustrate such a project by explaining step-by-step the involvement of students in the creation and maintenance of the (Unnamed Corpus), the protocols followed to facilitate the sharing of data, and the proposed application of the corpus in the language classroom to give back to the community.

The (Unnamed Corpus) is a student-based corpus of sociolinguistic interviews of Spanish-English bilingual speakers from a border community in the U.S. Southwest. Inspired by Labov's (1984) model of neighborhood studies, we offer internships and classes to graduate and undergraduate students in which they are trained to conduct sociolinguistic research. Similar to Nagy's initiative at the University of Toronto, this research experience provides students who are native speakers of Spanish as a heritage language with the opportunity to collect data in their communities and document and analyze their home dialects. The sociolinguistic interviews are then transcribed following standard protocols by students and are rechecked multiple times to achieve maximum accuracy. In order to facilitate the sharing of sociolinguistic data, a large amount of metadata is gathered from the participants yielding a plethora of possible factors for future researchers to consider. This metadata includes a Bilingual Language Profile (Birdsong, Gertken, and Amengual, 2012), a demographic questionnaire, speakers' networks, interviewer information, and field notes. The transcribed interviews and metadata are then uploaded to a website available to scholars and educators. Finally, in order to maximize student and community engagement (Labov, 1982; Schilling, 2013, Wolfram, 2013), we have applied the pedagogical models of "Learning by Design" (Kalantzis & Cope, 2008) and "Available Designs" (Kern, 2000) to propose a framework for teachers to use the (Unnamed Corpus) to incorporate elements of sociolinguistic diversity in the Spanish heritage language classroom.

By maintaining the Unnamed Corpus with sociolinguistics students, following protocols to share sociolinguistic data, and incorporating the analysis of the corpus in the Spanish language classes, we are able to complete the full cycle of the sociolinguistic enterprise: data collection, data analysis, student engagement, and classroom practices targeted at raising sociolinguistic awareness. In doing so, this project also counters standard language ideologies propagated in the language classroom and the community, while combating bilingual speakers' linguistic insecurity, which, in turn, is known to contribute to the maintenance of minority languages (Martínez, 2003; Leeman, forthcoming).

References:

- Birdsong, D., Gertken, L.M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. COERLL, University of Texas at Austin.
<https://sites.la.utexas.edu/bilingual>.
- Kalantzis, M., & B. Cope. (2008). *New learning: Elements of a science of education*. New York: Cambridge University Press.
- Kern, R. (2000). *Literacy and language teaching*. Oxford: Oxford University Press.
- Labov, W. (1982). Building on empirical foundations. In W. Lehmann & Y. Malkiel (Eds.), *Perspectives on Historical Linguistics*. Amsterdam/Phila: John Benjamins, 17-92.
- Labov, W. (1984). Field methods of the Project on Linguistic Change and Variation. In J. Baugh & J. Sherzer (Eds.), *Language in Use*. Englewood Cliffs: Prentice Hall.
- Leeman, J. (Forthcoming). Critical language awareness in SHL: Challenging the linguistic subordination of US Latin@s. In K. Potowski (Ed.) *Handbook of Spanish as a Minority/Heritage Language*. New York: Routledge.
- Linguistics Society of America. (2012). The 86th Annual Meeting of the *Linguistic Society of America*. Portland, Oregon.
- Linguistics Society of America. (2016). The 90th Annual Meeting of the *Linguistic Society of America*. Washington, DC.
- Malinson, C. (2013). Sharing data and findings. In C. Mallinson, B. Childs, & G. Van Herk (Eds.) *Data Collection in Sociolinguistics: Methods and Applications*. NY: Routledge, 253-257.
- Martínez, G. (2003). Classroom based dialect awareness in heritage language instruction: A critical applied linguistic approach. *Heritage Language Journal*, 1(1).
- Nagy, N. *Heritage Language Variation and Change in Toronto (HLVC)*.
http://projects.chass.utoronto.ca/ngn/HLVC/0_0_home.php.
- National Science Foundation (NSF).
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=12816&org=NSF
- Schilling, N. (2013). *Sociolinguistic fieldwork*. Cambridge, UK: Cambridge University Press.
- Torres Cacoullos, R & G. M. Berry. (forthcoming). Language variation in U.S. Spanish: Social factors. In K. Potowski (Ed.), *Handbook of Spanish as a Minority/Heritage Language*. New York: Routledge.

Wolfram, W. (2013). Community commitment and social responsibility. In J.K. Chambers and N. Schilling (Eds.), *Handbook of Language Variation and Change*, 2nd edn. Malden/Cambridge: Wiley/Blackwell, 557-76.