Uriel Cohen Priva

Brown University

uriel_cohen_priva@brown.edu

Relative low informativity licenses lenition

Relative low informativity licenses lenition

Weinreich et al. (1968) proposed several related *problems* of sound change, one of which being the *actuation problem*: what causes a potential sound change to occur in a particular language at a particular time? American English deletes /t/ word-finally (Guy, 1991, among others), several dialects of Spanish lenite or delete /s/ word-finally (Hochberg, 1986, among others), and Indonesian lenites or deletes /k/ word-finally (Soderberg and Olson, 2008). Each language lenites one particular segment while preserving the other two. Although different aspects of sound change, such as phonetic limitations, are well studied (e.g. Kiparsky, 1995; Pierrehumbert, 2001; Ohala, 2003, among others), current models do not predict which language is likely to be affected by which process. Recent work has proposed information theoretic accounts for sound change (e.g. Hume, 2008; Wedel et al., 2013), which involve a balance between phonetic and communicational pressures (as modeled in Sóskuthy, 2015). I build on the notion of *informativity* (the average information provided by a segment, Cohen Priva, 2008) to argue that *relative* low informativity licenses the actuation of weakening processes (Cohen Priva, to appear).

Study 1 compared the relative informativity of segments in seven languages: American English, Spanish, Indonesian, Arabic, Japanese, Mandarin Chinese, and Korean. The lenitions of interest were /t/ in American English, /s/ in Spanish, and /k/ in Indonesian. Informativity of leniting segment was compared with their average informativity in the other six languages. The prediction is that if in some language a segment weakens, its informativity will be lower than in languages in which it does not weaken. For American English word counts were taken from the Fisher, Switchboard and Buckeye corpora. The CALLHOME lexicons were used for Spanish, Egyptian Arabic, and Mandarin. Indonesian had no spoken data and the Berita Satu newspaper written data was used. Segment informativity was estimated as in Cohen Priva (2015). Indeed, the informativity of /t/, /s/, and /k/ was lowest in American English, Spanish, and Indonesian, respectively, both relative to the cross-linguistic average and similar segments within the language, mirroring word-final weakening patterns. This study therefore provides converging evidence that relative low informativity does correlates with the actuation of sound change in specific languages.

Study 2 replicated the results of Study 1 using a set of over 40 languages for which only written data is available. Word counts were taken from the Crúbadán project (Scannell, 2007), and phonemic representation was derived from their alphabet. The sample is biased, but still has only 50% European languages. The informativity of American English /t/, Spanish /s/, and Indonesian /k/ were each in the bottom 10% of the language sample, and several of the languages in which parallel segments had lower informativity have either had a process leniting the same segment, or did not allow the segment in word-final positions.

Study 3 used the Buckeye Corpus to predict word-final deletion of all post-vocalic pre-consonantal American English obstruents while controlling for phonological features, rate of speech, and word frequency. Word and speaker were used as random effects. Low informativity significantly promoted word-final deletion in English ($p<.01$). The contextual predictability of segments did not contribute to predicting word-final deletion, providing support to the *exceptionless* properties of sound change.

All three studies show that when a segment's informativity in a given language is relatively low, that segment is more likely to weaken. The first two studies compared languages to an estimate of the cross-linguistic average, and the last investigated the effects of informativity on a richer data set while controlling for a number of phonological factors. The results suggest that higher information motivates stability, and low information translates to propensity to weaken.

References

Cieri, Christopher, Miller, David, and Walker, Kevin. 2004. The Fisher corpus: a resource for the next generations of speech-to-text. In *Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 69–71.

Cohen Priva, Uriel. 2008. Using information content to predict phone deletion. In Abner, Natasha and Bishop, Jason, editors, *Proceedings of the 27th West Coast Conference on Formal Linguistics*, pages 90–98, Somerville, MA. Cascadilla Proceedings Project.

Cohen Priva, Uriel. 2015. Informativity affects consonant duration and deletion rates. *Laboratory Phonology*, 6(2):243–278. doi: 10.1515/lp-2015-0008.

Cohen Priva, Uriel. to appear. Informativity and the actuation of lenition. *Language*. URL http://urielcpublic.s3.amazonaws.com/Informativity_Actuation-resubmitted.pdf.

Godfrey, John J. and Holliman, Edward, 1997. Switchboard-1 release 2. Linguistic Data Consortium, Philadelphia.

Guy, Greogry. 1991. Explanation in variable phonology: an exponential model of morphological constraints. *Language Variation and Change*, 3(1):1–22.

Hochberg, Judith G. 1986. Functional compensation for /s/ deletion in Puerto Rican Spanish. *Language*, 62:609–621.

Hume, Elizabeth. 2008. Markedness and the language user. *Phonological Studies*, 11.

Kiparsky, Paul. 1995. The phonological basis of sound change. In Goldsmith, John A., editor, *The Handbook of Phonological Theory*, pages 640–670. Blackwell Publishers, Cambridge, MA.

Ohala, John J. 2003. Phonetics and historical phonology. In Joseph, Brian D. and Janda, Richard D., editors, *The Handbook of Historical Linguistics*, pages 669–686. Blackwell.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, Joan and Hopper, Paul, editors, *Frequency and the Emergence of Linguistic Structure*, pages 137–157. John Benjamins Publishing Company.

Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., and Fosler-Lussier, E., 2007. Buckeye corpus of conversational speech (2nd release). Department of Psychology, Ohio State University.

Scannell, Kevin P. 2007. The Crúbadán project: Corpus building for under-resourced languages. In *Building and Exploring Web Corpora: Proceedings of the 3rd Web as Corpus Workshop*, volume 4, pages 5–15.

Soderberg, Craig D. and Olson, Kenneth S. 2008. Indonesian. *Journal of the International Phonetic Association*, 38(02):209–213. doi: 10.1017/S0025100308003320.

Sóskuthy, Márton. 2015. Understanding change through stability: A computational study of sound change actuation. *Lingua*, 163:40–60. doi: 10.1016/j.lingua.2015.05.010.

Wedel, Andrew, Kaplan, Abby, and Jackson, Scott. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, 128(2):179–186. doi: 10.1016/j.cognition.2013.03.002.

Weinreich, Uriel, Labov, William, and Herzog, Marvin I. 1968. Empirical foundations for a theory of language change. In Lehmann, Winfred P. and Malkiel, Yakov, editors, *Directions for Historical Linguistics*, pages 95–18. University of Texas Press, Austin.