

Probabilistic Inference and Frequency Effects in Language Change

Vsevolod Kapatsinski (University of Oregon)

vkapatsi@uoregon.edu

Language change often involves the gradual encroachment of an innovative variant form on the contexts previously occupied by another form. For example, in American English, *going to* is encroaching on *will* in the context of future marking, the pronunciation *-in'* is encroaching on the contexts that used to favor *-ing*, and the flap [r] has encroached on [t] in post-tonic intervocalic contexts.

Much research has shown that words that often occur in contexts that favor the innovative variant become associated with that variant, so that the innovative variant becomes likely to be used with such words even outside of contexts that otherwise favor it (Bybee, 2002; Brown, 2004; Forrest, 2017; *inter alia*). For example, the [n] at the end of *-ing* is favored by informal speech style, and following coronal consonants. However, words that frequently occur in informal styles and before coronal consonants favor the [n] pronunciation even when they are used in a formal style and before a vowel. Following Brown, this is usually called the frequency in favorable contexts (FFC) effect. FFC effects are found to be stronger in frequent words (Forrest, 2017). The present work studies the conditions under which FFC and its interaction with frequency emerge from rational probabilistic inference.

We make use of a recently developed model of sound change in which online reductive pressures are combined with rational probabilistic hierarchical inference, which distributes credit for a pronunciation between the sublexical unit undergoing the change (the linguistic variable), and the larger lexical units (words and phrases) that contain it (Kapatsinski, 2021, in press). The model implements cycles of production and learning across generations. Production is assumed to have a bias in favor of the innovative variant, such that every token of a word's use increases the probability of selecting the innovative variant. The resulting productions constitute a corpus from which the next generation learns the language. Each learner is assumed to infer a hierarchical logistic regression model in which words, phrases and sublexical units are nested random effects (see also Vetchinnikova, 2024), alongside fixed effects of context and grammar. I report on an extension of the model in which 1) the production pressure for the innovative variant is stronger in innovation-favoring contexts and may be reversed in innovation-disfavoring contexts, 2) each word has a probability of occurring in a reduction-favoring context, sampled from a beta distribution, and 3) learners may not detect the relevant context in a particular token. As in the original model, word frequencies are Zipfian-distributed.

The FFC is shown to emerge only when learners misattribute (some of) the effect of context to lexical idiosyncrasy. That is, it emerges only if learners are prone to missing the context that an observed word token occurs in (faulty perception), or do not take the influence of context into account when building their mental model of variant choice (imperfect learning). The model also provides an interesting direction for future work on when FFC effect do and do not emerge. Specifically, the distributions of words across innovative and conservative contexts have yet to be studied. Yet, the model suggests that these distributions are crucial because the FFC effect emerges only when words have polarized distributions across reduction-favoring and disfavoring contexts (top row of Figure 1 vs. bottom row).

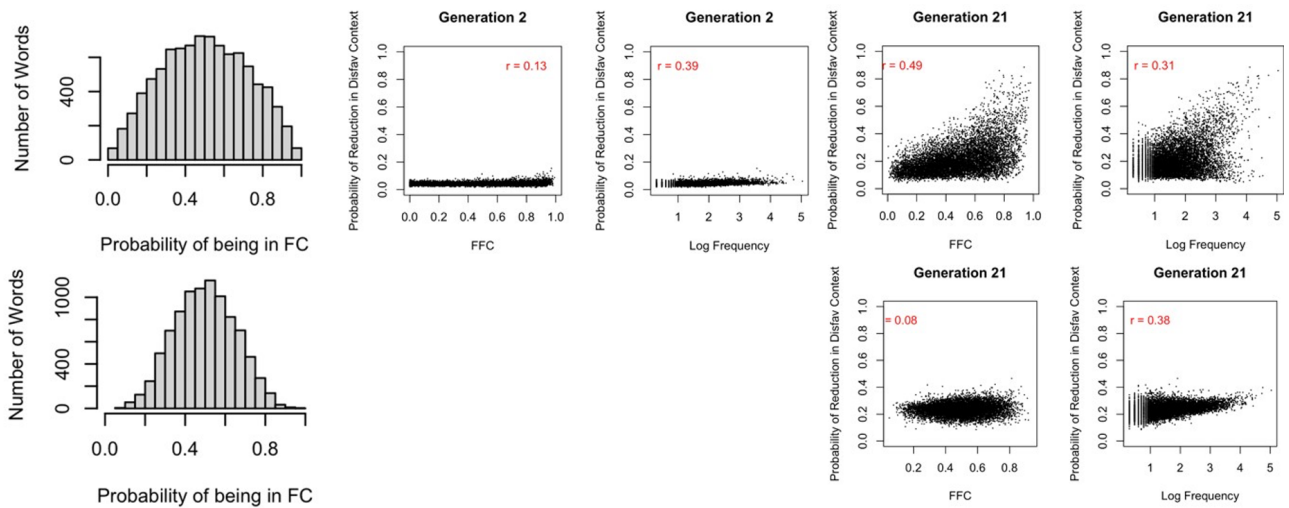


Figure 1. The effect of frequency in favorable context (FFC) emerges over the course of the sound change (by Generation 21) in the top row where the words' probabilities of being in a favorable context are variable enough. Starting from the same initial point (Generation 2), it does not emerge in the bottom row, where the words' distributions across contexts are less variable.

References:

- Brown, E. L. (2004). The reduction of syllable-initial/s/in the Spanish of New Mexico and southern Colorado: A usage-based approach. Ph.D. Dissertation, University of New Mexico.
- Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14(3), 261-290.
- Forrest, J. (2017). The dynamic interaction between lexical and contextual frequency: A case study of (ING). *Language Variation and Change*, 29(2), 129-156.
- Kapatsinski, V. (2021). Hierarchical inference in sound change: Words, sounds, and frequency of use. *Frontiers in Psychology*, 12, 652664.
- Kapatsinski, V. (In press). Lexical frequency and diffusion. In A. Ledgeway, E. Aldridge, A. Breitbarth, K. É. Kiss, J. Salmons and A. Simonenko (Eds.), *The Wiley Blackwell Companion to Diachronic Linguistics*. Wiley.
- Vetchinnikova, S. (2024). Idiosyncratic entrenchment: tracing change in constructional schematicity with nested random effects. *Corpus Linguistics and Linguistic Theory*.

