Principal components of lexical prominence in Munster Irish: analysis of nonword data

Connor McCabe[1]
[1]*School of Irish, Celtic Studies & Folklore, University College Dublin (Ireland)*

Munster Irish varieties (MI) are traditionally described as placing lexical prominence/'stress' on a non-initial non-light syllable (e.g. [3]). Syllables containing short vowels are light (L), those with long vowels or diphthongs are heavy (H), and some suggest a medium-heavy status for the sequence /ax/ (e.g. [2]). Limited work has examined phonetic correlates of stress location as traditionally described (e.g. [4]). However, a fundamental problem exists with the impressionistic identification of stress, given cross-linguistic diversity in lexical and phrasal prominence-marking and phonological significance thereof. As part of a larger investigation of prominence production in MI, a nonword experiment was devised to test productivity of weight-sensitive stress. Here, a principal component analysis (PCA) is used for these nonword data to examine correlation(s) amongst potential phonetic exponents of stress. This is a key step in developing empirically grounded phonetic stress diagnostics for Irish varieties independent of impressionistic description.

12 native speakers of MI (9 female, 3 male; age range 20-76) representing the three regional subvarieties of Counties Kerry, Cork, and Waterford were presented with 36 di- and trisyllabic nonwords, with each item produced five times per participant. Stimuli were devised using permutations of (*b*)*a* (/bʸa/; light), (*b*)*á* (/bʸɑː/; heavy), and (*b*)*ach* (/bʸax/, here considered light), and presented in prenuclear/pre-focal position in a carrier phrase. After discarding errors, a total of 1728 tokens were obtained. For each of 4688 syllables, maximum intensity, maximum F0, F0 range, vowel duration, mean F1, and mean F2 were extracted. F1 and F2 were only considered statistically for a subset of vowel-quality controlled targets, not reported on here.

First, 10 Bayesian mixed-effect linear regressions were constructed in R with `brms` [1], which examined change in each prominence measure for di- and trisyllables as a function of syllable position, with a random slope included for target weight-structure. Target-specific estimates for cross-syllable change in each measure were then extracted from the models and subjected to a PCA using `prcomp()`. The first five principal components (PCs) account for 97% of variance in the data, with 71% accounted for by PCs 1 and 2 alone (48% and 23%; loadings in Figure 1). PC1 is dominated by cross-syllable changes in maximum intensity, being contributed to by increases in this measure from first to second and/or final syllables. PC2 relates positively to increase in vowel duration (measured relative to contrastive length category) between first and second/final syllables. The two PCs show virtually no correlation.

PC1 patterning aligns with presence and distribution of heavy syllables in a target (Figure 2), something not evident for PC2. This is consistent with granular analysis of the initial models, in which (i) intensity change most closely aligned with described stress location, (ii) vowel duration changes seemed to relate to sporadic rallentando, and (iii) changes in F0 height and range did not systematically distinguish nonword structures. By-target analysis of scores for PCs 1 and 2 shows only partial alignment with previous descriptions of stress placement. Most consistently, targets with second-position heavy syllables show relatively high PC1, versus low PC1 for structures with no heavy syllables beyond initial position (if any). Notably, many targets expected to receive non-initial stress (such as LL**H** *baba**bá***, *baba**chá***, *bacha**bá***, and *bacha**chá***) show ambiguously low PC1.

It is planned to collect equivalent production data from other Irish varieties, as well as data on perception of prominence. In the meantime, these preliminary findings indicate intensity as the feature most robustly manipulated by MI speakers in applying prominence to unfamiliar (pseudo)lexical targets, encourage scepticism of received descriptions, and underscore the need for empirical evaluation of stress phenomena in Irish and cross-linguistically.

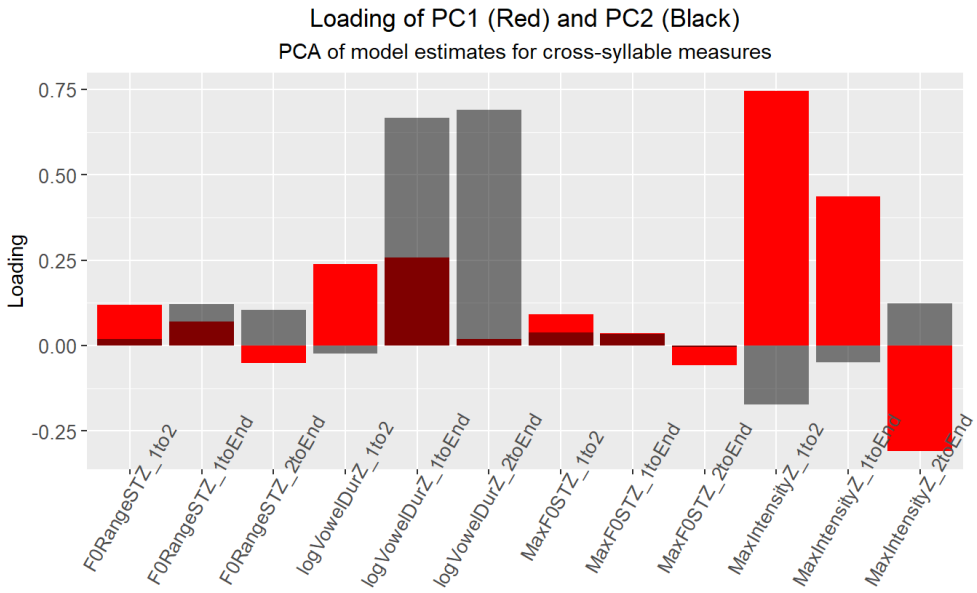

Fig. 1. *Loadings of the first and second principal components across the 12 cross-syllable prominence measures extracted from the nonword data.*

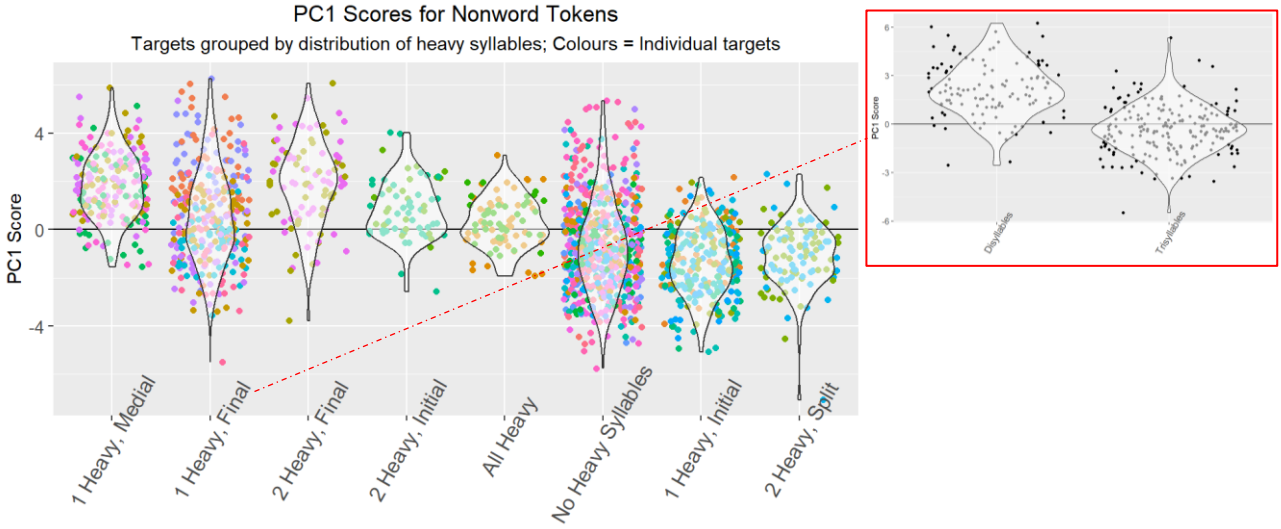

Fig. 2. *PC1 scores for the 1728 nonword tokens, grouped by distribution of heavy syllables and arranged in descending order of maximum score. Within* 1 Heavy, Final *(see inset), light-light-heavy trisyllables are mostly negative and light-heavy disyllables are mostly positive.*

References

[1] Bürkner, P.-C. 2017. "brms: An R package for Bayesian multilevel models using Stan." *Journal of Statistical Software* 80(1), 1-28.

[2] Doherty, C. 1991. "Munster Irish stress". *Proceedings of the 10th West Coast Conference on Formal Linguistics*, 115-126.

[3] O'Rahilly, T.F. 1932. *Irish Dialects Past & Present*. Dublin: Browne & Nolan.

[4] Windsor, J.W., S. Coward & D. Flynn. 2018. "Disentangling Stress and Pitch Accent in Munster Irish". *Proceedings of the 35th West Coast Conference on Formal Linguistics*, 430-437.