INTRODUCTION

Phonotactic probability (PP) and phonological neighborhood density (ND) have been claimed to be relevant for a wide range of phenomena:

- acceptability judgments (Bailey & Halle 2001)
- spoken word recognition (Vitevitch & Luce 1999)
- speech production errors (Vitevitch 1997) and latencies (Vitevitch 2002)
- word learning (Storkel 2001, Freedman & Barlow 2012)
- hyper- and hypo-articulation (Wright 2004) and coarticulation

Because there are many ways of quantifying PP and ND, and because they are highly correlated in natural lexicons (Vitevitch et al. 1999), it is challenging to disentangle their effects on a given aspect of performance (but see Luce & Large 2001, Pykkanen et al. 2002, Storkel et al. 2008, 2012).

PRODUCTION STUDY 1 (SCARBOROUGH 2013)

English participants (N=10) produced C′VC′N (and C′NVC′) monosyllables within carrier sentences to listeners in a dictation task.

<table>
<thead>
<tr>
<th>higher ND (hard)</th>
<th>lower ND (easy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VN heart (24)</td>
<td>champ (24)</td>
</tr>
<tr>
<td>NV must (24)</td>
<td>nose (24)</td>
</tr>
</tbody>
</table>

Items were all of high Hoosier familiarity (Nusbaum et al. 1984) and balanced for lexical frequency, vowel, and nasal consonant across conditions.

In the experimental design, neighborhood density $ND_{\text{nn}}$ was defined as the inverse of $R_{\text{nn}} = \text{binarized log}_{2}(1+\text{number of neighbors})$, where $N'(w)$ is single-edit neighborhood.

Degree of nasal coarticulation on the N-adjacent vowel was quantified by $A1-P0$: amplitude of F1 - amplitude of low frequency nasal peak (Chen 1997) NS. Greater vowel nasality damps A1 and increases P0, resulting in lower A1-P0.

STATISTICAL REANALYSIS

Results of the production study were previously analyzed with separate repeated-measures ANOVAs, here replicated with a single mixed-effects linear regression:

$$\text{fixed} \ A1-P0 \sim \text{Set (VN/VN)} + ND_{\text{nn}} + Vowel + Nasal$$
$$\text{random} \ + (1 + \text{Set } ND_{\text{nn}} + \text{speaker}) + (1 + \text{word})$$

with outliers removed ($\pm 2.5$ sd within participant $\times$ stimulus type), dependent variable centered, and binary predictors scaled to mean 0 and sd 0.5.

Significant effect of $ND_{\text{nn}} (\beta = -0.25, \text{SE} = 0.08, t = -3.33)$ and of Set ($\beta = -0.27, \text{SE} = 0.126, t = -2.20$, no sig. interaction). Also significant effects of nasalized vowel (e.g., /ax/ vs. /eI/), $\beta = -0.60, t = -3.88$, and nasal consonant (/j/ vs. /β/ = 0.46, $t = -3.07$), which were not previously investigated but do not eliminate the effect of ND.

DOES NASAL COARTICULATION TRACK TREND?

- Could phonotactic probability provide an alternative account of the coarticulation data in spite of item matching? (suggestion by Baese-Berk & Goldrick 2009)
- More generally, can species of lexical and sublexical factors be distinguished by their effects on phonetic realization? (see also Baese-Berk & Goldrick 2009, Galé et al. 2012)

VARIANTS OF ND AND PP

A large number of ND variants were calculated for the items of the production study from CLEARFOND (Maranon et al., 2012; clearfond.northwestern.edu):

- count of all phonological neighbors (PTAN) as well as those differing from the target by only substitution (PSAN), deletion (PDAN), or addition (PAAN)
- mean frequency for each phonological neighbor type (PTAF, PSAN, PDAN, PAAF)

The PP variants were maxent harmony scores computed from BLICK (Hayes 2012), Onset and Rime relative frequencies across word types (Cohen & Pirennebelieumont 1997) from the same CMU dictionary used to train BLICK (Wade, 1999; www.speech.cs.cmu.edu/cgi-bin/cmudict).

Positional n-gram sums and averages (Bosch & Luce 1994), and margin complexity (NC′ or C″).

Unsurprisingly many of the measures are highly correlated, making discrimination among them with this set of items statistically challenging.

Random forest regression (Breiman 2001) used to identify the numerical variant of each predictor type with maximal importance (inspired by the method of Burki et al. 2011)

Random forest evaluation & model comparison

Low correlation between the two best representatives of the lexical (PTAF) and sublexical (avg bigram probability) types: $r = 0.089$ n.s., and each survives a Likelihood Ratio Test for nested models ($p < 0.05$).

PTAF $\chi^2(1) = 5.51$, bigram avg $\chi^2(1) = 5.16$

PTAF and bigram avg together achieve a log-likelihood (−1056.13) similar to that for the original predictor $ND_{\text{nn}}$, alone (−1056.578).

Discussion

- Comparison of PP and ND contributes to the general problem of distinguishing sublexical / structural vs. lexical / holistic accounts of linguistic sound patterns. (e.g., Allbright & Hayes 2001, Burki & Gaskell 2012, Sadat et al., 2014, Becker et al.)
- Lexical properties (ND) condition phonetic realization in at least some cases, acting alongside sublexical properties (PP) to determine coarticulation & hyperarticulation.
- High lexical correlation of PP and ND motivates the search for converging evidence across studies, and statistical tools for predictor selection and model comparison: methodologies common to many other fields (e.g., Gourg & Elison 2003, Muyong & Pat 2011).