Articulatory strategy as a source of variation in acoustic vowel dynamics

Patrycja Strycharczuk[1], Justin J. H. Lo[2] & Sam Kirkham[2]
University of Manchester[1] & Lancaster University[2]

Extant theories of individuality in speech propose that speaker-specific articulatory targets determine specific movements between those targets, which can result in speaker-specific formant dynamics [3, 4, 5]. While influential and consistent with some forensic evidence, this idea has not yet been directly evidenced. Our study presents explicit evidence in support of it, by demonstrating a systematic link between the articulatory strategy for palatal vowel production and the shape of formant transitions in diphthongs with a palatal offglide, produced by the same speakers.

Data The study is based on synchronised ultrasound and audio data from 36 speakers of Northern English from a pre-existing corpus [7]. The stimuli were *bead, bade, bide* and *buoyed*.

Analysis We focus on the articulatory strategy for /i/ production (in *bead*), because /i/ articulation is known to be highly individual, due to anatomical variation in palate shape [1]. We parametrised the mean /i/ tongue shape for each speaker, using a morphometric analysis that involves Procrustes normalisation and principal component analysis [2]. This yielded three PCs that, combined, captured 78% of variance in tongue shape across speakers. We used the PCs as predictors in Generalised Additive Mixed Modelling of normalised F1 and F2 trajectories in /eɪ/, /aɪ/ and /oɪ/. The GAMMs were built using maximal random effect structure.

Results We find that speakers who share specific tongue shape features in their /i/ production also share some features of their acoustic formant trajectories in selected I-diphthongs. Speakers who have a more anterior constriction, or more domed tongue shape in /i/ (high PC1, low PC2) produce earlier formant transitions (especially F2) between onglide and offglide in /eɪ/ and /aɪ/. This relationship holds after tongue size is normalised by scaling. We also find that the speaker's mean tongue shape in /i/ is very strongly correlated with their tongue shape in I-diphthong offglides. We find no significant effect of the PCs on vowel duration.

Discussion We propose that speakers use consistent strategies for producing palatal vowels, including /i/ and /ɪ/. These strategies link articulatory movement with resulting tongue shape in a systematic way. Speakers with more tongue doming or more anterior constriction in palatal vowels need to traverse a larger section of the vocal tract when producing diphthongs such as [aɪ]. Given a fixed time window, such speakers need to produce higher velocity, and faster movement at the onset of the palatal gesture. This behaviour is consistent with the predictions of the linear harmonic oscillator model [6]. The described properties of articulatory movement cause the observed acoustic differences: higher velocity translates into faster and steeper formant transitions.

Conclusion Our findings support the link between vocal tract constraints on the one hand and individual differences in vowel dynamics on the other, also highlighting the limits of articulatory compensation: while articulatory variation is largely conditioned by compensation, some aspects of the relevant variation are acoustically recoverable.

References [1] A. Lammert, M. Proctor, and S. Narayanan. "Interspeaker variability in hard palate morphology and vowel production". In: *J. Speech Lang. Hear. Res.* 56.6 (2013). [2] J. J. H. Lo, P. Strycharczuk, and S. Kirkham. "Articulatory strategy in vowel production as a basis for speaker discrimination". In: *Proc. Interspeech 2025*. 2025. [3] K. McDougall. "Dynamic features of speech and the characterization of speakers: Toward a new approach using formant frequencies". In: *Int. J. Speech Lang. Law* 13.1 (2006). [4] K. McDougall. "Speaker-specific formant dynamics: An experiment on Australian English /aɪ/". In: *Int. J. Speech Lang. Law* 11.1 (2004). [5] F. Nolan and C. Grigoras. "A case for formant analysis in forensic speaker identification". In: *Int. J. Speech Lang. Law* 12.2 (2005). [6] E. L. Saltzman and K. G. Munhall. "A dynamical approach to gestural patterning in speech production". In: *Ecol. Psychol.* 1.4 (1989). [7] P. Strycharczuk et al. "Dimensionality reduction in lingual articulation of vowels: Evidence from lax vowels in Northern Anglo-English". In: *Lang. Speech* 68.3 (2025).