Dynamical modelling of vowel diphthongisation

Sam Kirkham[1], Patrycja Strycharczuk[2] and Emily Gorman[1]
[1]*Lancaster University (UK)*, [2]*University of Manchester (UK)*

The variable diphthongisation of vowels in English is a widely attested form of synchronic variation, such as the monothongisation of GOAT and PRICE in the dialects of Northern England [1], as well as diphthongisation of tense monophthongs, such as FLEECE and GOOSE [2, 3]. Diphthongisation also underlies many diachronic sound changes, such as the development of high vowels into diphthongs during the English Great Vowel Shift [4]. While these descriptive facts of vowel variation are well known, it remains challenging to provide a convincing account of vowel diphthongisation that can capture the wide range of gradient synchronic variation in dialects and the apparently categorical shifts of long-term sound change. In this paper, we develop a theoretical account of vowel variation and change, grounded in a dynamic neural field account of speech planning [5, 6] and a task dynamic model of articulatory execution [7,8].

We first model all long vowels as containing two targets [9, 10, 11]. In this view, a long monophthong is long because it is comprised of two sequential gestures with identical targets. A diphthong has the same underlying structure (two targets), but different target parameters. We illustrate this using task dynamic simulations based on the model in [7], specifying a vowel as two concatenated gestural activation intervals of 250 ms in duration, which are coupled anti-phase. Our model predicts that variation between a long monophthong and a diphthong can be captured via gradient variation in the nucleus target (Figure 1).

Our second analysis focuses on what needs to change in the phonological representations of individual speakers for gradual sound change to occur. We advance a dynamic neural field (DNF) model [12] of phonological planning. A DNF model situates phonological planning in an activation field over a range of phonetic parameters [13, 14, 5, 6]. A dynamical equation specifies the evolution of field activation until some value reaches a threshold, which is then selected as the parameter value for speech production. We then model production and perception as inputs to the field and track how the field develops over time.

We show that such simulations also generate a gradient continuum between a monophthong and a diphthong, providing a clear mechanism for variation and change. Following from this, our DFT model then defines /i/ as two planning fields (one for the nucleus, one for the offglide).We model production-perception as (i) a speaker producing a value from their activation field; (ii) hearing a speaker whose nucleus has a phonetic bias towards /e/; (iii) this perceived token is integrated into memory with a small amount of noise; (iv) this process repeats [13]. After a number of interactions with this ‘biased’ speaker, the activation field shifts away from the initial state (representing an /i/ nucleus) towards a new peak (representing an /a/-like nucleus), as in Figure 2. We simulate articulatory trajectories based on these activation fields and show that /i/ eventually changes into /ai/, with no recourse to categorical rules. Specifically, when the vowel nucleus moves away from its initial state towards a different state, the outcome of the task dynamic equation is more likely to be a diphthong.

Our model combines an autonomous model of gestural dynamics with a dynamic field planning model in order to simulate the processes of synchronic and diachronic vowel diphthongisation. We propose that the accumulation of gradient variation in targets can lead to long-term sound changes that look like categorical changes over historical time. In conclusion, we identify a shared mechanism for synchronic variation and diachronic change in vowels – gradient variation in gestural targets – and propose a mechanism for how individual phonological representations change. We also discuss potential variability in both vowel targets as part of a broader stochastic model of sound change.

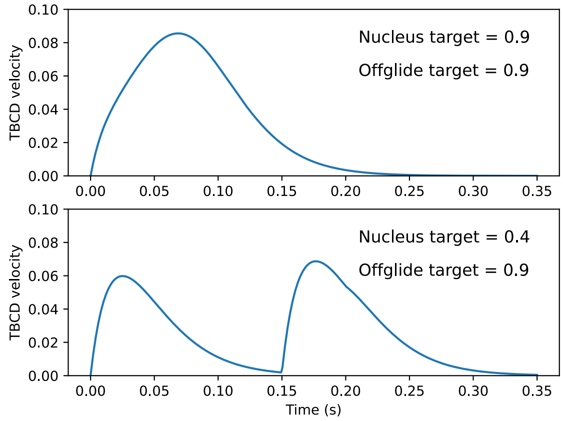

Fig. 1. Velocity trajectories of simulated two-target vowels, with identical (top) and different (bottom) targets.

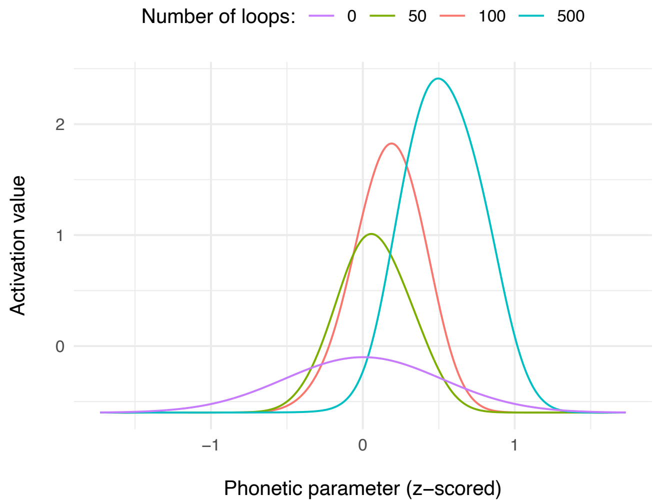

Fig. 2. DFT activation memory field for one speaker's vowel nucleus after different numbers of repeated production-perception loops with a biased interlocutor.

References

[1] Hughes, A., P. Trudgill & D. Watt. (eds) (2012). *English Accents and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles*. London: Hodder.

[2] Strycharczuk, P., M. López-Ibáñez, G. Brown & A. Leemann (2020). General Northern English: Exploring regional variation in the north of England with machine learning. *Frontiers in Artificial* Intelligence 3(48): 1-18.

[3] Wells, J.C. (1982) *Accents of English: Volumes 1–3*. Cambridge: Cambridge University Press.

[4] Jespersen, O. (1909). *A Modern English Grammar on Historical Principles*. London: George Allen & Unwin Ltd.

[5] Roon, K.D. & A.I. Gafos (2016). Perceiving while producing: Modeling the dynamics of phonological planning. *Journal of Memory and Language* 89(2): 222-243.

[6] Tilsen. S. (2019) Motoric mechanisms for the emergence of nonlocal phonological patterns. *Frontiers in Psychology* 10: 1-25.

[7] Sorensen, T. & Gafos, A.I. (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology* 28(4): 188-215.

[8] Saltzman, E. & K.G. Munhall (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology* 1(4): 333-382.

[9] Labov, W., S. Ash & C. Boberg (2006). *The Atlas of North American English: Phonetics, Phonology and Sound Change*. Mouton de Gruyter.

[10] Popescu, A. & I. Chitoran (2022). Linking gestural representations to syllable count judgements: A cross-language test. *Laboratory Phonology* 13(1): 1-48.

[11] Strycharczuk, P., S. Kirkham, E. Gorman & T. Nagamine (submitted). Towards a dynamical model of English vowels: Evidence from diphthongisation.

[12] Schöner, G., J.P. Spencer & The DFT Research Group (2016). *Dynamic Thinking: A Primer on Dynamic Field Theory*. Oxford: Oxford University Press.

[13] Kirov, C. & A.I. Gafos (2007). Dynamic phonetic detail in lexical representations. *Proceedings of the 16th International Congress of Phonetic Sciences* 637-640.

[14] Shaw, J.A. & K. Tang (2023). A dynamic neural field model of leaky prosody: proof of concept. *Proceedings of the Annual Meeting on Phonology* 1-12.