Using acoustic-phonetic simulations to model historical sound change

Toby Hudson, Jonathan Wei, and John Coleman
University of Oxford

Arising from [1], we employ audio analysis and synthesis methods to model and try to understand processes of sound change. Using audio recordings and synthetic speech simulations frees us from some restrictions inherent in text, e.g. the limitations of discrete symbols to capture phonetic variation or gradient change. Preliminary work in statistical phylogenetic modelling of acoustic sound change [2] has shown only moderate success in reconstructing ancestral pronunciations, partly because of the high dimensionality and many degrees of freedom in spectrographic representations; our subsequent experiments (e.g. [3], [4]) have therefore focussed on the more tractable task of modelling paths of acoustic change in between cognates in paired languages.

Main technique. Audio recordings of a pair of words, w_1 and w_2, are transformed into spectral representations s_1 and s_2 (e.g. Fourier spectrograms, LPC spectral vectors, MFC cepstrograms; we refer to all of these as "spectrograms"). Spectrograms s_i intermediate between s_1 and s_2 are computed at e.g. 5% or 10% increments; s_1 … s_i … s_2 are then converted back to audio files, yielding a continuum of sound files exhibiting gradually-changing pronunciations w_i, on a continuum in between w_1 and w_2. When w_1 resembles a pronunciation hypothesized as an older form and w_2 captures a later pronunciation, the sequence of interpolants w_i simulates a path of sound change. If sound change is hypothesized to arise from misperception or from phonological reanalysis by a new generation, acoustic simulations can be used as stimuli for use in e.g. categorical perception experiments such as [5].

Case studies. (1) Author 3 simulates sequences of sound change between Proto-Indo-European, Modern English, and other languages, e.g. the vocalic reflexes of the "laryngeals" in Greek, vs. their loss in other branches. (2) Author 2 uses simulation techniques to address the problem of vowel breaking (diphthongization) in Old English [6: 16-20], creating sound-change continua from e.g. Proto-Germanic **alda* to Old English *eald* 'old' in order to determine the acoustic-phonetic parameters involved in the diachronic development of the Proto-Germanic monophthongs <æ e i> to Old English diphthongs <ea eo ie>. While previous studies of short diphthongs suggest various phonetic plausibilities based upon textual evidence, author 2 recreates this sound change using modelled continua to examine the perception of vowel breaking in a lab setting. (3) Author 1 uses simulation techniques to make trajectories for a) pairs of words in Italian which differ by stress placement, such as pássero ('sparrow') ~ passeró ('I shall pass'), and b) the same, re-synthesized using the *pitch synchronous overlap and add* technique to remove pitch movement and any durational differences. A 2-forced-choice perception experiment examines how Italian listeners categorize a gradient of pitch peaks into perceptually discrete patterns. Results are interpreted with respect to diachronic left-to-right-edge shift in word stress, as in Old to Classical Latin (e.g. *fárci:re* to *farcí:re* 'to stuff'; see [7: 92-4] for details).

Critical appraisal of the utility of simulations. Simulation methods enable further interesting questions to be addressed. We find that the trajectories of simulated sound change are sometimes consistent with philologically attested intermediate forms; where they are not, we must reexamine the model assumptions. We can measure the acoustic-phonetic directions and magnitude of changes, which shows that divergent, parallel or convergent paths of development are indeed divergent, parallel or convergent in acoustic-phonetic space. From the magnitudes of change and estimates of the time period over which changes took place, we measured the rates of various sound changes. From these we infer that a

"punctuated equilibrium" model of sound change (jumps in between long periods of stability) is more plausible than slow, gradual drift. Our perception experiments show that the changes examined so far need to become quite large to be perceived, and that gradual continua are perceived categorically.

Keywords. Simulation; phonetics; sound change; Indo-European; acoustics; perception experiment

References

[1] The Functional Phylogenies Group (J. A. D. Aston, D. Buck, J. Coleman, C. J. Cotter, N. S. Jones, V. Macaulay, N. MacLeod, J. M. Moriarty, and A. Nevins). 2012. Phylogenetic inference for function-valued traits: speech sound evolution. *Trends in Ecology and Evolution*, 27 (3), 160–166. http://dx.doi.org/10.1016/j.tree.2011.10.001

[2] Hadjipantelis, P. Z. 2013. *Functional Data Analysis in Phonetics.* PhD Thesis. Department of Statistics, University of Warwick.

[3] Coleman, J., J. Aston and D. Pigole. 2015. Reconstructing the sounds of words from the past. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: the University of Glasgow. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0296.pdf

[4] Pigoli, D., P. Z. Hadjipantelis, J. S. Coleman and J. A. D. Aston. 2018. The statistical analysis of acoustic phonetic data: exploring differences between spoken Romance Languages. *Journal of the Royal Statistics Society, Series C (Applied Statistics)* 67, Part 5, 1103–1145.

[5] Beeley, H. 2014. Kwod Erat Demonstrandum: an empirical investigation of labiovelar historical developments. *SCIHS (Sound Change in Interacting Human Systems): 3rd Biennial Workshop on Sound Change*. 28-31 May. University of California, Berkeley. http://linguistics.berkeley.edu/SCIHS/abstracts/3_ThursdayPosters/Beeley.pdf

[6] Hogg, R. 1992. *A Grammar of Old English, Volume 1: Phonology*. Malden, MA: Wiley-Blackwell.

[7] Clackson, J. P. T. and G. Horrocks. 2007. *The Blackwell History of the Latin Language*. Malden, MA: Wiley-Blackwell.