Acoustic-phonetic modelling of historical and prehistoric sound change

John Coleman, University of Oxford
john.coleman@phon.ox.ac.uk

The comparative method reconstructs hypothesised ancestral forms of words based upon their derived modern and/or historically-attested written forms. This method is governed by four leading principles: (a) common traits in two or more daughter languages, if not loans or coincidences, are attributed to descent from a shared ancestral form; (b) the processes of sound-change by which the daughter languages are derived from the ancestor apply across the lexicon, and (c) are phonologically natural; (d) inversion of the sound-change rules ought to converge on reconstructed transcriptions of words in the past. For some language families we are fortunate to have historical texts as well as modern spoken languages; for others, the evidence is limited to transcriptions of modern languages.

Developments in computational modelling of language phylogenies, inspired by computational models of biological evolution, enable hypotheses concerning ancestral forms, sound change rules, phylogenetic trees and chronologies to be tested (e.g. Ringe *et al*. 2002, Nakhleh *et al*. 2005, Bouckaert *et al*. 2012, Bouchard-Côté *et al*. 2013, Chang *et al*. 2015). In all of those works, the data used is text: alphabetic transcriptions of words and/or features representing linguistic traits. Recently, substantial progress has been made in phylogenetic modelling of how continuous functions (curves and surfaces) change and diverge across generations (The Functional Phylogenies Group 2012, Jones and Moriarty 2013, Hadjipantelis *et al*. 2013), modelling size, shapes, growth rates etc. As signal parameters of spoken words such as formant frequencies, amplitude contours, or surfaces such as spectrograms can be represented using continuous functions, it is becoming possible to model linguistic history and prehistory by these methods. This makes it possible to reconstruct audible sound files instantiating hypothesized possible spoken forms from the past, including distant ancestral pronunciations and the intermediate forms at each generation.

Though we are only just starting to realise such a possibility, many practical challenges must be addressed. Over the last several years, our group has proposed methods for:

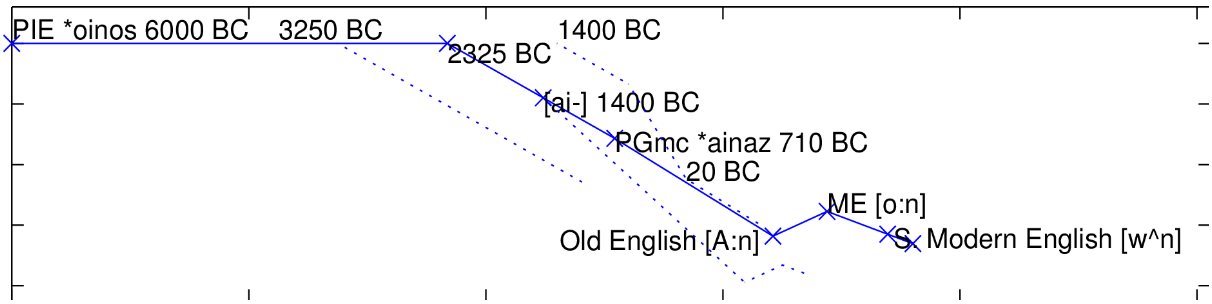
Input data, and its acoustic representations. Our data is recordings of discrete words across the Indo-European languages. As there is considerable variation across speakers and dialects, we constrained the vocabulary to be a small set of citation forms of cognate terms (words for the digits "one" to "ten") covering a variety of phonological patterns. To generate audible examples of putative ancestral forms derived by transformations from this data, acoustic representations are needed that are invertible, so that audible reconstructions may be synthesized. We have examined Fourier spectrograms, linear prediction reflection coefficients and Mel-Frequency Cepstral Coefficients for analysis–transformation–resynthesis, and shall discuss the relative merits of these.

Ancestral forms. As sound recordings from the distant past do not exist, we must base our analyses of sound change on proxies of several kinds: (a) Use modern recordings that resemble how ancestral pronunciations are thought to have been, e.g. the *un-* portion of Spanish or Italian *uno* is a proxy for the corresponding part of Latin *unus.* (b) Compute or (c) edit hybrid forms from two modern recordings. Thus, we generate Proto-Indo-European **dwoh* from the continuum between Elfdalian [two] and a Russian token of [dva] with laryngealized offset; the first syllable of Italian *quindici* spliced with the second syllable of *cinque* simulates Latin *quinque.* (d) Use statistical regression over a phylogenetic tree to extrapolate back from modern recordings to a hypothetical ancestral form (Hadjipantelis 2013).

Statistical operations over acoustic representations. Statistical acoustic phylogenetics needs methods for time-registration (dynamic time warping), smoothing, and computation of central tendencies and (co)variances. Interpolation of intermediate forms in between a putative ancestral proxy and a modern form provides a method for modelling sound change over arbitrarily many generations. These are sometimes testable against observations or other sources of data, such as texts from the past. For example, the continuum from [ũ] (a supposed successor of Latin *un-*) to French [ɛ̃] includes intermediate stages like [œ̃], an attested, conservative French prounciation. But linear interpolation between 'Latin' [tre:s] and French [tʁwa] does *not* pass through intermediate [trois], even though a form containing a diphthong [oi] and a final [s] is suggested by the historical spelling; linear interpolation is not a good model of sound change in this case.

Modelling the changing rate and direction of sound changes. Many previous computational models of sound change assume that the rate of change is approximately the same everywhere and at all times. This is evidently not correct, as some words in some languages are more conservative whereas others are more innovative. For example, Lithuanian *penki* is much closer to Proto-Indo-European **penkwe* than is the English word *five*. Gradual, incremental modelling of sound changes in the acoustic domain enables us to estimate the varying rate at which words have changed over millennia. We characterise the direction of sound changes using cosines of spectral vectors; for example, the maximum spectral distance between Old English [a:n] and Middle English [o:n] is about –24° (minus because it is in the opposite direction from Proto-Indo-European **oin(os)* to [a:n]) over about 500 years (841 µrad/y). Such estimates depend somewhat on the choice of spectral representation, but tracing the twists and turns in the evolutionary pathways informs us about the shape of the landscape of change (Fig. 1). As the conference call says, "phonological representations are dynamic, shaped by forces on diverse timescales." In this paper we examine those dynamics over very long timescales, of hundreds and thousands of years.

*Figure 1: Changing directions in the development of English "one" from Proto-Indo-European *oinos. Clockwise/downward angles represent opening changes; anticlockwise/upward angles represent closings. Length of line segments is proportional to time.*



References

Bouchard-Côté, A., D. Hall, T. L. Griffiths and D. Klein (2013) Automated reconstruction of ancient languages using probabilistic models of sound change. *PNAS.* 110: 4224–4229.

Bouckaert. R., P. Lemey and 7 others (2012) Mapping the origins and expansion of the Indo-European language family. *Science* 337 no. 6097.

Chang, W., C. Cathcart, D. Hall and A. Garrett (2015) Ancestry-constrained phylogenetic analysis supports the Indo-European steppe hypothesis. *Language* 91 (1), 194–244.

Hadjipantelis, P.-Z. (2013) *Functional Data Analysis in Phonetics.* PhD thesis, Univ. of Warwick.

Hadjipantelis, P.-Z., N. S. Jones, J. Moriarty, D. A. Springate and C. G. Knight (2013) Function-valued traits in evolution. *J. Royal Society Interface*, 10 (82).

Jones, N. S. and J. Moriarty (2013) Evolutionary inference for function-valued traits: Gaussian process regression on phylogenies. *J. Royal Society Interface*, 10 (78).

Nakhleh, L, D. Ringe and T. Warnow (2005) Perfect phylogenetic networks: a new methodology for reconstructing the evolutionary history of natural languages. *Language* 81 (2), 382–420.

Ringe, D., T. Warnow and A. Taylor (2002) Indo-european and computational cladistics. *Transactions of the Philological Society* 100 (1), 59–129.

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.