

Charting the rise and demise of a phonotactic change

Rhona Alcorn, Warren Maguire

Diachronic corpora are rarely annotated at the sound level, putting historical phonologists at a distinct disadvantage when compared to historical syntacticians and morphologists. In this paper we introduce a new technique for transforming historical corpora into data which support the quantitative study of all types of phonological change.

The technique we describe has evolved from an ongoing project which aims to reconstruct the phonology of some of the earliest attested varieties of Scots. The project, *From Inglis to Scots: Mapping Sounds to Spellings* ('FITS'), draws its data from the corpus of legal texts which underpins *A Linguistic Atlas of Older Scots, Phase 1, 1380-1500* ('LAOS', Williamson 2008). Each text in this corpus has been transcribed diplomatically from an original manuscript (or facsimile), and so preserves all original features of spelling, abbreviation, deletion, insertion, punctuation, etc. In addition, each token of each transcribed word—or, in many cases, morpheme—is lexico-grammatically tagged, allowing us to identify all tokens of the same word or morpheme regardless of its spelling.

Our technique, termed 'grapho-phonological parsing' (Kopaczyk *et al.*, under review), involves resolving each lexico-grammatically tagged token in LAOS into a series of spelling units ('graphemes'), which we individually annotate with the following information:

1. A corresponding sound value, e.g. [θ] for <y> in *yefis* 'thieves', [ð] for <y> in *fayer* 'father', [ʃ] for <sch> in *flesch*, [ø:] for <ui> in *guid* 'good'.
2. An etymological sound value, i.e. that of the likely input variety (or varieties), e.g. Old English [d] for <y> in *fayer* 'father', Old English [o:] for <ui> in *guid* 'good'.
3. The graphotactic, morphographic and morphological context.
4. The date, place of origin and sub-genre of the source text.

With regards to points 1 and 2, we follow Laing & Lass (2003: 268) and use square brackets to represent "poorly resolved broad phonetic realizations."

The data thus annotated are stored in a fully-searchable database. The database may then be accessed to determine, for example, which sounds are attested in sequence and, for each attested sequence, we may additionally retrieve its frequency, lexical distribution and position within the word (see Molineaux *et al.* 2016 for a demonstration of this database in action).

In this paper, we use our database of annotated grapho-phonologically parsed spellings to chart the rise and demise of a change in the stem-level phonotactics of Older Scots (OSc). Our target is the apparent devoicing of OSc [v] in word-final and pre-inflectional position, as suggested by the use of the <f(f)> grapheme in words such as *luff(f)* 'love' and *giffyn* 'giving' (Johnston 1997: 104). Since words of these types have [v] in both Old English and Modern Scots, the question arises as to whether there was, in fact, a process of devoicing of [v] (to [f]) in the development of OSc as the spellings seem to indicate, or whether we are simply dealing with a change in spelling which does not reflect any change in pronunciation (so that OSc also had [v] in these words).

We find evidence of complex patterning of <f(f)>-type and <v>-type spellings for etymological [v] (e.g. *love*) and [f] (e.g. *life*) in our database. Specifically, it becomes clear that the choice between <f(f)> and <v> correlates strongly with the etymological consonant, the presence or absence of inflections, the source language of the words concerned, and the date of the texts under analysis. On this basis we argue that a final devoicing of [v] to [f] is more likely than a spelling change alone and is consistent with what we know about phonological change. Having concluded that <f(f)> spellings do indeed indicate devoicing of [v], we show:

(i) that this change was variable (perhaps due to its interaction with final schwa apocope); (ii) that it spread by structural analogy from word-final to pre-inflectional position; and (iii) that the devoicing was later reversed (during the 15th century in pre-inflectional position, and after the 15th century in word-final position).

References

- Johnston, Paul. 1997. Older Scots phonology and its regional variation. In Charles Jones (ed.) *The Edinburgh History of the Scots Language*, 47-111. Edinburgh: Edinburgh University Press.
- Kopaczyk, Joanna, Benjamin Molineaux, Vasilis Karaiskos, Rhona Alcorn, Bettelou Los & Warren Maguire. Under review. Towards a grapho-phonologically parsed corpus of medieval Scots. Database design and technical solutions.
- Laing, Margaret & Roger Lass. 2003. Tales of the 1001 nists: the phonological implications of litteral substitution sets in some thirteenth-century South-West Midland texts. *English Language and Linguistics* 7.2, 257-78.
- Molineaux, Benjamin, Joanna Kopaczyk, Warren Maguire, Rhona Alcorn, Vasilis Karaiskos & Bettelou Los. 2016. Tracing L-vocalisation in early Scots. *Papers in Historical Phonology* 1: 187-217.
- Williamson, Keith. 2008. LAOS, A Linguistic Atlas of Older Scots, Phase 1: 1380-1500, [<http://www.lel.ed.ac.uk/ihd/laos1/laos1.html>] Edinburgh: University of Edinburgh.