



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Phonotactics, graphotactics and contrast

Citation for published version:

Molineaux, B, Kopaczyk, J, Alcorn, R, Maguire, W, Karaiskos, V & Los, B 2020, 'Phonotactics, graphotactics and contrast: The history of Scots dental fricative spellings', *English Language and Linguistics*.
<https://doi.org/10.1017/S1360674319000479>

Digital Object Identifier (DOI):

[10.1017/S1360674319000479](https://doi.org/10.1017/S1360674319000479)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

English Language and Linguistics

Publisher Rights Statement:

This article has been published in a revised form in English Language and Linguistics: <https://doi.org/10.1017/S1360674319000479>. This version is free to view and download for private research and study only. Not for re-distribution, re-sale or use in derivative works. © The Author(s), 2020. Published by Cambridge University Press.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Phonotactics, graphotactics and contrast:

The history of Scots dental fricative spellings

BENJAMIN MOLINEAUX*, JOANNA KOPACZYK**, RHONA ALCORN*,

WARREN MAGUIRE*, VASILIS KARAIKOS* and BETTELOU LOS*

**The University of Edinburgh*

***The University of Glasgow*

ABSTRACT

The spelling conventions for dental fricatives in Anglic languages (Scots and English) have a rich and complex history. However, the various – often competing – graphemic representations (<þ>, <ð>, <y> and <th>, among others) eventually settled on one digraph, <th>, for all contemporary varieties, irrespective of the phonemic distinction between /ð/ and /θ/. This single representation is odd among the languages’ fricatives, which tend to use contrasting graphemes (cf. <f> vs. <v> and <s> vs. <z>) to represent contrastive voicing, a sound pattern that emerged nearly a millennium ago. Close examinations of the scribal practices for English in the late medieval period, however, have shown that northern texts had begun to develop precisely this type of distinction for dental fricatives as well. Here /ð/ was predominantly represented by <y> and /θ/ by <th> (Jordan, 1934; Benskin, 1982). In the late fifteenth and early sixteenth centuries, this “Northern System” collapsed, due to the northward spread of a London-based convention using exclusively <th> (Stenroos, 2004). This paper uses a rich body of corpus evidence for fifteenth-century Scots to show that, north of The North, the phonemic distinction was more clearly mirrored by spelling conventions than in any contemporary variety of English. Indeed, our data for Older Scots local documents (1375-1500) shows a pattern where <y> progressively spreads into voiced contexts, while <th> recedes into voiceless ones. This system is traced back to the Old English positional preferences for <þ> and <ð> via subsequent changes in phonology, graphemic repertoire and letter shapes. An independent medieval Scots spelling norm is seen to emerge as part of a developing,

proto-standard orthographic system, only to be cut short in the sixteenth century by top-down anglicisation processes.

KEYWORDS: Scots, phonotactics, graphotactics, phonology, spelling

1 INTRODUCTION

The late medieval and early modern period saw the rise of a relatively stable, standard English orthographic system, such that it is still recognisable to most readers today. At a similar time, and in parallel, Scots had also entered the early stages of standardisation (cf. Meurman-Solin 1997), forging its own distinct path at a time when the two nations were still following separate political courses. For Scotland, however, this process was cut short by growing influence from southern England – anglicisation – which, as Murison puts it, meant that Scots “lost spiritual status at the Reformation, social status at the Union of the Crowns, and political status with the Parliamentary Union” (1979: 9). This resulted in its independently developing orthographic system being derailed during the sixteenth century, as writing in the language became less prestigious and ceased to be as widely used. In time, the spelling conventions of Scots became anglicised, via the adoption of southern orthographic practices and the signalling of phonic differences through apostrophes and other markers. Finally, while the nineteenth century saw new spelling proposals, no single set of conventions has emerged for the language, so the matter of present and future Scots orthography remains far from settled, as indeed does the status of the language itself (cf. Bann & Corbett 2015).

This paper deals with the variable spellings for the dental fricatives in early Scots, in the heyday of Scots as a multi-purpose, independent language (cf. Murison, 1979: 8-9). We will argue that the “(th) variable” (cf. Stenroos 2004) in Older Scots (OSc) shares some aspects of the late Northern Middle English (NME) system first described by Jordan (1934) and elaborated by Benskin (1977, 1982), Stenroos (2004, 2007), Jensen (2012) and Adamczyk (2016). However, OSc appears to go a step further towards its own internally coherent system which more transparently represents dental fricative voicing contrasts via spelling contrasts, a development unique in the Anglic world, past or present.

The lack of a voicing-based contrast in present-day orthographies for the dental fricatives in both Standard English and Scots writing is surprising when compared to the orthography of other obstruents in these languages. Among the dental fricatives, <th> is used for both /θ/ and /ð/, while labio-dental and alveolar fricatives (mostly) distinguish voicing by using <f> for /f/ and <v> for /v/ or <s> for /s/ and <z> for /z/.¹ The reasons behind this particular gap have historical roots, as we shall see, but it is also probably related to the somewhat predictable distribution of voicing in the dental set and, hence, the low functional load of the contrast, at least for lexical discrimination.

The distribution of voiced and voiceless dental fricatives in present-day varieties of Scots and English has unique characteristics. For historical reasons, dental fricatives are found mostly in native Germanic forms (either Old English or Scandinavian). Their voicing today is highly dependent on phonotactic position, as well as on the functional nature and morphological category of the lexical item they appear in (see Table 1).

Table 1: *Distribution of voiced and voiceless dental fricatives in present-day English*

¹ Among the alveolars, however, root-final and inflectional [z] is overwhelmingly spelled <s>. The case of the post-alveolar fricative is different, since [ʒ] is extremely rare in both Scots and English.

		INITIAL		MEDIAL		FINAL
[ð]	FUNCTION	[ð]ese, [ð]em	ALL	o[ð]er, ga[ð]er	VERB	brea[ð]e, ba[ð]e
[θ]	CONTENT	[θ]eme, [θ]under	(Greek >	e[θ]er, Ca[θ]olic)	NOUN	brea[θ], ba[θ]

Word-initial dental fricatives tend to be voiced in functional items and voiceless in content words. Medial dental fricatives are overwhelmingly voiced in all word categories, with the major exceptions being Greek borrowings, compounds (cf. *bath tub*, *toothache*) and derived words (cf. *breathalizer*, *wrathful*).² Word-finally, voicing mostly depends on word category: voiced in verbs and voiceless in nouns. For all categories there are exceptions and there is dialectal and contextual variation.³ Nevertheless, the distribution of voicing is highly predictable.

Voicing-based lexical contrasts in dental fricatives are very rare, with minimal pairs like *thy~thigh* and *wreathe~wreath* being primarily distinguished by their grammatical category, for which voicing acts almost as a morphophonological exponent. As a result, at least by the metric of complementary distribution, voicing contrast is notably weak amongst the dentals (especially when viewed in isolation from other fricatives). Under such conditions of weak phonemic contrast, the present-day Scots and English “orthographic lacuna” (Lass 1991–3: 21) for voicing in dental fricatives seems unsurprising or even unavoidable. Nonetheless, we will show here that a different path of development – where “marginal” phonemic contrast (Lass 1991–3: 11) leads to spelling contrast – is possible and was, indeed, near categorical in the scribal practices for dental fricatives in late-fifteenth-century Scots.

² Of course, in compounds and derived words, the fricative is still morpheme-final.

³ Variation in pre-inflectional position is particularly widespread, cf. MacKenzie (2018).

In what follows, we examine the evidence for spellings of dental fricatives in pre-Modern Anglic varieties. We begin by examining the system for Old English (OE), and its eventual collapse (§2). We then survey scholarship on the development of the (th) variable in NME, where a phonemically-based spelling contrast is claimed to have emerged and then disappeared (Benskin 1982) (§3). In §4 we turn to the OSc situation, surveying the admittedly sparse literature, in order to move on to new corpus-based data for the spelling of our target segments (§5). In our discussion of the data, in §6, we propose a new pathway for the development of dental fricative spellings in OSc. Finally, §7 offers some conclusions on the matters of spelling change, “weak” phonemic contrast, standardisation and methodology.

2 OLD ENGLISH DENTAL FRICATIVES

2.1 *The Old English dental fricative spelling system*

The distribution of voicing in OE fricatives is predictable in relation to their phonological contexts, such that in medial position, between voiced sounds, the fricative is voiced, while elsewhere – and in the case of geminates – it is voiceless.⁴ Such a pattern is summarised in Table 2. Importantly for our later discussions (especially §5.2.2), the medial voicing pattern creates paradigmatic alternations through inflection, e.g. *að* ‘oaths (nom sg)’ (with [θ]) ~ *aðas* ‘oaths (nom pl)’ (with [ð]).

Table 2: Voicing of Old English fricatives by phonotactic position

INITIAL	MEDIAL	FINAL
---------	--------	-------

⁴ This distribution may be further refined both in terms of phonological and morphological structure. For an overview, see Minkova (2011).

[θ]~[ð]	[θ]: <i>þing</i> ‘thing’	[ð]: <i>ōðer</i> ‘other’	[θ]: <i>bæð</i> ‘bath’
	[θ]: <i>ðū</i> ‘thou’	[θθ]: <i>niððas</i> ‘men’	
[f]~[v]	[f]: <i>fisch</i> ‘fish’	[v]: <i>drīfan</i> ‘drive’	[f]: <i>hrōf</i> ‘roof’
		[ff]: <i>offrian</i> ‘offer’	
[s]~[z]	[s]: <i>sweord</i> ‘sword’	[z]: <i>nosu</i> ‘nose’	[s]: <i>hus</i> ‘house’
		[ss]: <i>blissen</i> ‘bless’	

The orthographic situation was such that, while labials and alveolars were generally represented by one grapheme each (<f> and <s>, respectively), two graphemes (<þ> and <ð>) were used interchangeably for dentals. In some of the earliest texts from the north of England, scribes additionally used <th>, a convention that was already in use for Greek <θ>-words in Latin. <th> continued to be applied throughout the Middle Ages to Anglo-Saxon names in Latin (cf. Benskin 1982: 18–19). Another fairly widespread, early convention for dental fricatives was the use of <d>, which eventually was made distinctive from /d/ by the addition of a strikethrough, thus bringing about the <ð> grapheme (cf. Campbell, 1959: §58).⁵ As for <þ>, this was brought in from the Germanic runic tradition, where it also represented dental fricatives. The exact dynamics that led to the relatively stable and exclusive use of <þ> and <ð> throughout most of the OE period may be impossible to reconstruct. Nevertheless, it appears that their use was semi-systematic, not in their phonic distribution, but in their positional preferences (cf. Minkova 2014: 23–4; Lass 1991–3: 6, Campbell 1959: §58; Stenroos 2004: 272–3), i.e. in terms of their *graphotactics*.

⁵ The earliest OE uses of <th>, <d> and <ð> have all also been claimed as spelling conventions borrowed from the Old Irish tradition (cf. Campbell, 1959: §55; Strang 1970: 363).

Given the lack of quantitative work on such distributions, we conducted a rough search of the *Dictionary of Old English Corpus* (DOEC, 2009) for instances of <ð> and <þ> at the start, middle and end of words. Results, in Figure 1, represent a view of the entire corpus and show <ð> and <þ> not to be random across positions, but that <þ> is used far more frequently at the start, while <ð> is used far more frequently in medial and final positions. A comparison of the allophonic distribution of voicing (Table 2) and graphemic choice (Figure 1) shows a misalignment between phones and graphemes. The pattern in Figure 1 points, rather, to a strong graphotactic preference among scribes for <þ> in initial positions, as opposed to <ð> in all other positions.⁶

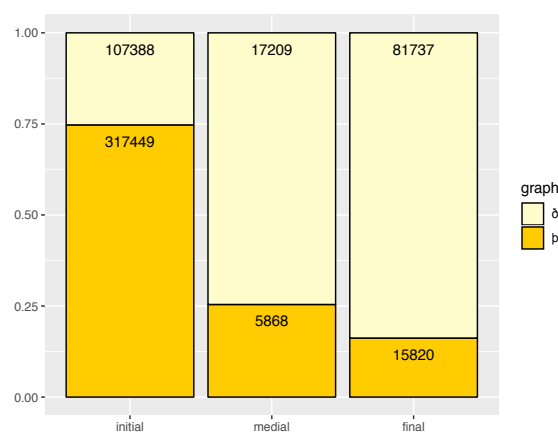


Figure 1: Proportion of <ð> and <þ> spellings for dental fricatives by position in the word in the *Dictionary of Old English Corpus* (DOEC 2009)

2.2 The collapse of the OE spelling system

⁶ Of course, this could be studied in further detail as a function of particular lexical or grammatical items and according to different sources, either scribal, temporal and regional, but this falls beyond the scope of this paper. As an anonymous reviewer points out, the DOEC relies on edited texts, some of which may fail to faithfully represent scribal usage for varying orthographic practices. While we are unable to perform a full check of the 3,060 corpus texts, we assume that editorial practice follows the main forms in each text, tending towards balance throughout the entire dataset.

The profound administrative, ecclesiastical and (high) cultural upheavals following 1066 brought about fundamental changes in the nature, dispersion and transmission of written material. Compared to the late OE *Schriftsprache* of the West-Saxon scriptoria, the decentralised, local early Middle English (eME) material evidences dramatic variation at all levels of the grammar, alongside an abundance of innovative spelling systems. As Benskin (1982: 20) puts it, there was “no longer any pressure on writers of English to conform to a single written standard”. The tradition of continuous vernacular writing saw gaps and loss of prestige, both of which contributed to substantial changes in the conventions for spelling. This is particularly true of the early NME scribal traditions, which were probably the main source for OSc conventions as well (cf. Kniezsa 1997). In the north, indeed, we see that the link to classical West-Saxon spellings was far less straightforward. Here, most textual production was done in Latin, which, as we shall see, exerted an important influence upon vernacular spelling. For the case of dental fricatives, the changes in the orthographic repertoire are accentuated by important changes in the phonological makeup of post-Conquest varieties.

Amongst fricatives, the eME period saw formerly allophonic voicing patterns become phonemicized. Indeed, both alliterative and rhyme evidence shows voicing is no longer predictable by phonotactic context alone (cf. Minkova 2011), in a pattern that strongly resembles that of present-day Anglic varieties (Table 1). Crucially, this period also sees the rise of distinct spellings for voiced and voiceless fricatives (the emergence of <v/u> for [v] and <z/ʒ> for [z]), providing further evidence for the phonemic distinction.

The emergence of phonemic contrast amongst fricatives has traditionally been attributed to both external and internal factors (cf. Lass 1992: 58–9, Minkova 2011).

Externally, the influx of French loanwords with an initial voiced fricative broke down the OE phonotactic restriction – particularly in the labial series – generating a number of minimal pairs (*vēle* ‘veal’ vs. *fēle* ‘many’). Internally, phonemicisation was the result of degemination, schwa loss and prosodic-based voicing. Degemination meant that medial geminate fricatives – which were always voiceless in OE – became singleton voiceless medial fricatives in ME (cf. OE *bli[ss]en* > ME *bli[s]e* ‘to bless’), contrasting with the etymological singletons, which were voiced in OE (cf. OE *wi[z]e* > ME *wi[z]e* ‘wise’). Loss of final schwa ‘exposed’ a preceding voiced fricative to the word edge in most verbal forms, contra the OE phonotactic restriction on final voiced fricatives and creating the contrast with their nominal counterparts as in *mou[ð]_V~mou[θ]_N*.⁷

The key language-internal factor for the phonemic split among dental fricatives, however, is the lenition (in this case, voicing) of fricatives in prosodically weak positions. As a result, the initial [θ] of pronouns, demonstratives and determiner such as *they*, *thy*, *there* or *the* became voiced.⁸ The dating of this change is difficult as there is no clear spelling evidence. However, the poetic record for OE suggests that, at least as far back as *Beowulf* and throughout the period, function and content words alliterated with each other (see Minkova 2011). By the time of Chaucer, however, we find rhymes such as *soothe/to the* ‘the truth/to thee’ (*The Canon Yeoman’s Prologue* 662-3) and *swithe/hy the* ‘swiftly/hasten the’ (*The Canon Yeoman’s Tale* 1294-5) (cf. Minkova 2014: 95 fn. 24, Jordan 1934: §207), indicating initial voicing of *the*, *thee*, etc. Despite the evidentiary gap between the OE and late ME data, it is reasonable to assume that dental fricative voicing

⁷ The status of the word-final voiced fricative, however, is problematic for Anglic varieties that also underwent a more general process of final fricative devoicing, such as OSc, as we shall see in §5.2.1 (see also Maguire *et al.* 2019).

⁸ We also find this at the end of high-frequency words and in other fricatives, such as in *of*, *was*, *is*, *has* and, variably, *with*.

contrast did not emerge in isolation, but did so in tandem with other fricatives, for which spelling evidence is more revealing. It is therefore possible to conclude that by the end of the eME period this change was well underway, if not complete.

In terms of spellings, the longstanding OE <ð, þ> system for dental fricatives quickly collapsed in eME. Most notably, the use of <ð> saw a sharp decline. While we still find it in many texts in the *Linguistic Atlas of Early Middle English* (LAEME 2013) – though rarely in the North, see Figure 2 – this variant is not recorded in any of the linguistic profiles for the *Electronic Version of the Linguistic Atlas of Late Mediaeval English* (eLALME 2013) except extremely rarely in its capital form: <Ð>. In parallel, we find that <th> makes a bold reappearance, surfacing frequently in the earliest ME material, such that it can be found in most LAEME texts (Figure 2). The spread of <th> is usually attributed to Norman influence, as is the preference for <sh> and <ch> (cf. Kniezsa 1997: 38), however it is more likely that the digraph followed Anglo-Latin usage, which gained ground during the post-Conquest drop in vernacular writing. Regardless, this does not explain the preservation of <þ>, in most texts, at the expense of <ð>. Here, the strong association of <þ> with initial position – including highly frequent function words – probably tipped the balance in its favour, as we will argue in §6.

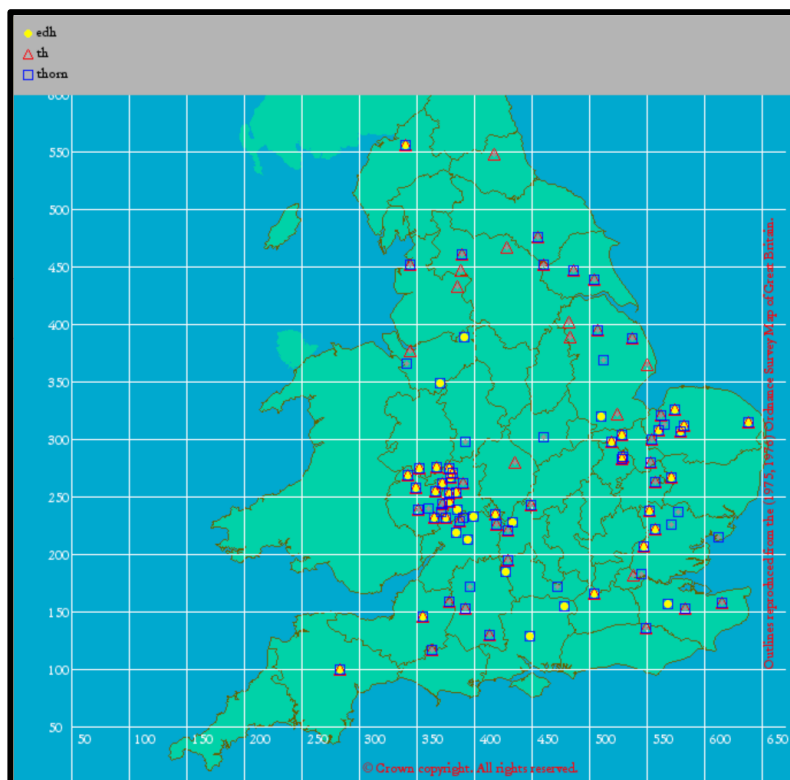


Figure 2: The distribution of $\langle\delta\rangle$, $\langle\beta\rangle$ and $\langle th\rangle$ spellings in LAEME⁹ (colour version available online)

3 THE “NORTHERN SYSTEM” FOR ME DENTAL FRICATIVE SPELLINGS

While the early ME period witnessed a proliferation of scribal practices, late ME ushered in a reduction in variation and the establishment of broader regional patterns and conventions. A striking orthographic change as regards the dental fricative began as natural variation in the shape of letters (i.e. *figurae*, cf. Benskin 1997, Laing & Lass 2003), which ultimately led to a merger of $\langle\beta\rangle$ and $\langle y\rangle$ shapes (Benskin 1982, Laing & Lass 2009, 2013), through the ‘lobing’ of $\langle y\rangle$ and the loss of the $\langle\beta\rangle$ ascender (particularly in *textura* script). In such scribal systems, a more $\langle y\rangle$ -like shape often became dominant. There are also cases where the overlap in shape may be a cline between

⁹ Early instances of $\langle y\rangle$ -shaped- $\langle\beta\rangle$ (see §3, below) also occur in LAEME, and are coded under $\langle y\rangle$ (see Laing & Lass 2013: *Introduction* §3.3.3), however, such cases have not been included in the map.

a more canonical <þ> and a more canonical <y>, but where all these shapes may be found in either the [ð]~[θ] etymological category or in [j]-words. This <y>-shaped-<þ> system is found predominantly in the North of England and Scotland, and is conspicuously absent in the south, as noted by McIntosh (1974: 608–9) and Benskin (1982: 14–16). The contrast between the two systems – the ones that merge <þ> and <y> (here given as <y>) and those which keep them separate – is evidenced by the map in Figure 3, showing a fairly clear late ME *isograph* (Benskin 1982: 16).

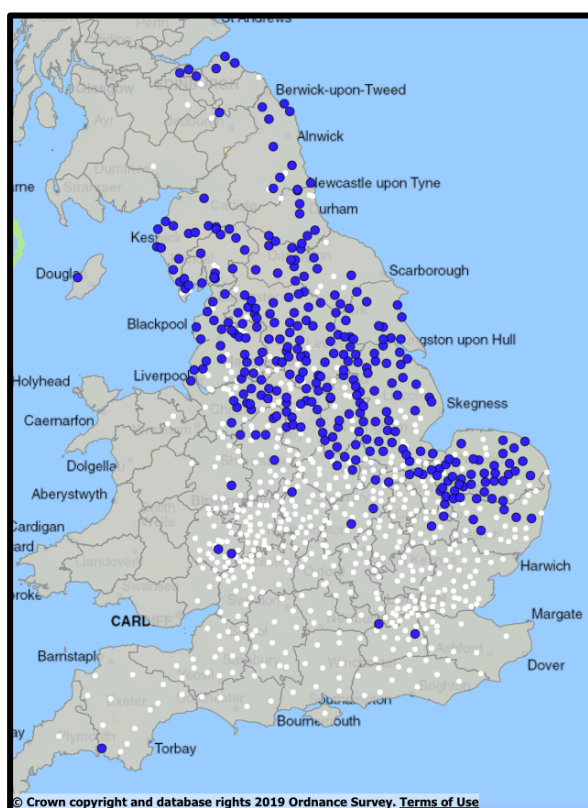


Figure 3: eLALME map representing <y>-initial spellings of the items *the, these, those, than, they, then, their, though, them, think, there, through, thee, thy, thou, thence, thither, three, third, thousand* (blue/dark dots). White dots represent all other survey points that do not contain such variants (colour version available online)

As a result of the shape merger, both <þ> and <y> are transcribed as <y> in LALME (cf. Vol. 2: xvii–§6.5) and LAEME (Laing & Lass 2013 :§3.3.3) for writing systems which

do not distinguish them. This is a departure from non-philological transcription practices criticised by Benskin (1977: 506–7, fn. 9), who cites Simpson as saying that, in early Scots, <y> was a “convenient way of writing *th*” (1973: 42)¹⁰. Benskin insists “it is *y/b* that is replaced by *th*” (1977: 506–7, fn. 9). A closer look at the early ME data, however, shows both accounts to be problematic. On the one hand, the idea that <th> was somehow the default spelling for dental fricatives is clearly historically “misguided” (Benskin 1977: 506–7, fn. 9). On the other hand, Benskin’s view assumes a period where dental fricatives were predominantly spelled as <þ> (and <y>) in all positions and were only subsequently replaced by <th>, towards late NME. The *LAEME* data in Figure 4 – compiled in the same way as that for *DOEC* in Figure 1 – suggests a more complex picture where <þ> coexisted with <ð> and <th> in many eME manuscripts. Indeed, it is more than likely that many of the scribal systems – especially in the north (see Figure 2) – the <þ> spelling convention never became the main variant in non-initial position. Instead, it seems that Anglo-Latin <th> filled the void left by loss of non-initial-<ð> starting relatively early, rather than simply ‘replacing’ a well-established <y/þ>-system in late NME.

¹⁰ By the 2009 edition of Simpson’s *Scottish Handwriting 1150-1650*, the relevant passage had been removed.

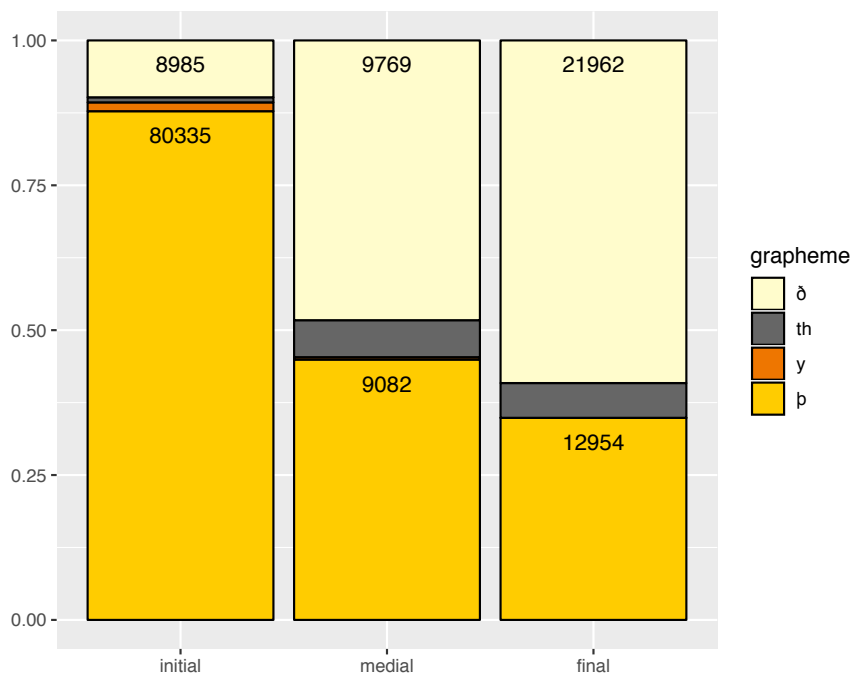


Figure 4: Proportion of δ, th, y and $þ$ and spellings for dental fricatives by position in *LAEME*.

Crucially, as shown in Figure 4, eME saw an overall increase in the proportion of word-initial $þ$ as compared to the OE system (cf. Figure 1), from 74.7% to 87.8%. At the same time, $þ$ and (to a lesser extent) th take the place of δ in medial and final positions.¹¹

A key NME aspect of the increase of th to represent dental fricatives, according to Benskin (1977, 1982), is that it first spread to word-final position and then to word-initial content words, thus bringing about a short-lived, voicing-based spelling distinction:

In ME, *th* is first generalised in place of word-final *-þ* and this *-þ*, as it happens, is in OE and ME nearly always voiceless ('soft'). In most northerly dialects, *th* is

¹¹ An analysis of these distributions across genre, period and, particularly, geographic origin might shed further light on these distributions. However, such analyses exceed the scope of this paper.

then extended from final position to all other voiceless *þ* contexts; and except for medial cases arising in inflected forms, these are confined to word-initial position. There thus arises a system whereby (1) words like *thing*, *through*, *thousand* are spelled *th*-, but (2) words like *they*, *then*, *there* are spelled *þ*- or *y*-. The use of *þ* (or *y* for *þ*) and *th* is hence phonetically conditioned in the orthographies of a great many scribes, an observation which seems to have eluded most scholars. (Benskin 1977: 506–7, fn. 9)

The observation of what, indeed, would be a phonemic – and not just phonetic – spelling contrast in the Northern System, was also made by Jordan (1925: §203) and raises the question as to why the dental-fricative spelling system did not settle on a voicing-based contrast, as in the case of other fricatives. Nevertheless, it is Benskin’s claim that the relentless takeover of <th> did not settle for representing only voiceless segments, but went on to spread – in both NME and Scots – first to medial position and then to the initial position of function words (1982: 18). The result is a four-stage process, illustrated in Figure 5.

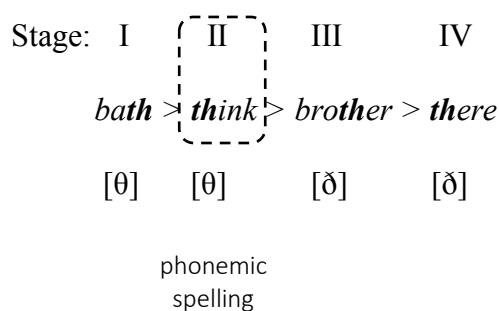


Figure 5: Spread of <th> in the ME “Northern System” (Benskin 1982)

Two decades after Benskin’s claims, and with corpus methods established as a central tool in historical linguistics, Stenroos (2004) took a closer look at the ME Northern

System. Examining the hypothesised stages, she found that, indeed, for most NME texts in her sample – which also have the <y>-shaped-<þ> – there was a tendency for voiceless fricatives to be written as <th>. This pattern consolidated over the course of the three half-centuries covered by her materials, as seen in Figure 6. However, while voiced fricatives were at first consistently written with <y> (or <þ>), by the second half of the 15th century, the frequency of <th> in these contexts reached about 30% (Figure 7). Unfortunately, the reported data are not fine-grained enough to distinguish whether most of the changes are occurring word-medially (as predicted by Benskin) or word-initially in function words. It is apparent, however, that the NME system as a whole was never purely phonemic.

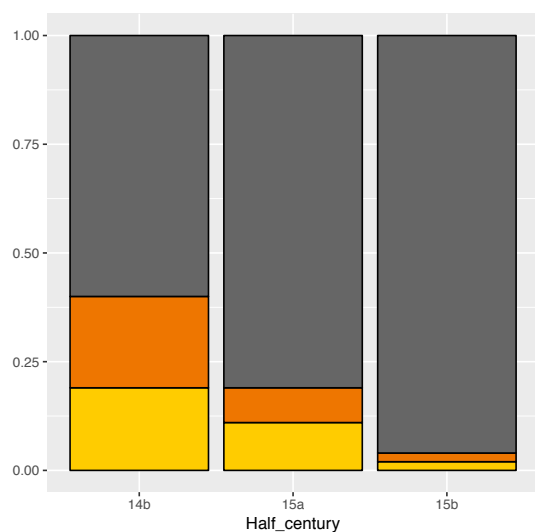


Figure 6: Based on Stenroos (2004: 272): NME distributions of the (th) variable in voiceless contexts by half-century

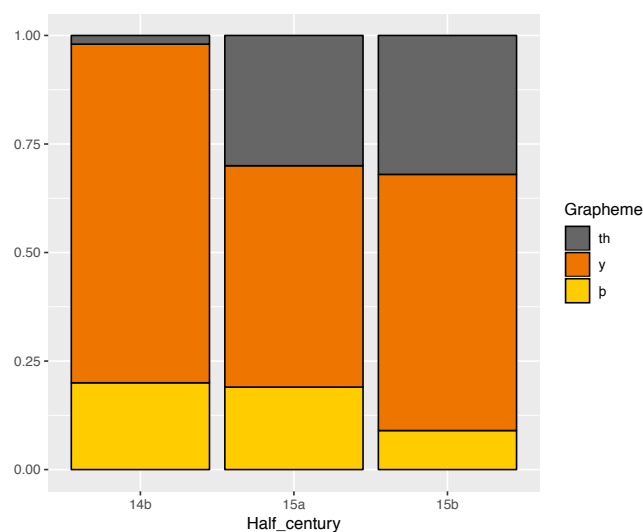


Figure 7: Based on Stenroos (2004: 272): NME distributions of the (th) variable in voiced contexts by half-century

Stenroos (2004: 275–9) went further, however, and looked at the effects of text-type in the spread of <th>. As seen in Figure 8, she found that for voiced contexts (as well as voiceless ones), <th> progressed much faster in documents than in literary materials. This, she tells us, was mainly due to the spread of <th> to medial fricatives (83% are <th>) in words such as *other* and, to a lesser degree, to the fricative-initial functional items such as *they* (47% are <th>). In other words, Stenroos' data suggests that the transition into Benskin's final two stages of the <th> takeover was led by those producing documentary materials.

Again, Stenroos' data is not granular enough to establish whether the stages proposed by Benskin were step-wise in either literary or documentary material, and hence to decide whether the rough correspondence of phonological and spelling facts were ever systematic. Importantly, however, Stenroos does engage with the question of why – if there was a

phonemic spelling stage in NME – it was so short lived.¹² The evidence points to the spread of southern spelling practices, probably related to the courts and London. In such systems <þ> rarely merged with <y>, and <th> was incorporated relatively late, in the

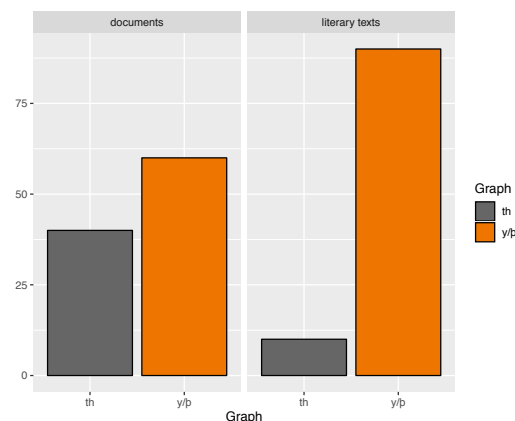


Figure 8: NME distributions of the (th) variable in voiced contexts, by genre (Stenroos 2004: 279)

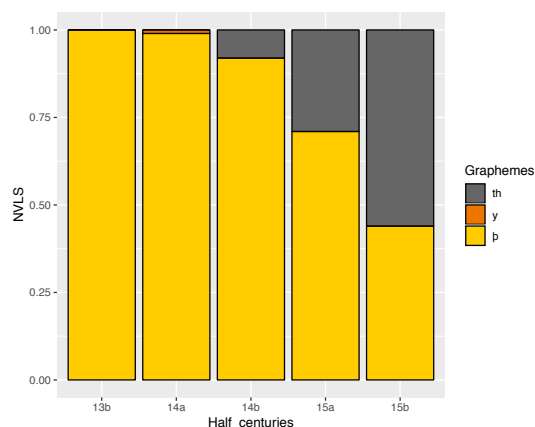


Figure 9: Based on Stenroos (2004: 274) Southern ME distributions of the (th) variable by half century

¹² The claim that printing is responsible for the rise of <th>, since typesets lacked <þ> (cf. Scrag 1974: 2) was rejected by Stenroos (2007: 9–11). Given that <th> was making inroads into Anglic spellings well before the invention of printing, it cannot be its cause.

mid-fourteenth century. Since there were no phonographic or graphotactic restrictions on the distribution of <þ> and <th> in those southern varieties, <th> quickly replaced <þ> in all contexts alike (Figure 9). This growing convention would have quickly spread to other scribes working in the legal profession further north, affecting documents first and leading to the final breakdown of the Northern System. Faced with this pressure from southern scribal practices, medial voiced fricatives probably caved in to the incoming convention first, while <y>-initial function words held out a while longer, due to their frequency, thus following the final two stages in Benskin's proposal.

4 PREVIOUS ACCOUNTS OF DENTAL FRICATIVE SPELLINGS IN OLDER SCOTS

The history of Scots spelling conventions is impossible to reconstruct earlier than the later middle ages, given the three-century gap between the tenth century Old Northumbrian materials¹³ and the first OSc glosses, documents and literary works, which surface in the late fourteenth century. However, the spelling system that emerges in Scotland in the fourteenth century is not a *de novo* adaptation of the Latin alphabet to Scots but “a bleeding of common Middle English spellings, special northern English scribal traditions and, in some cases, native innovations or graphemes not evidenced elsewhere” (Kniezsa 1997: 46). Ultimately, early Scots orthography is part of a continuum of spelling variation reaching northwards beyond the NME practices, that is, a system with its roots in Old English dialect writing, interspersed with Anglo-Latin and Anglo-Norman conventions.

Not only is the Scots written record late to emerge, but it also soon lost its unique identity. Indeed, one of the areas where the influence of Southern English first made

¹³ Northumbrian texts all originate, geographically, south of the present-day Scottish-English border, with the exception of the runic Ruthwell Cross, which may be dated as far back as the eighth century.

inroads into OSc was spelling, via the large influx of English books and the introduction, at the start of the sixteenth century, of the printing press (see Bald 1926; Kniezsa 1997: 44–6). While some Scots orthographic conventions were preserved in early printed texts, this was short-lived, especially in the context of religious publications and the Reformation, dominated by English and its spelling conventions. The Union of Crowns and eventually the Parliamentary Union tipped the scales in favour of anglicisation in both written and spoken Scots. The key period for an independent Scots spelling, then, is before anglicisation set in, i.e. before the sixteenth century.

Referring to what must be precisely this period, Benskin (1977: 506–7, fn. 9) gives a general overview of the (th) variant:

the system of Early Scots is as follows: (i) *y* and *þ* are almost invariably confused ... (ii) *y* for *þ* is restricted to word-initial voiced context, the *they-them-there* group; and elsewhere *th* is written, except for (iii) medial contexts which are immediately followed by a suspension, thus *oyer* ('other, or') regularly so written, and occasionally *broyer* ('brother').

This pattern is taken to represent the Scottish continuation of the Northern System. However, while further studies of this pattern have given a corpus-based view of this feature of NME (Stenroos 2004, 2007, Jensen 2012, Adamczyk 2016), the OSc system has not been subject to the same close scrutiny. This is particularly important in order to ascertain whether the same step-wise spread of <th> claimed for NME can be reconstructed for Scots, and hence whether the period of phonemic spellings – as well as its subsequent dissolution – can be confirmed.

More recent accounts of the dental fricative spellings in Scots tend to oversimplify matters. Kniezsa, for instance, claims that “Old English <ð> and <þ> are replaced with

the grapheme <th>, introduced under French influence ... It appears in all positions in words” while at the same time, a “defectively formed <þ> is written mainly in grammatical words, well into the fifteenth century” (1997: 38). Similarly, Bann and Corbett claim that in English “<th> did not reappear regularly in texts until Caxton reintroduced it in printing in the late fifteenth century” and that, while in OSc <y> forms “were common ... particularly word-initially”, “<th> was eventually used in all positions” (2015: 34). Not only are these accounts problematic in their treatment of the origins of <th>, there is also little consideration of the reasons behind the persistent distribution of <th> and <y> while it lasted, and the exact nature of the processes by which one form rose and the other fell in its usage. The following sections attempt to fill these gaps.

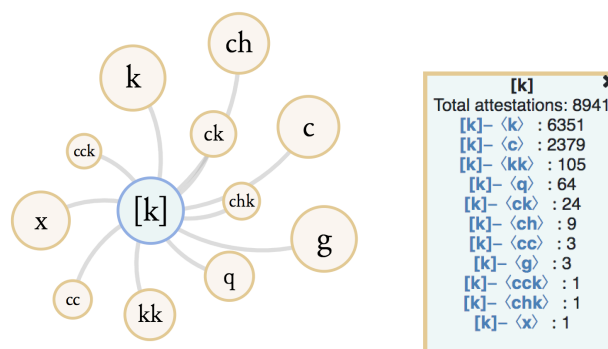
5 NEW DATA FOR SCOTS DENTAL FRICATIVE SPELLINGS

5.1 FITS and grapho-phonological parsing

In order to trace the development of dental fricative spellings in pre-anglicised Scots, we need a fine-grained dataset for the temporal, regional and contextual variation. To this end, our analysis relies on the c.1,250 local documents – c.400k words – contained in *A Linguistic Atlas of Older Scots* (LAOS 2008), dated between 1380–1500. These materials were accessed via the *From Inglis to Scots* Corpus (FITS, Alcorn *et al.* forthcoming), which provides a grapho-phonologically parsed version of the Germanic root elements in LAOS (see Kopaczyk *et al.* 2018), facilitating a triangulation of spellings, reconstructed sounds, etymological origins, and texts. Crucially, this allows us to map detailed spelling repertoires on to reconstructed sound values, and see how these are distributed over time

and space. The fact that we are dealing with local documents – writs, burgh records, land charters, etc. – is particularly relevant, given Stenroos (2004) has shown the spread of positionally/phonologically unbound <th> to be further advanced in NME documentary materials (see §3, above).

Our analysis of spelling variation in OSc materials relies on the well-established view that medieval scribal practices are overwhelmingly systematic as regards their graphemic repertoires, graphotactic distributions, and spelling-to-sound mappings (see Laing 1999, Laing & Lass 2003, 2009, Kopaczyk *et al.* 2018). This said, non-standard spelling systems may use a variety of graphs in order to represent a single sound and those same graphic elements may be used for multiple sounds as well. So, for any given sound we can reconstruct a *spelling substitution set* (e.g. Figure 10) and for every grapheme, a *sound substitution set*. The many-to-one and one-to-many patterns mean that a single word may be spelled in a multiplicity of ways depending on the graphemes selected from the spelling substitution set for each of its sounds.



kened 'known'; *scaith* 'scathe'; *nakkit* 'naked'; *quene* 'queen'; *kyrck* 'church'; *chynde* 'kind';
speccar 'speaker'; *kinrig* 'kinrik'; *brocczyn* 'broken'; *spechk* 'speak'; *aixman* 'axman'

Figure 10: A spelling substitution set for [k] across the FITS Corpus

Turning to our specific focus, Figure 11 shows the range of spellings for [ð] and [θ] in our corpus. We assume that the overall phonological distribution of voicing in ME is also in place in OSc, i.e. historically medial and function-word-initial fricatives are voiced,¹⁴ while elsewhere they are voiceless.¹⁵ Bubble size represents relative frequency of the grapheme. Line thickness represents relative frequency of the sound's association to the grapheme (blue/dark lines) and of the grapheme's association to the sound (mustard/light lines): this reveals that some associations are highly infrequent, e.g. [θ] is rarely spelled <hth>, although <hth> commonly signifies [θ]. Most importantly for our purposes, <th> appears to be slightly more frequent for [θ] than for [ð], while <y> is substantially more frequent for [ð] than for [θ]. A closer examination of the numbers follows.

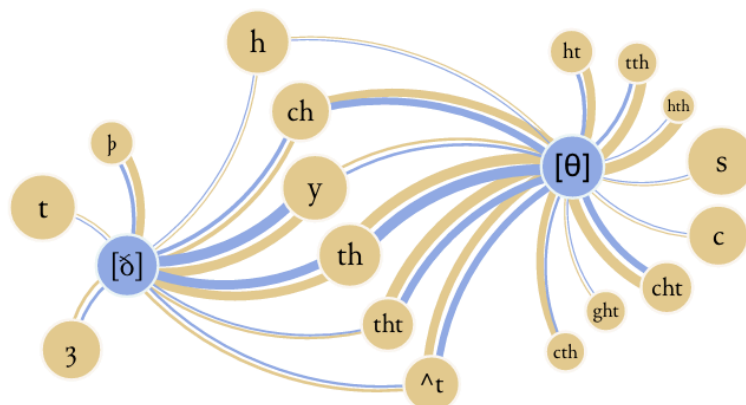


Figure 11: overlapping spelling substitution sets for [ð] and [θ] in the FITS corpus data (colour version available online)

5.2 Positional alternation among dental fricative spellings

¹⁴ The case of fricatives that became final due to schwa loss are probably an exception to this pattern as seen in spellings such as <luf> ‘love’ and morphologically related <luffis> ‘loves’. Such cases are discussed in Maguire *et al.* (2019) for labio-dentals; for dentals see §5.2.1 and §5.2.2 below.

¹⁵ *with* is omitted from our analysis since in PDE, where the fricative survives, it varies between voiced and voiceless realisations, while in Scots the fricative is mostly lost.

For this study, we extracted all spellings for etymological dental fricatives in the FITS corpus, distinguishing five categories: (1) initial position in content words; (2) initial position in function words; (3) morpheme-internal (‘medial’) position; (4) morpheme-final, pre-inflectional position; (5) word-final position. The distribution of spellings across the entire corpus is presented in Figure 12 (38,672 tokens). Following Benskin’s analysis for NME, we examine the relevant categories by the diachronic order in which, according to him, the <th> variant became dominant. We examine pre-inflectional cases – not considered by Benskin – immediately after those in final position.

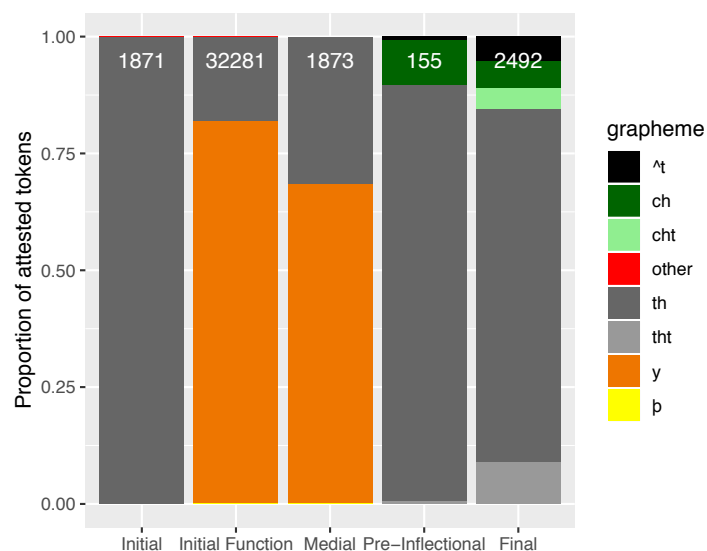


Figure 12: Proportions of dental fricative spellings in the FITS corpus by phonotactic position (total N per category at top of bars — colour version available online)

5.2.1 *Word-final fricatives*

According to Benskin’s proposal, NME saw the consolidation of <th> forms first in final position. In examining the rightmost column in Figure 12, we find the <y> grapheme is

almost completely absent, the only exception being one instance of <boye> ‘both’. Following the four-stage proposal for NME, therefore, the Scots documentary record begins having already completed the first stage.

What is surprising about this category – and this is something that Benskin and Stenroos do not mention – is that by late ME/Osc the ‘final position’ included many words in which the fricative had become final as a result of final vowel loss, e.g. verbs like *scaith* ‘scathe’ (<ON *skaða*), *freith* ‘frith’ (<OE *friðian*) and *couth* ‘could’ (<OE *cūðe*). Given that many present-day Anglic varieties have voiced fricatives in such verbs, this appears to be evidence for the misalignment of phonemic and graphemic systems in this earliest of Scots documentary evidence. However, as has been argued by Maguire *et al.* (2019) based on spelling differences for labiodental fricatives, loss of final schwa in early Scots did not necessarily lead to final voiced fricatives – at least not at first. Rather, it seems that the OE phonotactic constraint banning final voiced fricatives survived in Scotland for some time following the onset of schwa loss, resulting in a period where newly-final fricatives effectively devoiced (cf. Figure 13). This view is supported by the almost complete lack of <y> spellings for word-final dental fricatives in the FITS corpus.



Figure 13: Proposed diachronic overlap between constraints on fricative voicing and final vowel deletion, following Maguire *et al.* (2019: 54)

Compared to the data for final [f] and [v] in Maguire *et al.* (2019), the dentals show far less spelling variation in final position, both for etymologically voiceless, i.e. etymologically final, and etymologically voiced, i.e. etymologically [ðə], forms. Crucially, etymologically voiced forms are rare among dentals (22% of all final dental fricatives in FITS) as compared to the etymologically voiced labiodentals (77.5% of all final labio-dental fricatives in FITS, according to Maguire *et al.* 2019). Indeed, if the actual rate of <v> vs. <f>-type spellings in etymologically [və] contexts (24.5%) were transposed to etymologically [ðə] contexts, we would expect to find no more than 5.4% spellings representing voiced dentals (24.5% of the 22% etymologically voiced dental fricatives). It is therefore reasonable to assume that the rarity of this pattern, coupled with the lack of a clear-cut spelling contrast between [ð] and [θ] in the language more generally, probably facilitated the generalisation of <th>-type spellings in final position.

For the additional spellings that surface in final (and in pre-inflectional) positions, i.e. <tht>, <cht>, <ch> and <^t>, we follow Molineaux *et al.* (in press). To summarise briefly: the trigraphs <tht> and <cht> probably bear a hypercorrective final <t>, related to the more general process of /t/-deletion in the cluster /xt/ (Johnston 1997a: 101, Romaine 1984). However, it is unclear whether the excrescent element has phonic content or is merely orthographic. For the <ch(t)> forms, the claim is that these are likely a consequence of the confusability, in most OSc scripts, of <c> and <t>, coupled with the frequent fronting of /xt/ to /θ(t)/ (cf. Johnston 1997b: 505). As a result, transcribers of ambiguous <c~t> do not follow etymological lines in rendering such bi- and trigraphs. However, while this practice is justified for etymological /xt/ words, it is questionable for etymological /θ/. Ultimately this leads to cases where [θ] is mis-transcribed as <ch(t)>, which should be treated as equivalent to <th(t)>. Finally, <^t> (a superscript <t> in

manuscripts) is used consistently only for word-final position amongst historically voiceless dental fricatives and may be straightforwardly considered an abbreviation of <th>.

Given this analysis, word final dentals are likely overwhelmingly voiceless, and are represented consistently with the <th> digraph, its abbreviation, <ʰ>, and a transcribed form with a hypercorrective <ʰ>, which may or may not represent phonological substance.

5.2.2 Pre-inflectional fricatives

As a continuation of the OE voicing pattern, we would expect pre-inflectional fricatives to be voiced. However, as in final position, the only spelling category we find is <th> (and its variants <ch>, <tht> and <ʰ>), which has thus far appeared to represent voicelessness. In a close analysis of the pre-inflectional [v]~[f] alternation, Maguire *et al.* argue the predicted allomorphy between inflected roots (here e.g. *trouthis* ‘truth’s’, originally containing [ð]) and uninflected roots (e.g. *trouth* ‘truth’, originally containing [θ]) “is exactly the place we expect to see analogical levelling (Hock 1986: 167–171) ... from the basic form to the inflected form (e.g. [li:f]~[li:fəs])” (2019: 54). Given the sporadic nature of analogy, the authors find a pattern of variation in pre-inflectional spellings which reflects the proportions of voicing spellings in the absolute final position, from whence phonological voicelessness was presumably analogised.

Following Maguire *et al.* (2019), we would expect some variability in the voicing of pre-inflectional dental fricatives. However, the FITS data for dentals in this position lacks <y> spellings altogether, suggesting that, if there is any variation, this is not represented by the spelling. As in the word-final cases, this might be a result of the rarity

of voiced elements before an inflectional morpheme. Indeed, if we follow the findings for labio-dentals, we see that where the final fricative was unambiguously voiceless, i.e. in words like *lif* ‘life’, 86% of the corresponding pre-inflectional forms were spelled with the voiceless variant, <f(f)>, presumably as a result of a sound-based analogical levelling towards the uninflected root. Furthermore, among etymologically [və]-final words, the FITS pre-inflectional counterparts display 53% <f(f)>-type spellings. Extrapolating these proportions to the dentals, we would expect voiceless fricatives in about 94% of pre-inflectional tokens. Given this overwhelming majority of voiceless forms, maintaining a spelling distinction for the voicing contrast in this context would have been challenging for scribes. Recall, of course, that unlike <v> vs. <f>, a clear spelling indication of voicing distinction between <y> and <th> had not yet developed across the language. As a result, scribes probably turned to using <th>-type spellings exclusively to represent the dental fricatives in this position, which in practice were voiceless in the vast majority of cases.

5.2.3 *Word-initial fricatives in content words*

Stage II in Benskin’s proposal for NME dental fricative spellings entails the ‘spread’ of <th> to initial position in content words. Following the phonemicisation of the OE fricative voicing pattern, this category is expected to contain only voiceless dental fricatives, as there were no borrowings with initial [ð] to create a positional contrast. The expected consistency in the phonology for this category is mirrored by spelling of words such as *think*, *thatch*, *thing*, *thorn* and *threat* which begin almost categorically with <th> (Figure 12). Indeed, only two tokens of <y> are attested in this context, both for ‘thieves’ in the same text.

Following the purportedly stepwise spread of <th> in Benskin's Northern System, Stage II represents the consolidation of a phonemic contrast in the spelling of dental fricatives. The data in FITS suggest that this stage had already been reached from the start of the OSc period, since the two main positions where we would expect voiceless variants are spelled almost universally with <th>.

5.2.4 Medial fricatives

The picture provided by the FITS data for morpheme-internal fricatives – such as those in *other*, *gather*, *brother*, *smithy*¹⁶ – contrasts starkly with that for fricatives that become medial due to inflection (i.e. those in §5.2.2, where a morphological alternant exhibits a word-final fricative). Given that all lexemes for this category in FITS have an etymologically simplex, word-medial fricative, we would expect these to be realised as [ð], continuing the voicing pattern from OE. In fact, the majority (68.3%) of tokens in this position are spelled with <y> which suggests that Benskin's Stage II is still mostly intact, reflecting a phonemic spelling pattern. Following the NME analysis, we would expect that the <th> spellings in this category (31.4%) are a result of the continued spread of this spelling beyond the bounds of voiceless fricatives, entering Stage III and the collapse of the spelling contrast. Of course, such an analysis relies on the idea that <þ>, and later <y>, were at some point the main variant in this category – a claim that we have

¹⁶ A slightly problematic group of items crops up in this category, that is, words such as *mother*, *father* and *hither* which have their origins in OE forms with <d(d)> and continue to be spelled as such well after the FITS period. While the majority forms for these lexemes show the traditional <d(d)> spellings, the FITS corpus shows some of the earliest instances spelled with <y> and <th> which represent [ð] as the result of processes of pre-/r/ dentalisation and [dər] lenition, as described by Maguire (2016: 335). Here we include the cases spelled as <th> and <y> only and remain agnostic as to the possibility of some of the <d(d)> spellings representing [ð].

seen is hard to uphold, given the OE data in Figure 1 and the eME data in Figure 4. A more fine-grained look at the corpus-internal diachronic distribution of <y> and <th>, in §5.3.1 below, allows us to reconstruct a more plausible account for this development.

5.2.5 *Word-initial fricatives in function words*¹⁷

If the last category for consolidation of <th> spellings in NME is word-initial dental fricatives in function words, as claimed by Benskin, then we would expect this category to see the lowest rate of <th> in OSc. Such is the case in the FITS data, where fricatives are spelled as <y> in 81.6% of such tokens (Figure 12).¹⁸ Unsurprisingly, this position also preserves the largest number of <þ> spellings (n=114), in manuscripts where <þ> and <y> have not merged. Indeed, recall that word-initial dental fricatives in OE and eME are overwhelmingly spelled <þ>, a pattern most easily preserved in what are high-frequency items such as function words.

The overall pattern shows that <th> made few inroads into the spelling of OSc dental-fricative-initial function words, which is particularly striking in the light of the near-categorical shift to initial <th> amongst content words. As with the case of medials, we will need to look at corpus-internal chronological patterns in order to see how well the cases of <th> in the corpus fit the stepwise pattern proposed by Benskin (Figure 5).

¹⁷ As the FITS corpus excludes function words which do not have forms that also surface in major word categories (cf. ‘there’ in *yairefyr* ‘thereafter’), only *thence*, *there*, *thither*, *though* and *thus* were available via the FITS corpus tools. The remaining items (*the*, *this*, *that*, *they*, *them*, *their*, *thou*, *thee*, *thy*) constitute a far larger proportion (94.8%) of the data analysed here and were gathered directly from the LAOS database by searching for the relevant lexels. As a result, this latter group is not included in the spelling substitution set in Figure 11.

¹⁸ Interestingly, an important proportion of the <th> spellings in initial position for function words are those found in capital letters. Indeed, capital <y> for what we assume to be [ð] is rare across FITS data.

5.3 Corpus-internal diachrony and the Older Scots dental fricative spellings

A bird's eye view of relevant spellings across the entire FITS corpus, focusing only on position and grammatical category, yields a pattern not unlike that of a late Stage II in NME, where <y> and <th> may be reliably taken to represent voiced and voiceless fricatives, respectively. This is interesting insofar as it falls in with more general findings for this period: features of OSc are difficult to distinguish from the same features in the NME varieties south of the political border (cf. Williamson 2002: 253). Nevertheless, the FITS data also represent the period where emerging differences with NME should be visible.

Indeed, the fifteenth century might be precisely when we would expect to find conventions setting Scots apart from its southern neighbours. Uncovering such loci of change, however, requires a more fine-grained approach. Given that it is made up of documentary materials, for which we are generally able to retrieve specific dates and locations, the FITS corpus affords us this unique, targeted view. For the dental fricatives, we focus on medial and function-word-initial spellings, as these categories display relevant variation between <y> and <th>.

5.3.1 Medial fricatives over time

If we group the data for medial fricative spellings by decade of manuscript composition and view the results chronologically, we are able to obtain an overview of the proportions of <y> and <th> spellings across time

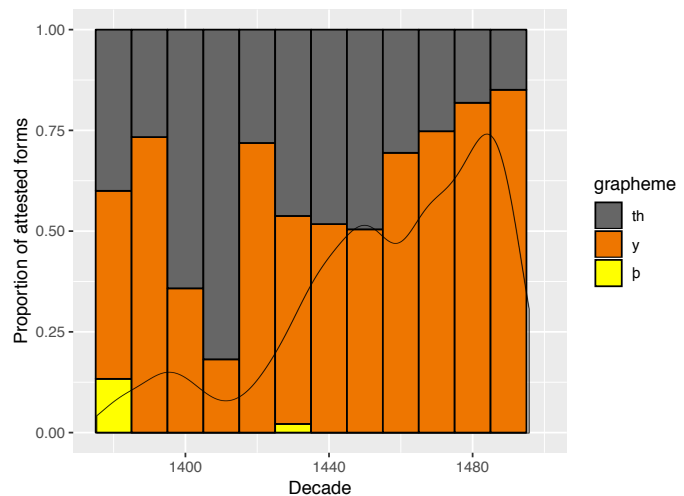


Figure 14: Proportions of word-medial fricative spellings in the FITS corpus, by decade and spelling. Density plot for all the data overlain.

(Figure 14). Importantly, as the availability of the data is limited for the earliest decades, a density plot for attestations across the entire corpus has been overlaid on the histogram, giving an idea of the reliability of the data by period.

The pattern that emerges is one where the first half of the period, while sparser in data, shows a greater proportion of <th> spellings. The final four decades of the corpus, nevertheless, where the data is more abundant and thus more reliable, see a sustained rise in the use of <y> over <th>, peaking at over 80% of the tokens in the final decade of the fifteenth century.

Needless to say, this is not the expected direction of change following Benskin's claims for NME. Indeed, we find that the earliest documentary evidence for Scots displays a mixture of <th> and <y> spellings in this category, tending to consolidate towards <y>. In other words, medial fricative spellings, rather than passively forming part of the contrastive spelling pattern by 'remaining' as <y> (or <p>), are actively shifting to <y>, which never appears to have been the default spelling for this category.

5.3.2 Word-initial fricatives in function words over time

A diachronic look at this category, which has the highest rates of <y> spellings (cf. §5.2.5), by decade (Figure 15) is initially somewhat difficult to interpret. Clearly, the proportion of <th> is not increasing in the period overall. In

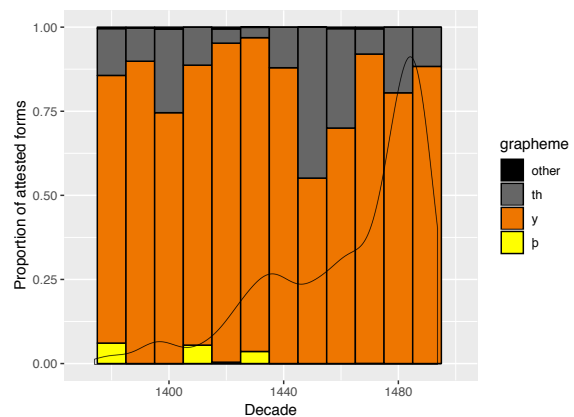


Figure 15: Proportions of word-initial fricative spellings in the FITS corpus, by decade and spelling. Density plot for all the data overlain.

fact, we find that the pattern is most robust towards the end, especially considering this is where we have more data to go on, as indicated by the peak in the density plot. Most importantly, the incidence of <th> is similar at the beginning and at the end of the corpus period, providing no evidence for a shift towards the Westminster-led pattern proposed by Stenroos (2007).

6. DISCUSSION: A NEW ACCOUNT OF THE DENTAL FRICATIVE SPELLINGS IN OLDER SCOTS

As we have seen, the overall distribution of dental fricatives by position across the FITS corpus (§5.1) shows spelling to be an excellent predictor of phonological voicing. In this sense, the OSc documentary data seems roughly consistent with Benskin's proposals for NME, placing the FITS period at some point in the transition from Stage II to Stage III (§5.2). However, this correspondence quickly breaks down once we examine the corpus-internal diachrony of the spelling variables (§5.3). The trend towards growth of <y>, which we find in word-medial position and the maintenance of <y>-dominance for function-word-initial positions, contradicts the general pattern of stepwise <th> spread

argued for by Benskin for NME (1977, 1982), highlighting that a naïve presupposition of <y>~<þ> as the traditional variable and <th> as the innovative variable is fundamentally at odds with the corpus data for OSc.

Our proposal for the development of the dental fricative spellings in Scots assumes that orthographic systems are largely conservative and, as such, they rely on previous spellings in the language, borrowing occasionally from conventions of other written languages in the scribes' repertoires. In this sense, the OSc spelling tradition has its roots in OE and eME practices alongside conventions from (Anglo-)Latin and varieties of French. Here we do not argue with Benskin's (1977) claims regarding the origins of the <y> and <th> spellings, but endeavour to underscore how their early distribution foreshadows their later phonemic specialisation. We suggest three abstract stages – which would have seen variation from one scribal language to another – allowing for the transition from OE to the late fifteenth-century system represented by the end of the FITS corpus (Figure 16).

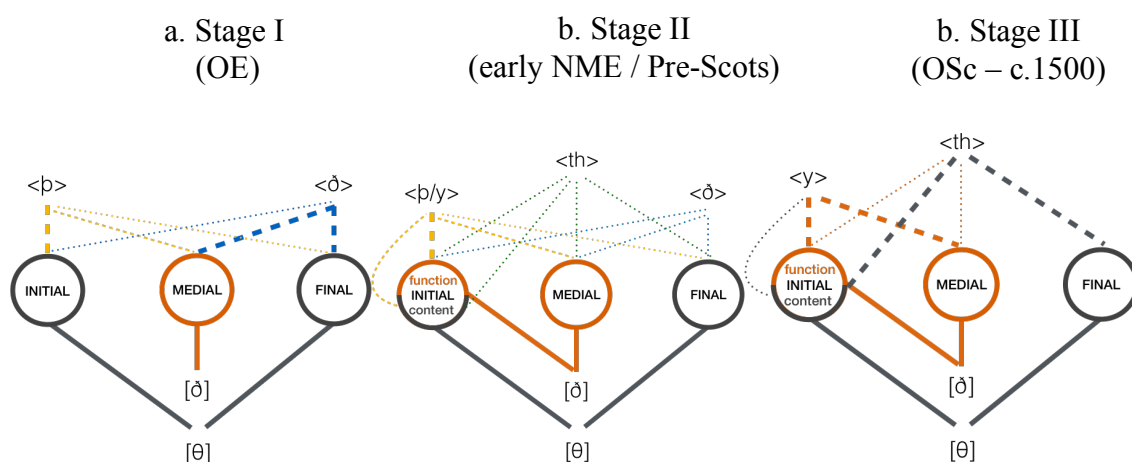


Figure 16: Proposed stages for the development of Scots dental fricative sounds (below – solid lines) and spellings (above – dotted lines) by position in the word. Main variants represented by thick lines, minor variants by thinner lines. Colours for spelling lines (online) represent independent variants with no distinct phonic values. By Stage III colours in spelling lines match those of the sound-association lines.

Going back to the OE period – our Stage I (Figure 16a) – we find that while the distribution of fricative voicing is governed by phonotactics, the spelling conventions are mostly based on what we have been referring to as *graphotactics*, that is, the positional distribution of graphemes in the word. Needless to say, these two domains remain misaligned throughout. Eventually, however, with the transition to eME – our Stage II (Figure 16b) – two things happened in the phonology: a) fricative voicing became phonemicised and b) initial dental fricatives voiced in function words. In parallel, changes in the power dynamics within the languages of Britain led to a reduction in the prestige and widespread use of Anglic varieties as written vernaculars. This brought about a reconfiguring of the spelling system such that non-Latin characters were dispreferred. A quick decline in the use of <ð> ensued in tandem with the gradual – though ultimately

incomplete – spread of Anglo-Latin <th> to all contexts for early NME/Pre-Scots.¹⁹ In contrast, <þ> (and later <y>) remained the main variant for initial position in function words, most likely due to their frequency of use.

We thus assume that the original graphotactic preference for initial <þ> was probably first to be pulled into a phonemically contrastive spelling pattern. Indeed, while frequency blocked the spread of <th> to initial position of function words, no such restrictions held for initial content words, which are generally far less frequent. In the same vein, the OE data shows that the least preferred graphotactic position for <þ> is final, leaving the field open for <th> to become the main variant, following <ð>-loss. This, in turn, would have matched a phonological position where fricatives were overwhelmingly voiceless either due to being historically final or becoming final due to loss of schwa and undergoing final devoicing (see §5.2.1). It is this pattern, where the dominant variant for initial voiced fricatives was spelled <y> and initial and final voiceless fricatives were mostly spelled <th>, that was taken by new generations of OSc scribes to represent a voicing contrast in the spelling. The result is the quick disappearance of <y> from voiceless contexts (initial content words, final and pre-inflectional fricatives) and its gradual spread into voiced contexts (initial function words and morpheme-medial fricatives) across OSc scribal systems. We thus reach the final stage of our narrative – Stage III (Figure 16c) – representing the end-state of the FITS data presented above.

While the data for medial and function-word-initial fricatives in FITS shows there is variation from text to text in terms of spelling choices, the trend we see in the data

¹⁹ Here, our focus is on the northern dialects that probably fed into Scots most strongly. The southern dialects, following Stenroos (2004) must have extended <þ> to all positions at some point during eME, only to receive more substantial influx of <th> spellings in the fifteenth century (Figure 11).

would predict that, left to its own devices, Scots would eventually have continued on the path towards a fully contrastive dental fricative spelling system.²⁰ As the data stand at the end of the fifteenth century, this is the closest case to such a pattern attested at any point in the historical development of Anglic languages.

7. CONCLUSIONS

In this paper we have shown that conventional patterns of grapheme distribution (i.e. graphotactics) can be co-opted by scribes in order to express new phonological contrasts. The history of dental fricative spellings in Scots illustrates this process, starting with the distribution of <þ> and <ð> in OE and leading to an emerging orthographic norm which consistently represents the OSc phonemic contrast between [ð] and [θ] via <y> and <th> respectively.

While a similar pattern to this one is claimed for late NME (Benskin 1982), Stenroos (2004: 275–9 and §3 above) has shown that spread of <th> to all positions, at least as early as the first half of the fifteenth century, would have led to a collapse of the phonemic spelling distinction, particularly among documentary material. The main reason for this collapse is considered to be the influence of southern scribal practices associated with an emergent London-based standard court orthography. In this context, we observe that the Scots scribes, rather than following the southern norm, elaborated on their own spelling practices instead. Indeed, we find that the *y-for-voicing* convention is

²⁰ Some sixteenth and seventeenth-century Scottish hands continued to use <y> in functional elements such as *yair* ‘their’, *yir* ‘there’, *y’* ‘that’ etc. well after anglicisation had taken root at other levels of the grammar and spelling, though this became progressively more of a relic convention.

on the rise precisely among documents, the textual genre – here represented by the FITS corpus – where it first eroded in NME.

This emerging systematicity of spelling is noted by Aitken (1971: 187), who claims that “[i]n general, much that we know of Middle Scots spelling and phonology does seem to square with a belief in a reasonable fit between the orthographic and phonemic systems over a good deal of their area.” In the case of voicing in dental fricatives this tendency towards phonemic spelling is particularly significant, since the phonological variants remain mostly predictable on the basis of positional or grammatical considerations. In other words, scribes enshrined the distinction between voiced and voiceless fricatives in the spelling despite the fact that it had a low functional load for disambiguating minimal pairs in the lexicon. This is an important finding as regards scribes’ metalinguistic awareness of contrast, which appears to be independent from word-level complementary distribution, thus underscoring that the Standard English unitary spelling is not the only possible representation of marginal phonemic contrasts such as these.

As regards the more general picture of the language’s historical development, we may say that, if at the turn of the sixteenth century Scots displays “features characteristic of the first phase of standardisation” (Meurman-Solin 1997: 3), the (near-)contrastive spellings for voicing in dental fricatives should certainly be one of those features. Such developments, of course, were eventually cut short by the influx of southern orthographic conventions via a general process of anglicisation.

Finally, the paper has also shown that, from a methodological standpoint, fine-grained analyses of historical corpus data are essential to understanding the relationship

between sounds and spellings over time. In this sense, there is need for corpora – such as FITS – that seriously consider spelling variation and its relation to sound-substance.

Authors' addresses

Department of Linguistics and English Language

The University of Edinburgh

Dugald Stewart Building, 3 Charles Street

Edinburgh, EH8 9AD

Scotland, UK

benjamin.molineaux@ed.ac.uk

r.alcorn@ed.ac.uk

w.maguire@ed.ac.uk

vkaraisk@exseed.ed.ac.uk

b.los@ed.ac.uk

Department of English Language & Linguistics

The University of Glasgow

12 University Gardens

Glasgow, G12 8QH

Scotland, UK

Joanna.Kopaczyk@glasgow.ac.uk

REFERENCES

- Adamczyk, Michał. 2016. Realisations of the word-initial variable (th) in selected Late Middle English northern legal documents. In Ewa Waniek-Klimczak & Anna Cichosz (eds.) *Variability in English across time and space*, 11-41. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Aitken, A. J. 1971. Variation and variety in written Middle Scots. In A.J. Aitken, Angus McIntosh & Hermann Pálsson (eds.), *Edinburgh Studies in English and Scots*, 177–209. London: Longman.
- Alcorn, Rhona, Vasilis Karaiskos, Joanna Kopaczyk, Bettelou Los, Warren Maguire & Benjamin Molineaux. Forthcoming. *From Inglis to Scots: A Corpus of Grapho-phonological Correspondences (1380-1500) with Associated Corpus of Changes*. <http://www.amc.lel.ed.ac.uk/fits/>
- Bald, Marjory. 1926. The anglicisation of Scottish printing. *The Scottish Historical Review* 23(90), 107-115.
- Bann, Jennifer & John Corbett. 2015. *Spelling Scots. The orthography of literary Scots, 1700-2000*. Edinburgh: Edinburgh University Press.
- Benskin, Michael. 1977. Local archives and Middle English dialects, *Journal of the Society of Archivists* 5(8), 500–514.
- Benskin, Michael 1982. Letters <p> and <y> in later Middle English, and some related matters. *Journal of the Society of Archivists* 7, 13–30.
- Benskin, Michael 1997. Texts from an English township in late medieval Ireland, *Collegium Medievale* 1–2, 91–170.
- Campbell, Alistair. 1959. *Old English grammar*. Oxford: Oxford University Press.

DOEC: *Dictionary of Old English Web Corpus*, 2009. Compiled by Antonette diPaolo Healey with John Price Wilkin and Xin Xiang. <https://tapor.library.utoronto.ca/doecorpus/> Toronto: Dictionary of Old English Project (accessed 03 March 2019).

eLALME, *An electronic version of A linguistic atlas of Late Mediaeval English 2013–*. Compiled by Michael Benskin, Margaret Laing, Vasilis Karaiskos & Keith Williamson, The University of Edinburgh. <http://www.lel.ed.ac.uk/ihd/elalme/elalme.html> (accessed 03 March 2019).

Jensen, Vibeke. 2012. The consonantal element (th) in some Late Middle English Yorkshire texts. In Jukka Tyrkkö, Matti Kilpiö, Terttu Nevalainen & Matti Rissanen (eds.) *Outposts of historical linguistics: From the Helsinki Corpus to a proliferation of resources*. Helsinki: Varieng e-Series. <http://www.helsinki.fi/varieng/series/volumes/10/jensen/> (accessed 15 September 2019).

Johnston, Paul. 1997a. Older Scots phonology and its regional variation. In Charles Jones (ed.), *The Edinburgh history of the Scots language*, 47–111. Edinburgh: Edinburgh University Press.

Johnston, Paul. 1997b. Regional variation. In Charles Jones (ed.), *The Edinburgh history of the Scots language*, 433–513. Edinburgh: Edinburgh University Press.

Jordan, Richard. 1934. *Handbuch der mittelenenglischen Grammatik*. Heidelberg: Carl Winter's Universitätsbuchhandlung.

Kniezsa, Veronika. 1997. The origins of Scots orthography. In Charles Jones (ed.), *The Edinburgh history of the Scots language*, 24–46. Edinburgh: Edinburgh University Press.

- Kopaczyk, Joanna, Benjamin Molineaux, Vasilios Karaiskos, Rhona Alcorn, Bettelou Los & Warren Maguire. 2018. *Towards a grapho-phonologically parsed corpus of medieval Scots: Database design and technical solutions*. *Corpora* 13(2), 255–269.
- LAEME: *A Linguistic Atlas of Early Middle English, 1150–1325*. 2013–, compiled by Margaret Laing <http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html> Version 3.2. The University of Edinburgh (accessed 03 March 2019).
- Laing, Margaret & Roger Lass. 2009. Shape-shifting, sound-change and the genesis of prodigal writing systems. *English Language and Linguistics* 13(1), 1–31.
- Laing, Margaret & Roger Lass. 2013. Introduction: Part II, The Corpus, Chapter 3: Overview. In *A Linguistic Atlas of Early Middle English, 1150-1325*, compiled by Margaret Laing. The University of Edinburgh. http://www.lel.ed.ac.uk/ihd/laeme2/laeme_intro_ch3.html [accessed 03 March 2019].
- Laing, Margaret & Roger Lass. 2003. Tales of 1001 nists: The phonological implications of litteral substitution sets in some thirteenth-century South-West Midland texts. *English Language and Linguistics* 7(2), 257-278.
- Laing, Margaret. 1999. Confusion wrs confounded: Litteral Substitution Sets in Early Middle English writing systems. *Neuphilologische Mitteilungen* 100(3), 251–270.
- LAOS: *A Linguistic Atlas of Older Scots, Phase 1: 1380–1500*, 2008. Compiled by Keith Williamson. The University of Edinburgh. <http://www.lel.ed.ac.uk/ihd/laos1/laos1.html> (accessed 03 March 2019).
- Lass, Roger & Margaret Laing. 2013. Introduction, Chapter 2. In Margaret Laing (comp.) *A Linguistic Atlas of Early Middle English, 1150–1325*, Version 3.2. University of Edinburgh. <http://www.lel.ed.ac.uk/ihd/laeme2/laeme2.html> (accessed 03 March 2019).

- Lass, Roger. 1992. Phonology and morphology. In Norman Blake (ed.) *The Cambridge history of the English language*, vol. 2. 1066-1476, 23-155. Cambridge: Cambridge University Press.
- Lass, Roger. 1991–3. Old English fricative voicing unvisited. *Studia Anglica Posnaniensia* XXV-XXVII, 3-45.
- MacKenzie, Laurel. 2018. Variable stem-final fricative voicing in American English plurals: different pa[ð ~ θ]s of change. *Language Variation and Change* 30, 147-174.
- Maguire, Warren 2016. Pre-R dentalisation in Scotland. *English Language and Linguistics* 20(2), 315–339.
- Maguire, Warren, Rhona Alcorn, Joanna Kopaczyk, Benjamin Molineaux, Vasilios Karaiskos & Bettelou Los. 2019. Charting the rise and demise of a phonotactically motivated change in Scots. *Folia Linguistica Historica* 40(1), 37-59.
- McIntosh, Angus. 1974. Towards an inventory of Middle English scribes. *Neuphilologische Mitteilungen* 75, 602–624.
- Meurman-Solin, Anneli. 1997. Differentiation and standardisation in Early Scots. In Charles Jones (ed.), *The Edinburgh history of the Scots language*, 433–513. Edinburgh: Edinburgh University Press, 3-23.
- Minkova, Donka. 2011. Phonemically contrastive fricatives in Old English. *English Language and Linguistics* 15(1), 31-59.
- Minkova, Donka. 2014. *A historical phonology of English. Online companion*. Edinburgh: Edinburgh University Press.
https://edinburghuniversitypress.com/media/resources/Historical_Phonology_of_English_-_Online_Companion.pdf Accessed 13 June.

- Molineaux, Benjamin, Warren Maguire, Rhona Alcorn, Vasilis Karaiskos, Joanna Kopaczyk & Betelou Los. In press. Sound change vs. orthographic remapping: Visualising ‘excrement’ <ɾ> and <ɾ> deletion in fifteenth-century Scots.
- Murison David. 1979. The historical background. In A.J. Aitken & Tom McArthur (eds.), *Languages of Scotland*, 2–13. Edinburgh: Chambers.
- Romaine, Suzanne. 1984. The sociolinguistic history of t/d deletion. *Folia Linguistica Historica* 5(2), 221–255.
- Scragg, Donald G. 1974. *A history of English spelling*. Manchester: Manchester University Press.
- Simpson, Grant. G. 2009. *Scottish Handwriting 1150-1650: An introduction to the reading of documents*. Edinburgh: John Donald Short Run Press.
- Stenroos, Merja. 2007. A Middle English mess of fricative spellings. Reflections on thorn, yogh and their rivals. In Marcin Krygier & Liliana Sikorska (eds.) *To make his English sweete upon his tonge*, 9-35. Frankfurt am Main: Peter Lang.
- Stenroos, Merja. 2004. Regional dialects and spelling conventions in Late Middle English: Searches for (th) in the LALME data. In Marina Dossena & Roger Lass (eds.) *Methods and data in English historical dialectology*, 257-285. Bern: Peter Lang.
- Strang, Barbara 1970. *A history of English*. London: Methuen & Co. Ltd.
- Williamson, Keith 2002. The dialectology of “English” north of the Humber, c. 1380-1500. In Teresa Fanego, Belén Méndez-Naya & Elena Posse (eds.) *Sounds, Words, Texts and Change*, 253-86. Amsterdam / Philadelphia: John Benjamins.