Formal Variation and Semantic Change in the Middle English Demonstratives

by

Yoko Takenouchi

Submitted in fulfilment of the requirements for the
Degree of MRes

School of Critical Studies
College of Arts
University of Glasgow

October 2015

Abstract

This thesis investigates the history of the Middle English demonstratives *these* and *those* with a strong emphasis on their geographical distribution and use of various forms found in texts in the late Middle English period.

Most of the forms that are used at present are surviving descendants of myriads of variants which emerged in the Middle English period, and the plural demonstratives *these* and *those* are no exception. However, it is known that each of these two distinct but similar-looking words has traced a different path in its development which does not allow a simple explanation. In particular, the emergence of the present-day standard form *those* remains wrapped in mystery owing to its complexity.

This study has attempted to gain a better insight into this intricate history using the extensive mass of data gained from the notable linguistic atlas called *eLALME*. Findings presented in this study demonstrate diverse aspects of functional selection of variables which occurred and then would produce the present-day system. Above all, this study has shown that the *A-curve* distribution pattern provides a useful clue to the evolution of linguistic variations.

Acknowledgement

First of all, I would like to express my deep sense of gratitude to Mr Tetsuya Shimizu, a lecturer in English at Waseda Prep School, for inspiring me to discover such a profound world as the English Language. It goes without saying that without an encounter with him I would not be so fascinated by this subject that I decided to study it abroad.

Also, I am very much thankful to Professor Yasuyo Moriya for giving me the thrill of the history of English. Although she was not my supervisor at university, my academic life at ICU would have been a totally different one if I had not met her. I am thankful that she also encouraged me to study abroad and gave me countless pieces of valuable advice. I also express deep and sincere gratitude to Professor Yuhri Ishimaru, whose earnest and passionate attitude towards research made her my first role model at all times even if the subject is different.

I am deeply grateful to my friends, including Dr Ichitaro Saito and Dr Nasreen Amin for their constructive advice on my application for University of Glasgow and Ms Ann Lorenc for her warm support on proofreading. Needless to say, I am indebted to my family, especially my grandparents, for supporting the long years of my academic life, financially and materially.

Lastly, I express the deepest gratitude to Professor Jeremy J Smith for his valuable guidance, great care and encouragement at various stages of my short but very intensive study and research period. It was such a considerable pleasure and great honour to study under him, one of the most eminent professors in this field. Although my first encounter with him was on a textbook, but for it, my wonderful life in Glasgow would not have existed.

Formal Variation and Semantic Change in the Middle English Demonstratives

Table of Contents

Appendix

Chapter 1: Introduction

1.1 Aims and Background

The aim of this study is to investigate the evolution of the Middle English equivalent forms of the Present-Day English plural demonstratives *these* and *those*. The primary focus is on the geographical distribution and use of various forms found in texts in the late Middle English period (c. 1300-1500).

Language is not a perfect system. Although sometimes we are tempted to talk about language almost as if it is an actual substance, every language is inherently dynamic and its system is constantly changing and developing. This language change takes place suddenly over myriads of utterances which are fundamentally unpredictable, and the precise moment when a certain linguistic item becomes a factor of evolution is hard to determine if not impossible. Nevertheless, it is possible to trace in detail the complex processes involved, especially during the mediaeval period, using the rich body of surviving data. Such data is increasingly available for analysis because of the work of linguists and philologists over the last half century, most notably those working in the *Linguistic Atlas* tradition (e.g. McIntosh et al 1986; Laing et al 2014).

The Middle English demonstratives are the perfect example to show these complex processes involved in such linguistic change. Middle English is the period which demonstrates the greatest variation in writing owing to the lack of a standard form either written or spoken. It is presumed that this reflects variation in the spoken mode and, as has often been pointed out (e.g. by Samuels 1972), variation is a key factor in this linguistic change.

A demonstrative is, according to *Oxford English Dictionary* (*OED*), ‘a word used to indicate the location (spatially, temporally, or abstractly) of something or someone in relation to the discourse context; especially a demonstrative pronoun or determiner (as in Present-Day English, ***this***, ***that***, ***these*** and ***those***)’. These words are seen in the following sentences.

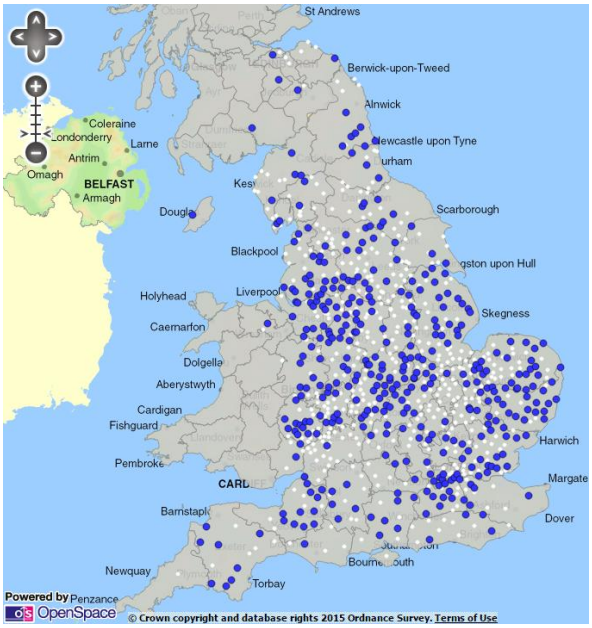
(1) *This* tree is about as high as *that* tree.

(2) *These* flowers are better than *those* which we planted last year.

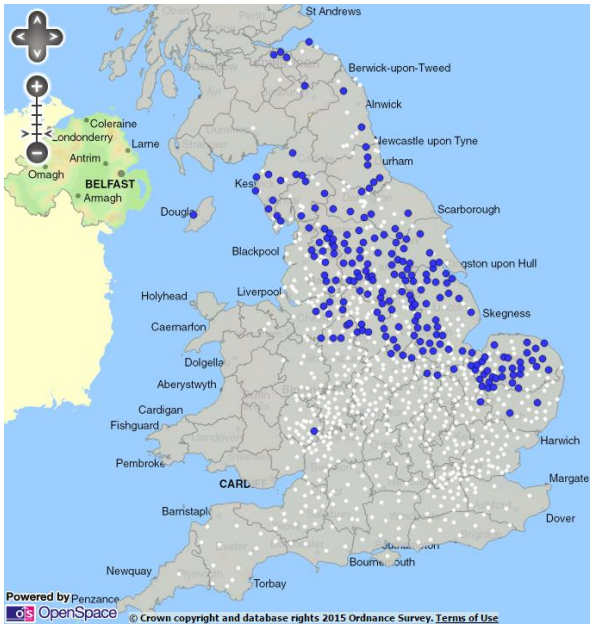
A noticeable feature of these words is that they are highly vulnerable to diachronic processes of analogy and functional selection (Smith 2012: 595). Middle English has numerous variations in form for the demonstratives, as demonstrated by *An Electronic Version of Linguistic Atlas of Late Mediaeval English* (*eLALME*), the most recent major output from the *Linguistic Atlas* tradition, which provides the largest amount of dialectal evidence yet assembled for forms of the items THESE and THOSE. There are recorded within *eLALME* at least 146 different tokens for THESE, ranging from widespread *thes* and *þese* to very local *dyse* (found in Sussex only), and no fewer than 65 items for THOSE, such as prevalent *þo* and very local *thaye* (found in Devon only).

The distribution of these various forms of demonstratives reveals an interesting and rather complicated history. Some intriguing examples are shown in figures below which are derived from *eLALME*. Map 1 and Map 2 show the distribution of texts containing the word for THESE which begin with <th> and <y>, respectively; Map 3 and Map 4 show the distribution of texts containing the word for THOSE which begin

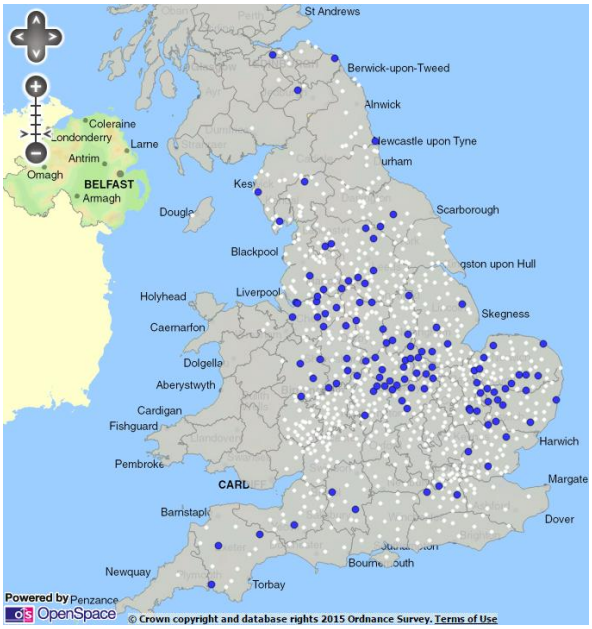
with <th> and <y> respectively. Each white dot represents a survey point where a particular document is originally written; a blue dot signifies a document in which each variation is found, regardless of frequency.



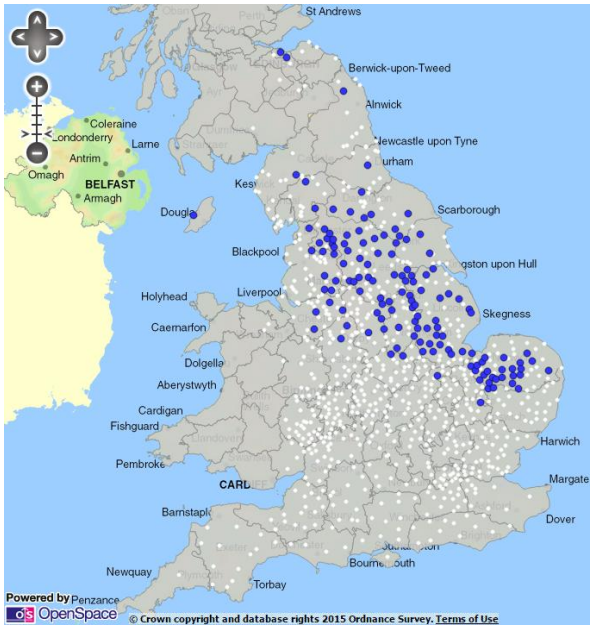
Map 1: THESE with initial <th> (e.g. **th**eer)



Map 2: THESE with initial <y> (e.g. **y**air)



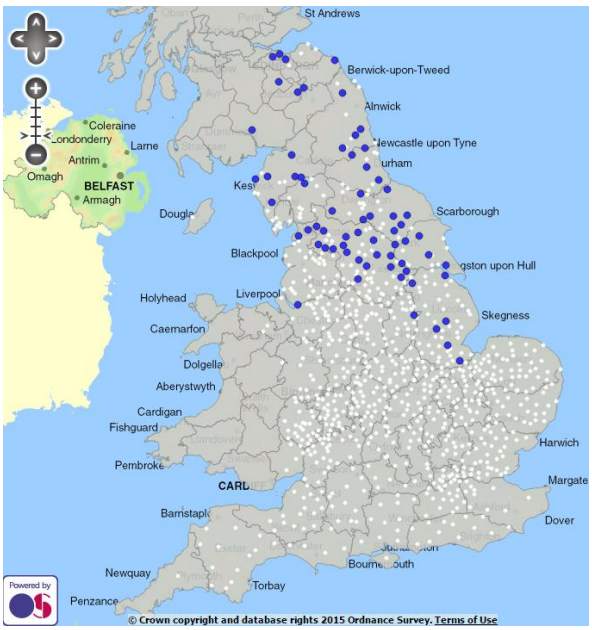
Map 3: THOSE with intial <th> (e.g. **th**o)



Map 4: THOSE with initial <y> (e.g. **y**a)

Interestingly enough, while initial <th> THESE seems to have been used throughout

most parts of the survey area, initial <y> THESE can only be found in the limited region of northern and some eastern parts of England; the same thing applies to THOSE. The distribution of THESE ending with <r> displays the similar tendency to the initial <y> THESE and THOSE as shown in Map 5.

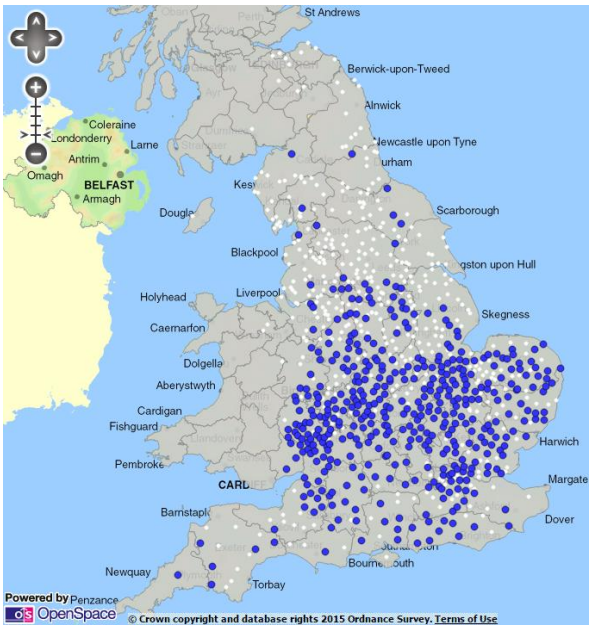


Map 5: THESE with ending <r> (e.g. thee**r**)

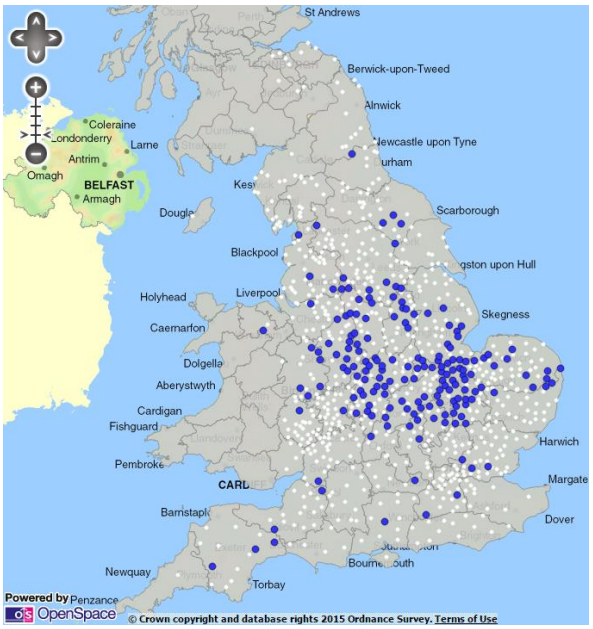
Benskin (1982) points out that this regional usage of <y> is due to the conflation of the letters <y> and <þ> by scribes, especially in northern and some parts of eastern England, where <y> is used in place of the <þ> as the graphemic expression of dental consonants (e.g. <yai> *they*, <yan> *then*, and <y^{e}> *the*). According to Benskin, it seems that a variety of the letter <y> had evolved quite independently of English script; <þ>-like <y> is found in continental Latin manuscripts from at least the later twelfth century and becomes relatively common in the fourteenth and fifteenth centuries when this usage was transferred to vernacular texts.

The distribution of the initial <þ> for THESE and THOSE shown in Map 6 and Map 7

below illustrates the relatively less frequent distribution of <þ> forms in the North, which suggests the dominant use of <y> forms by the northern scribes.

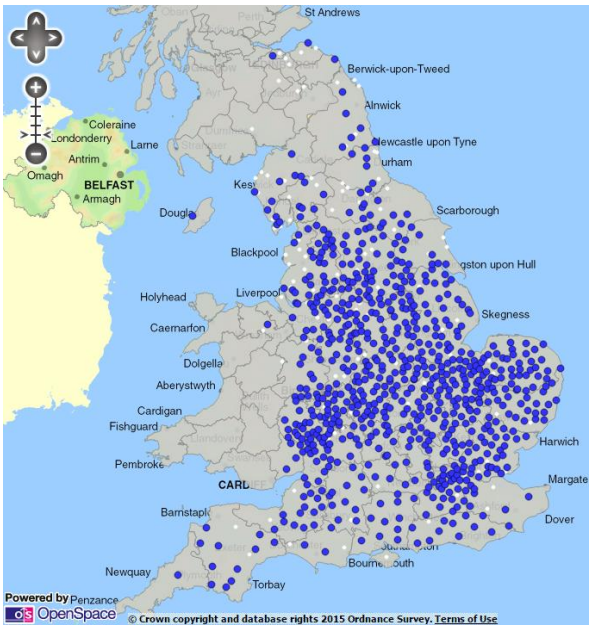


Map 6: THESE with intial <þ> (e.g. **þ**ees)

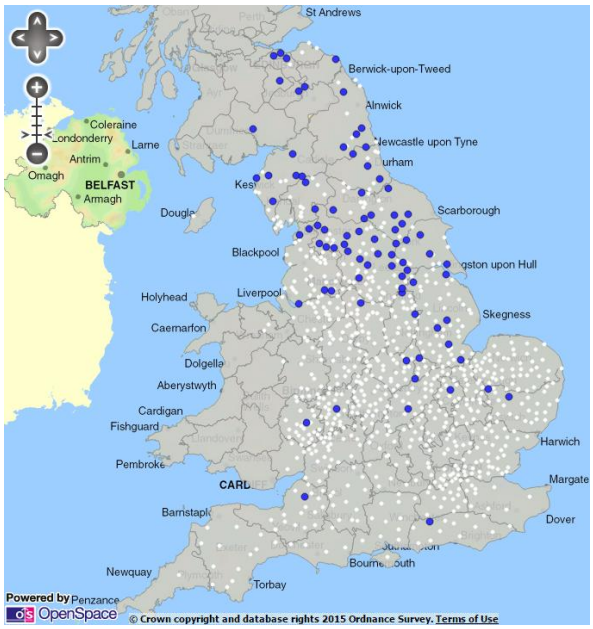


Map 7: THOSE with initial <þ> (e.g. **þ**a)

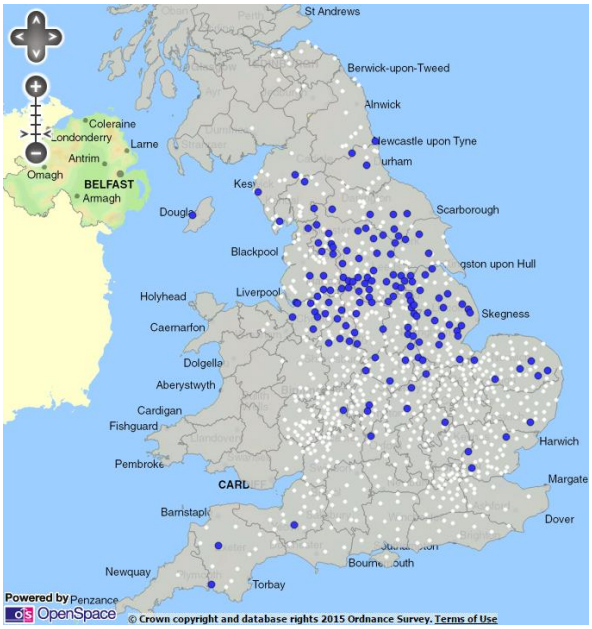
Another interesting issue is that, as seen in Map 8 and Map 9, THESE which ends with <s(e)> such as *thees* and *þeose* is found throughout the survey points on the map whereas THESE without the ending <s> can only be found mainly in the northern parts of England and southern Scotland, similar to the distribution of *r*-form THESE in Map 5. It seems that <s(e)> is a widespread plural marker for THESE. On the other hand, the equivalent distribution of THOSE, illustrated in Map 10 and Map 11 respectively, offers the fact that tokens without the ending <s(e)> are more dominant for THOSE.



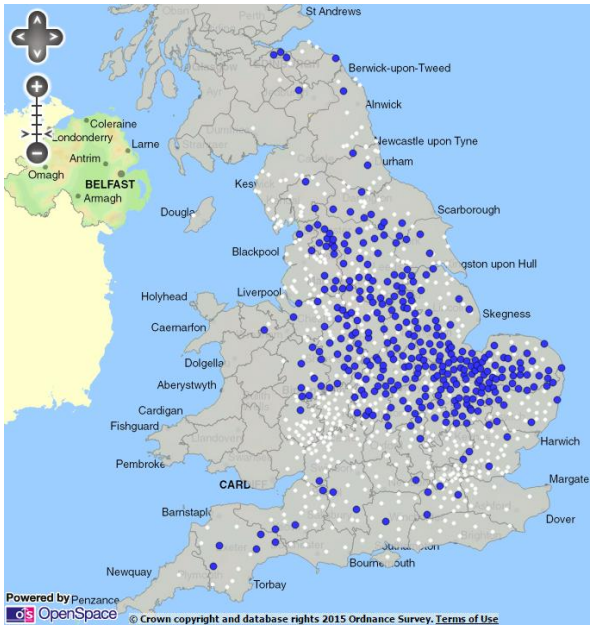
Map 8: THESE with ending <s> (e.g. thy**s**)



Map 9: THESE without ending <s> (e.g. the**r**)



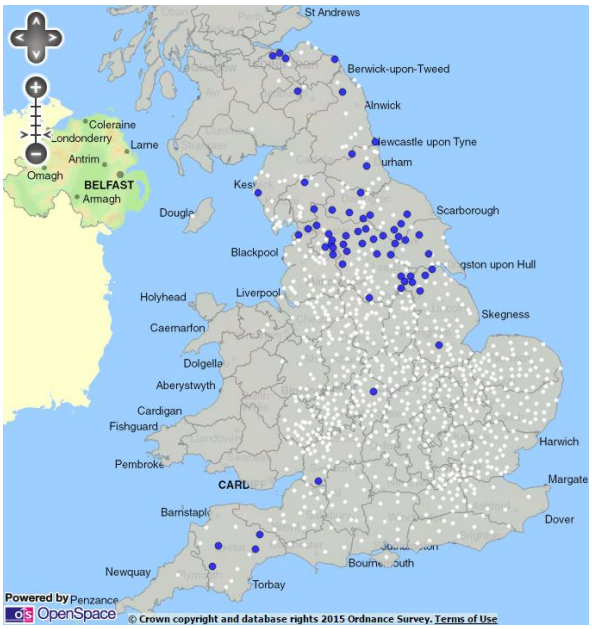
Map 10: THOSE with ending <s> (e.g. tha**s**)



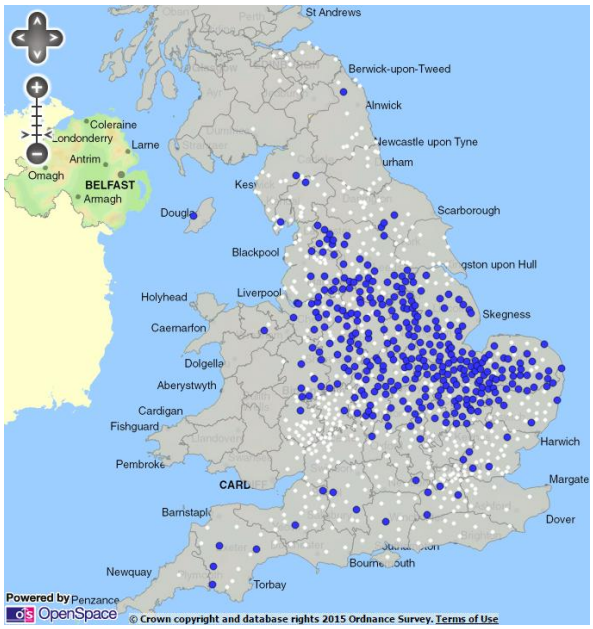
Map 11: THOSE without ending <s> (e.g. y**a**)

Lastly, these two figures below illustrate the comparison of the distribution of THOSE with ending <a> (Map 12) and <o> (Map 13) respectively. The texts which contain either *a*-type THOSE or *o*-type THOSE are found in a limited area, especially *a*-type THOSE being found only in the North. This indicates that the sound [ā] of *þā* in Old English developed differently in the North and the South, showing an important key

which has led to the modern form *those* from the historical equivalence.



Map 12: THOSE with ending <a> (e.g. th**ar**)



Map 13: THOSE with ending <o> (e.g. th**oe**)

The diachronic semantic development of demonstratives is similarly complex. It has long been recognised that the main functions of demonstratives are as markers of definiteness and deixis. Epstein (2010), on the other hand, argues that the distal demonstrative determiner in *Beowulf* also serves a variety of discourse-pragmatic functions, such as highlighting the importance of main characters and indicating chapter boundaries. Novelli (1957) asserts that Chaucer's use of the demonstrative adjective *this* conveys 'colloquialism' and 'informality'. However, this argument seems invalid considering the fact that the syllable-stress verse which Chaucer adopted in most of his work is not likely to reflect the rhythm of ordinary spoken English (Attridge 1995: 99). Smith (2012) also questions this notion by claiming that the colloquialism of the demonstratives does not give an explanation of the semantic change of the English demonstrative system, that is, the shift from a more/less emphatic distinction in the sense of this/that in Old English and Middle English, to the present-day

proximal/distal distinction. He explains that the more/less emphatic distinction in Middle English is also found in, and derives from, early Germanic dialects, as evidenced by Gothic.

1.2 Research Questions

In the light of these research contexts, this study will seek to answer the following research questions:

- How have the various forms of Middle English plural demonstratives *these*/*those* evolved throughout the English-speaking area?
- How have semantic changes in the Middle English plural demonstratives *these*/*those* evolved over time?
- What kind of extralinguistic/intralinguistic factors affected their usage and forms?

These questions will be addressed, although in a preliminary way constrained by the available extent of the thesis, in particular through the close analysis of a major Middle English text that survives in several manuscripts and dialects, viz. the *Cursor Mundi*.

1.3 Methodology

As noted above, *eLALME* draws upon a huge corpus of texts written in local usage in order to produce a set of maps. This database covers documents and literary texts found mainly in England and in some parts of Wales and southern Scotland written approximately between 1350 and 1450. This particular period is between the time

when the historical linguistic system of Old English collapsed and Standard English emerged (Benskin et al 2013), made possible by a variety of political, social and economic factors (McIntyre 2009: 47) from the fourteenth century onwards. During this period, local usage is still regularly reflected in writing by most scribes.

There is an important issue to be noted regarding a methodology for surveying the texts for *eLALME*, however. *eLALME* (and its printed predecessor *LALME*) has a collection of Linguistic Profiles (LPs) which are the result of examining texts for the occurrence of a set of predetermined linguistic criteria: the questionnaire. Each LP records all the forms for the items listed in the questionnaire. The problem here is the original questionnaire used in *LALME* was, for various reasons, made up of a merger of different questionnaires: the northern area of survey (north of the Wash) and the southern area of survey (south of the Wash). Items for the southern survey were listed originally in the separate Appendix and were not usually collected systematically. This disparity is very relevant to the present study, since the items for THOSE are collected using these different questionnaires, and as a result, not all variations for THOSE are fully covered for the southern texts. In short, the data for THOSE is not adequate and needs to be supplemented through further textual analysis.

As Heltveit remarks (1953: 14), until his time none of the questions regarding the history of demonstratives can be said to have been dealt with in an adequate way, and his claim that the development of demonstratives has only been touched upon more or less ‘incidentally’ remains surprisingly valid today. Furthermore, such studies have treated the period before 1300 and after 1475 exclusively, leaving the critical period untouched. In this way, there has been insufficient research on English

demonstratives to puzzle the problems of this important period between 1300 and 1475, when a dialect spoken in the London area gradually established its supremacy over the other regional dialects as the accepted means of written communication. It is exactly the period which *eLALME* deals with.

eLALME, like its predecessor *LALME*, offers a body of decontextualised data. The intention of this thesis is to place such data back in context and then investigate the deployment of demonstratives in a widely circulated text that survives in several manuscripts, namely the anonymous *Cursor Mundi*. *Cursor Mundi* is a long poetic version of Biblical history of some 30,000 lines that survives in nine manuscripts copied in a wide variety of Middle English dialects (Thompson 1998: 1). Four of these manuscripts, which are written in four different dialects of the northern and Midlands areas and provide an almost full text, form the textual basis for the present study. Comparing the various uses of demonstratives across parallel texts allows for the distinct patterns of demonstrative behaviour to be distinguished, and for our understanding of the various systems in various dialects used to be refined. Moreover, given its length, *Cursor Mundi* also supplies a substantial body of additional data for both kinds of demonstrative.

Chapter 2: Demonstratives in the History of English

In order to understand the deployment of the Middle English demonstratives, it is of great importance to place them in a wider context, both extralinguistically and intralinguistically.

2.1 A history of English and the Scandinavian influences

The earliest form of English is Old English, which emerged from the Germanic languages spoken by the Anglo-Saxons who settled in Britain in the mid-fifth century. Old English was to be spoken for over 600 years, during which it underwent constant change. Although the Anglo-Saxons succeeded in conquering the native Britons, they themselves were also exposed to attack from a series of raids by Scandinavian invaders between the late eighth and early ninth centuries. During this period, their Scandinavian language, i.e. Old Norse, had a great impact on the development of Old English. Close contact between the Scandinavians and the Anglo-Saxons led to major developments in vocabulary and grammar in particular, although the impact of Norse on English was largely hidden because of the dominance in the written record of Late West Saxon and the later influence of French and Latin. For this reason, we still do not have an adequate understanding of the degree to which these two peoples were mutually intelligible (Fell 1982: 88) and to which Old Norse had an impact on the 'internal' system of English as well as the 'product' of the language contact (that is, loanwords and place-names) (Townend 2002: 9).

Leaving aside military confrontations, Anglo-Saxon texts from the ninth to eleventh centuries record many peaceful encounters between English and Norse speakers since the first attack by the Vikings in 787. Furthermore, the contact between the

Anglo-Saxons and the Scandinavians must have been frequent or even daily, either individually or in groups, in various occasions. Within areas of the Danelaw, the members of the court must have been bilingual (Townend 2002: 8), and also there must have been plenty of regular contacts between merchants and Anglo-Norman families resulting from intermarriage.

There are quite a few scholars who have pointed out the resemblance of Old English and Old Norse; for instance, Jespersen (1956) remarks that an enormous number of words were identical in the two languages, and therefore the two peoples would have no great difficulty in communicating with each other[1]. However, it is not very hard to envisage that there must also have been a linguistic 'compromise' between speakers of two similar but different languages for a better mutual understanding.

A number of phenomena noted by Trudgill (1986) with regard to the development of dialects have occurred to the demonstrative systems as a result of the contact between speakers of Norse and English. In particular, the most prominent features are 'the process of *levelling*, which involves the loss of marked and/or minority variants; and the process of *simplification*, by means of which even minority forms may be the ones to survive if they are linguistically simpler' (Trudgill 1986: 126). The detailed discussion about these processes which happened to the English demonstrative systems will be made in the later sections; but what should be noted here is the idea that not only minor variants but also marked ones may disappear over the course of evolution. It is what exactly happened to the paradigm systems of the English demonstratives, and the fact the latter notion made by Trudgill is well

[1] Indeed, some of the elements of the demonstrative paradigms of Old English and Old Norse are very similar: for instance, *se* and *sa*, *þæt* and *þat*, *þes* and *þessi*.

demonstrated especially in the development of THOSE.

The Norman invasion of England in 1066 had far-reaching consequences for the development of the English language, leading to the next stage: Middle English. After the conquest, French became the vernacular language of the Royal Court, and Latin the language of administration and religion, making the status of English downgraded among the elites in particular. Although the impact of French on the English language is significant for the later development of the language, it is safe to say that it is less important for the development of the demonstrative systems. The Conquest means, in this context, the end of the dominance of Late West Saxon and the appearance in the written record of developments, including hitherto hidden Norse-derived forms. Amongst such 'hidden' forms were distinctive, dialectally-distinguished, demonstrative systems.

2.2 A history of English demonstratives

The Germanic languages possess a diverse range of particles functioning as definiteness/indefiniteness, proximal/distal, or for the purpose of emphasis, used anaphorically/cataphorically and in most cases pronominally (Smith 2012: 594). Such systems can be observed in the earliest stages of Germanic languages, such as Gothic, which survived approximately between the third and the eighth century, and other contemporary languages, such as Old Norse and Old High/Low German (Millar 2000). Each of them has an elaborate system of demonstratives, distinguishing between simple/compound, gender, singular/plural and so forth, and shares the same fundamental functions. There are slight differences in the structure among these

earliest Germanic dialects, however. Millar (2000) classifies them into three basic formal patterns which have: (1) simple demonstrative alone (distal), (2) simple (distal) and compound (proximal) demonstrative and (3) simple (distal), compound (proximal), and discrete definer.

Type (1) is only found in the written record of the Gothic language. There is some evidence that a suffix *–uh* was employed to the simple paradigm to give proximal meaning (e.g. Wright 1954: 266; Krause 1958: 181). Although it would be possible that all Germanic dialects might have shown such system at some point before the written record, this 'one-dimensional' system seems not to have survived into later Germanic dialects (Millar 2000: 17).

Type (2) paradigm is present in all of the West Germanic dialects in the earliest evidence, including Old High German, Old Low German (Old Saxon) and Old English (Millar 2000: 18). Old English abandoned the distinction of gender in the plural in preliterate times, in contrast to other dialects which display distinctions of genders for both of singular and plural (Heltveit 1953: 69). It seems that each of these languages developed article function within its simple demonstrative paradigm during its history (Millar 2000: 21).

With respect to type (3), only the North Germanic languages appear to have evolved this tripartite formal split which is semantically similar to the *this-that-the* distinction found in Modern English (Millar 2000: 22). All West Germanic dialects have moved towards the system closer to this, and in the case of Old English, the development of articles, both definite (*the*) and indefinite (*a*/*an*), was completed during the Middle

English period (Smith 2012: 598). Despite the fact that English determiners have always occurred as a premodifier, more common is an enclitic usage as observed in the North Germanic groups such as Old Norse (the detailed mechanism of this system is not going to be discussed at length in this study).

Table 1 and Table 2 below show the Old English demonstrative paradigms, which were marked for case, gender and number. Although West Germanic languages including Old English did not originally have a division of two deictic functions like Gothic, compound demonstratives appeared around the fifth century as a result of the need for stronger deictic pronouns (Heltveit 1953: 65). As for the initial consonant, the *s*-forms in the simple demonstrative gradually took on a semantically specialised role, e.g. *se ðe*. However, the *s*-forms were eventually replaced by *þ*-forms under the influence of analogy with all other forms of both diagrams by the Early Middle English period (Millar 2000: 205).

The paradigm of compound pronoun is made up of two types of word formation: that is, an *archaic* type and a *new* type. In the archaic type, the ones shaded in grey below, deixis is expressed by the corresponding form of the simple pronoun + the deictic particle *s* (e.g. þas = þa + *s*), regardless of case, gender and number. In the new type, on the other hand, deixis is expressed by a new stem + the corresponding inflectional ending of strong adjective (e.g. þæs → þiss + *es* = *þisses*). In Old English, there was considerable agreement between noun and adjective in two discrete paradigms (Millar 2000: 29), and there seemed to be no exception for the pronoun. The reason why two different systems are found within one paradigm is owing to the requirements of the system itself, according to Heltveit (1953: 68). Taking the nominative masculine

singular *þes* for instance, because the corresponding adjective has no inflectional ending, and the simple pronoun *se* is not possible to become either stem or formative, the archaic type has to remain in this paradigm.

	Masculine	Feminine	Neuter	Plural
Nominative	þes	þeos	þis	þas
Accusative	þisne	þas	þis	þas
Genitive	þisses	þisse	þisses	þissa
Dative	þissum	þisse	þissum	þissum

Table 1: Old English compound demonstratives for PDE ***this*/*these***

	Masculine	Feminine	Neuter	Plural
Nominative	se	seo	þæt	þā
Accusative	þone	þā	þæt	þā
Genitive	þæs	þære	þæs	þara (þæra)
Dative	þæm	þære	þæm	þæm (þam)

Table 2: Old English simple demonstratives for PDE ***the*,*that*/*those***

Millar (2000) explains in great detail the process by which this intricate system of case and gender inflection of demonstratives emerged and collapsed. According to Millar, major linguistic changes including the development of pronominal paradigms occurred in the 'transitional period' which started in the late Old English period. Especially, the contact between Norse and English in the northern England generated 'ambiguities' in the paradigms which led to the drastic simplification and reinterpretation of them during the period.

During this transitional period, Old English nominal/adjectival/pronominal morphology was under threat of system breakdown because of changes in the sound system of Old English; Old English inflections became mostly unstressed in spoken language

(McIntyre 2009: 13). First of all, a system based on morphemes functioning as case distinctions gradually broke down to the extent that formal integrity comes into question (Millar 2000: 161-162). For example, the distinction between accusative masculine singular *–ne* and dative masculine/feminine singular/plural *–m* in the compound paradigm was the most central ambiguity. Accusative masculine singular *þisne* and dative masculine singular *þissum* in the compound paradigm began to fall together at *þisse* (Millar 2000: 167), and the ending morphemes are eventually eliminated completely, being settled down at *þis*. Similarly, counterparts in the simple paradigm *þone* and *þæm* lost their case markers, falling together as *þe*.

Subsequently, the gender distinctions between *þe*(*s*)/*þeo*(*s*)/*þa*(*s*) distinguished only by their root vowels also fell together with each other, since, especially in unstressed contexts, they would have been pronounced in almost exactly the same way (Millar 2000: 208). It is quite possible to imagine then that the vowels of these three forms fell together at schwa /ə/ (Millar 2000: 214). In addition to this sound simplification, formal dislocation in the paradigms also occurred at the same time. It is easily envisaged that genitive masculine/neuter singular *þisses* with the case distinction marked by *–es* became *þis* due to the loss of the ending. Genitive masculine/neuter *þæs* in the simple paradigm was lessening in usage, being replaced by *þe* in the end (Millar 2000: 239). When it comes to the forms with inflectional endings containing *r*, that is, genitive/dative feminine singular and genitive plural, it seems that they were left unchanged for a certain amount of time because of their uniqueness, according to Millar (2000: 238). However, their separateness in the paradigm was also too marked for the forms to survive any longer (Millar 2000: 238).

In this way, the effects of ambiguity emerged in both paradigms of the demonstrative pronouns led to the significant formal simplification, *levelling* in Trudgill's term (1986). This transformation of pronominal paradigms as well as nominal and adjectival ones became a critical element which made the nature of the English language analytic in the Middle English and onwards from synthetic in Old English (Millar 2000: 11). In other words, the English language was to become free from such an elaborate grammatical case/gender system in exchange for the loss of flexibility in word order.

The Present-Day English standard forms, *these* and *those*, derive from the numerous variants created in the new system of Middle English. It is still very difficult, however, to trace their establishment because more than one derivation is possible (Heltveit 1953: 82). Moreover, the fact that *these* and *those* have developed independently also makes it intricate to uncover their processes of evolution. For these reasons, in the following sections the processes by which the modern forms were established will be discussed separately.

2.2.1 A history of compound pronouns

When inflection fell into disuse, the nominative accusative neuter *þis* came to be used as a singular regardless of gender and case (Heltveit 1953: 73). When the distinctions of gender and case were also abandoned in the strong declension of adjectives, the only distinct pattern on which the declension of compound pronoun could be remodelled was the plural indicated by the ending *–e* of adjectives (e.g. *gōd*: singular – *gōde*: plural). After the inflectional collapse, it was a very productive formative in Middle English (Heltveit 1953: 85). Consequently, <*þis*: singular – *þise*: plural> was

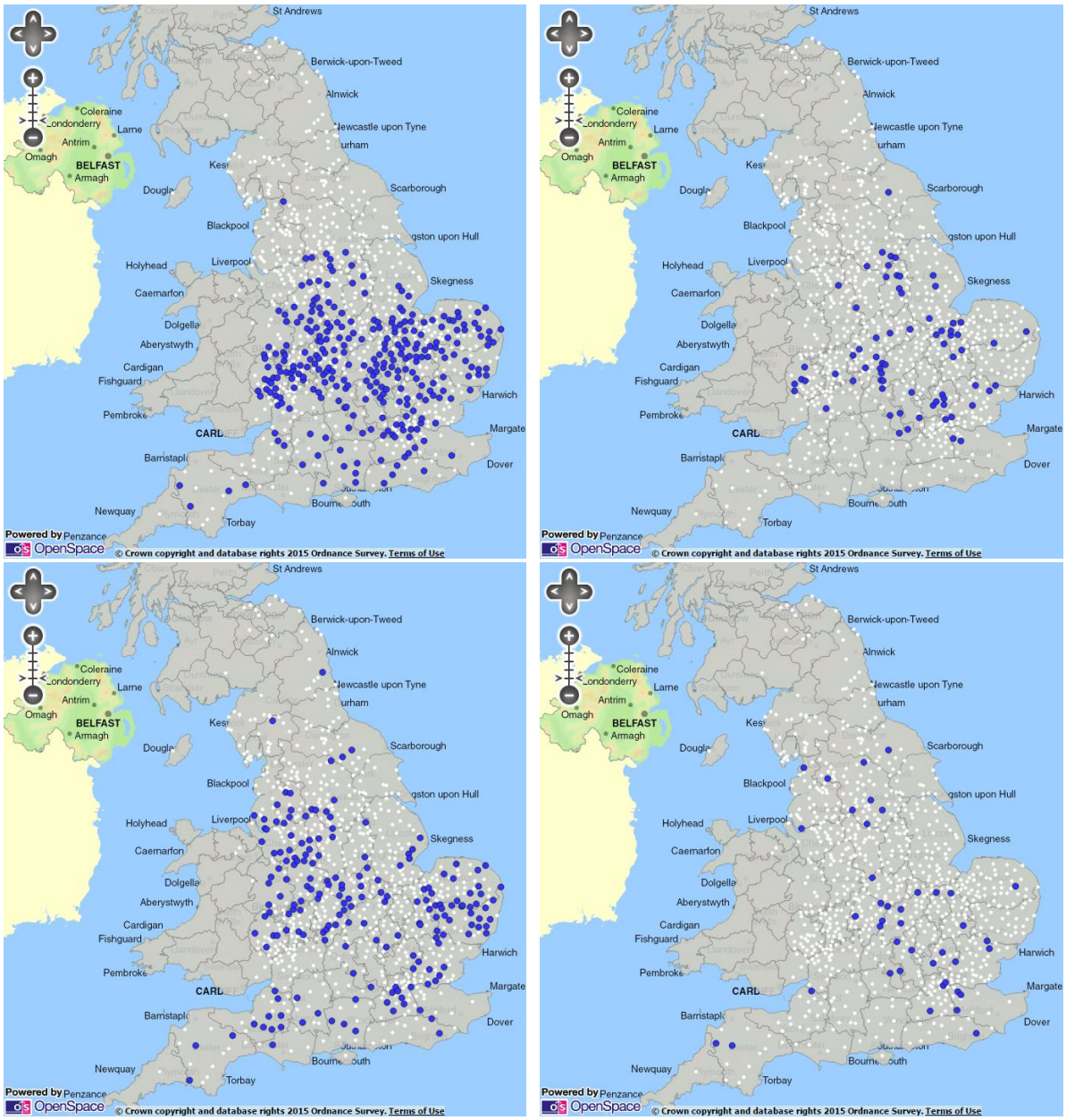
created, and soon the archaic *þas* was replaced with this new form *þise* (Heltveit 1953: 74).

This process may seem quite simple, but the reason why the *i*-type THESE was eventually chosen by discarding the archaic *þas* remains unsolved. Smith (2006: 12) argues that phonaesthesia could be a useful clue to understand this selection. Phonaesthetic association of front/close vowels with nearness and back/low ones with farness, such as *me*/*you* and *here*/*there*, has been discussed by many scholars (e.g. Pinker 1994; Firth 1964). For this reason, it would be at least arguable that this front/back formal distinction was favoured for the selection of vowels in the plural demonstratives.

With regard to the origin of Late Middle English *þese*, it seems to have appeared in the London area slightly later than when *þis*e emerged. According to Heltveit (1953: 82), *þese* goes back to a non-neuter singular form (probably masculine *þes*), and its origin is proved to be almost certain by the geographical distribution of the various plural forms of the compound pronoun in Middle English, demonstrating the strong preference of *þese* found in the London area by the thirteen century. He analyses that the form *þese* with plural inflective –*e* does not appear to have been frequent in Early Middle English[2], being restricted to the London area and the adjacent areas to the north of London. Therefore, considering the fact that it spread rapidly throughout the whole area of the survey by the Late Middle English period, as *eLALME* also shows that 6 out of the top-7 variants are of *e*-type, it must have become widely known during 100 years by the fourteenth century. By the mid-fifteenth century, the *e*-type

[2] Further discussion regarding this point is made in the later section.

became the prevalent form in texts from all Southumbrian dialect areas (Heltveit: 1953: 86), as *eLALME* also illustrates (see Map 14 below).



Map 14: The distribution of *þese* (top-left), *þise* (top-right), *these* (bottom-left) and *thise* (bottom-right).

The only exception found in the Southern area is Kent, which is known as one of the most linguistically conservative areas in the Late Middle English period (Millar 2000:

69). In the Southern areas including Kent, the archaic type *þas* developed as *þos* according to the sound change in the south, which became a characteristic dialectal form in the thirteenth-century texts of the area (Heltvelt 1953: 82). Consequently, in Kent, *þos* was replaced by *þise* during the first half of the thirteenth century. The *i*-type seems to have been the predominant variant in the fourteenth century in this area. Although a wave of the standardisation approached this area in the fifteenth century onwards, *thise* is still the most common token and *i*-type ones are also more dominant than *e*-type at the point of the mid-fifteenth century (see Appendices 5 and 7).

Although the form *þise* (*thise*) should be regarded as the most regular form in Middle English, particularly after *þis* (*this*) became established as the singular form (Heltveil 1953: 91), the status of *þise* and *þese* had become reversed by the Late Middle English period. Heltveit explains this reversal was brought about due to linguistic reasons rather than the overwhelming influence of the London dialect over others. According to Heltveit, when the unstressed final *–e* of *þise* was becoming mute, the *i-type* came to have a disadvantage that the notion of plural was left unexpressed, and thus the paradigm <*þis*: singular – *þese*: plural> was overcharacterised (1953: 91).

2.2.2 A history of simple pronouns

It would seem that the history of the simple demonstrative is more complicated than that of the compound. One complication is that simple demonstratives function as distal and less emphatic deixis (*that*) and definite article (*the*) at the same time. During the breakdown of the grammatical gender/case systems, a functional differentiation was established in the simple demonstrative paradigm, in which *se* (*þe*) took over the

article function while *þat* (and *þa*) shifted in meaning towards ‘pure demonstrative usage’ (Millar 2000: 311). The decisive factor behind this formal and functional split was Old Norse, which already had possessed the tripartite formal apparatus and had had close contact with Old English in the north of England. Although the notional distinction between *the* and *that* was still vague after the completion of the formal separation, the tripartite system of *the*/*that*/*this* was in full use in London English by Chaucer’s time in a manner very close to that found in Present Day English (Millar 2000: 318).

The second problem concerns the etymology of *those* itself, that is, where and when the form emerged. The archaic form *þas* (*þos* in the Southern area) = *these* seems to have become obsolescent in most areas in England by the middle of the thirteenth century (Heltveit 1953: 71-72); especially, it disappeared by the twelfth century in the North, being replaced by *i*-type tokens which originate in the single neuter *þis*, as discussed earlier. Instead, in the North, the new form *þas* = *those* is assumed to have arisen from the archaic *þa* + plural marker *–s* in the thirteenth century. Taking these incidents into account chronologically, Heltveit (1953: 107) concludes that there is no justification for the assumption that *those* is the direct continuation of Old English *þas* = *these* which acquired a new function at a later stage.

It is postulated that for the creation of *þas* = *those* there is a crucial relation between the fact that the *s*-form THOSE first appeared in Northumbrian and the fact that Old English [ā] was retained there while it was rounded to [ō] in Southumbrian (see Map 12 and Map 13), inducing a close connection of singular *þat* and plural *þas* (Heltveit 1953: 109). In addition, it has long been recognised that the inflectional power of *–s*

for plurality was more generally acknowledged in the North than in the South (Heltveit 1953: 109), as Map 10 and Map 11 also suggested above. In the North, therefore, it would seem that <*þat*: singular – *þa*: plural> was felt to belong to a paradigm where the stem *þa* was a shared stem and the singularity was characterised by the formative –*t*[3]. As a result, the marker –*s* for plurality which was already spread across the noun system became an essential inflective to describe the plurality of demonstratives.

Whereas the compound pronoun discarded its archaic plural form (*þas*) in the early stages of the history of the demonstratives, the simple pronoun retained its inherited plural form (*þa*, *þo*) throughout the Middle English period (Heltveit 1953: 111). Previous literature has revealed that although the transformation of the compound pronoun took place in the thirteenth century, southern *þos* = *those* arises in London English 'more than two centuries later' than the northern *þas*. Even Chaucer preferred *tho* to represent distal demonstrative meaning in the plural, in spite of the fact that a homophonic clash between *though* and *tho* due to a drop of the unpronounced –*gh* must have brought about further confusion (Millar 2000: 288; Smith 2006). Heltveit (1953: 113) proposes two possibilities for this time difference: namely, either the *s*-form was adopted into London English from the North taking as long as two centuries, or that it arose independently of the northern *þas* in the London area in the late fifteenth century, being created from the southern inherited *þo* + the plural formative –*s*, and a further plural marker –*e*.

As to the first suggestion, Morsbach (1930) stresses that there should have been a

[3] Millar (2000) casts doubt on this view, however. I will only remark here that, from the fact that Heltveit is a native speaker of Norwegian in which –*t* is still active as a grammatical morpheme, –*t* would give 'an impression' of a marker, if not grammatically established.

continuous influx of elements peculiar to the dialects of the north of London including Northumbrian into London English. Based on this idea, it is natural to explain that *þos* is an importation from the North as well. However, Heltveit (1953) disagree with this external influence on the emergence of *þos* in the London area. As Appendix 2 demonstrates, on the assumption that the *s*-form was an importation from the North, especially from Cheshire where the form *those* seems to have been originally born, it is difficult to account for the absence of this form in the Midlands by the early half of the fifteenth century (Heltveit 1953: 116). Even though it might not be impossible that the *s*-form spread incredibly quickly during the third quarter of the fifteenth century, it is obvious that much more common at that time are the *þo*-forms throughout the whole country, as Appendix 2 and Appendix 4 illustrate. Therefore, it cannot be said that there were plenty of scribes adjacent to the London area to affect Londoners' scribal habits.

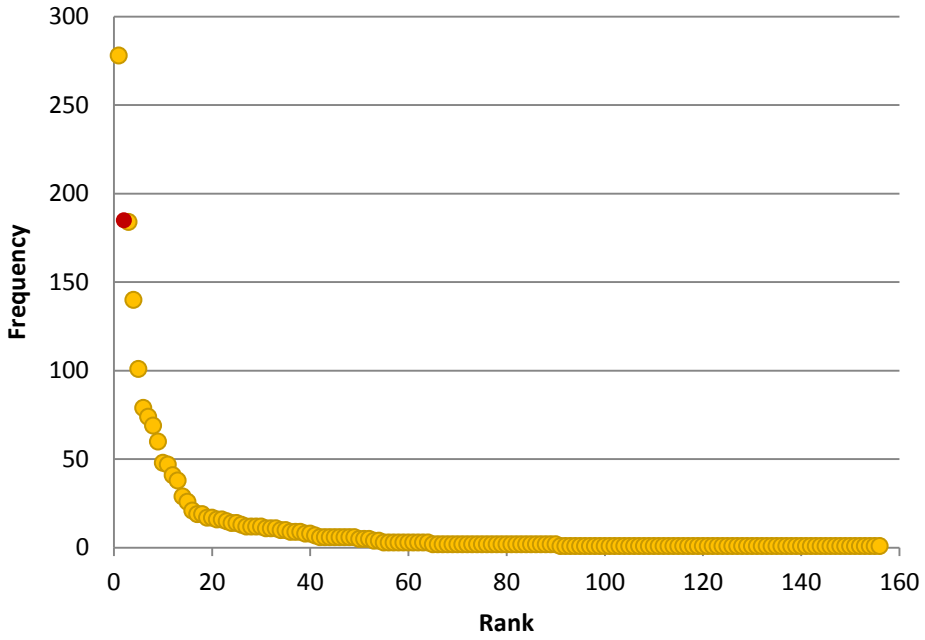
Considering these facts, Heltveit's second suggestion seems more reasonable that the *s*-form emerged in the London area completely independently of the Northern *þas* two centuries later with the same application as the Northern *þas* using the plural inflection *–s.* It is also understandable that the establishment of *these* by the middle of the fifteenth century with the additional plural inflective *–e* and a close association of these two companion pronouns helped the simple pronoun to catch up with the form of the compound pronoun which had already gained ground. In this way, the new *s*-form was rapidly established in about 1475 in the London area (Heltveit 1953: 123). By then, London English was being recognised as a standard language for every scribe to follow, and finally *those* replaced the rest of the existing forms in the early half of the sixteenth century, more than a century later than the establishment of

these.

Yokota (2006) argues that this time gap of the present-day standard written forms is due to the functional difference in THESE and THOSE. According to Yokota, THESE served mainly as a determiner, often entailing with a numeral adjective, while THOSE as a determiner and a pronoun, i.e. an antecedent. It is reasonable, therefore, to suppose that the addition of an adjective plural marker *–e* to the singular form was needed for THESE more urgently than THOSE, which might not have required a plural marker because of a stronger pronominal function and thus persisted *tho*-type variants until quite late.

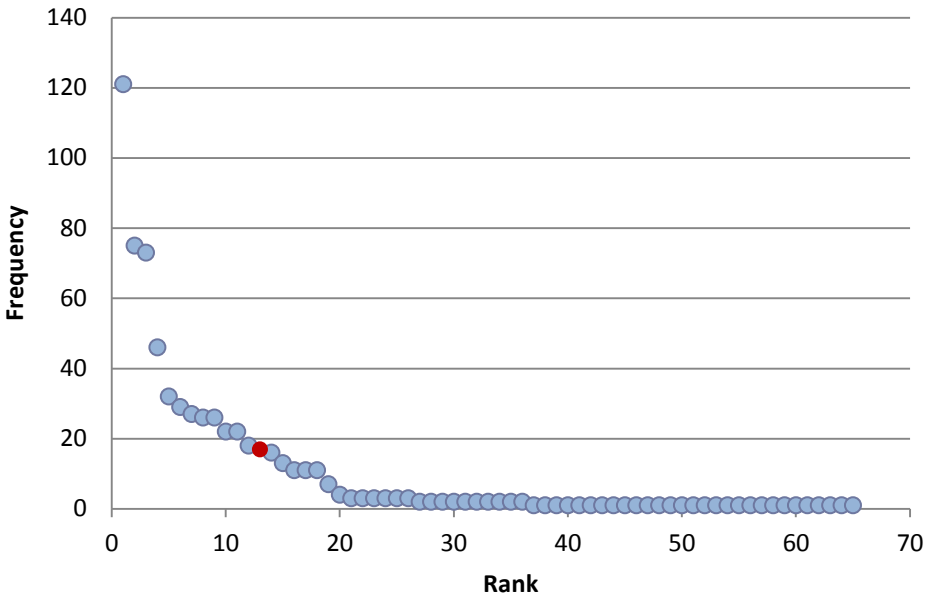
2.3 Survey of the Middle English Demonstratives

In this way, the Middle English demonstratives, which lost their well-structured case- and gender-based paradigms, were to survive in a huge number of formal recorded variants. The lists of all variants and their geographical distributions are available in Appendices 1, 2, 3 and 4. One of the most intriguing results obtained from these data is the frequency distribution, what is called ‘A-curve’ (Kretzschmar 2009), shown in Figure 1 and Figure 2 below. Kretzschmar considers that language has extensive variation in all features at all times, and the frequency distributions of these features form the same asymptotic hyperbolic curve. The frequency distributions of variants of the Middle English demonstratives obtained in this study clearly show the same curve, verifying that variation always demonstrates the same basic pattern of distribution (Kretzschmar 2009: 97).



þese	278	thyes	12	theȝ	4	yiese	2	theyes	1	yieȝ	1
these	**185**	y^{es}	12	þ^{es}	4	yi*re*	2	thier	1	yire	1
þes	184	þeise	12	Thys	3	yisse	2	thiese	1	yises	1
thes	140	þyse	12	theos	3	þ*er*	2	thieȝ	1	yyes	1
þis	101	These	11	there	3	þ*us*	2	thir*e*	1	þ*ere*	1
yese	79	ther	11	thus	3	þer	2	thues	1	þ^{eis}	1
yes	74	thir	11	yeir	3	þesen	2	thyese	1	þ^{e}s	1
þise	69	þeose	10	yeȝ	3	þeys	2	thyr	1	þe^{s}	1
yis	60	þus	10	yhese	3	þeyse	2	tys	1	þece	1
this	48	y*er*	9	yier	3	þeȝ	2	y*ese*	1	þeeȝ	1
yise	47	y^{is}	9	þesse	3	þues	2	y*ere*	1	þere	1
thise	41	þeis	9	þeȝe	3	ȝes	2	$\overset{e}{\text{y}}^{s}$	1	þeus	1
thys	38	theise	8	This	2	Thees	1	$\overset{e}{\text{y}}$se	1	þez	1
þees	29	yees	8	dese	2	Theos	1	$\overset{i}{\text{y}}$se	1	þeze	1
yir	26	y^{s}	7	theese	2	Theys	1	y*eis*	1	þiis	1
thees	21	Thes	6	theose	2	Thise	1	y^{e}es	1	þire	1
thyse	19	y^{e}se	6	theyse	2	Thuse	1	y^{e}ise	1	þisse	1
þys	19	yeis	6	thire	2	Thyse	1	y^{i}es	1	þiȝe	1
theis	17	yeise	6	thuse	2	dyse	1	y^{i}r	1	þoes	1
yies	17	þ^{s}	6	y^{e}is	2	theer	1	y^{i}se	1	þos	1
thies	16	þeese	6	y^{e}s	2	theeȝ	1	yair	1	þs	1
yer	16	þies	6	y^{e}ȝ	2	theiȝ	1	y*eus*	1	þuese	1
þuse	15	þir	6	y^{ies}	2	th*ere*	1	yece	1	þyȝe	1
theys	14	thesse	5	yeese	2	theses	1	yess	1	ȝese	1
þeos	14	yere	5	y*ere*	2	theseȝ	1	yesse	1	ȝeyse	1
þ^{e}se	13	þ^{is}	5	yeys	2	thesȝ	1	yeyse	1	ȝise	1

Figure 1: THESE by frequency (a red dot signifies the PDE equivalent form: *these*)



þo	121	þos	18	they	3	þay	2	thow	1	yoȝ	1
tho	75	**those**	**17**	yai	3	þei	2	thoys	1	þ^{o}se	1
yo	73	yas	16	þas	3	þoos	2	yaas	1	þ^{t}	1
þoo	46	y^{o}	13	þase	3	thar	1	yaise	1	þaa	1
yoo	32	thos	11	thai	2	thas	1	yaisse	1	þai	1
ya	29	yase	11	thay	2	thaye	1	yay	1	þais	1
þose	27	þ^{o}	11	thoes	2	the	1	yayes	1	þaye	1
yose	26	yais	7	y^{os}	2	thoe	1	yho	1	þhose	1
þoȝ	26	þa	4	yaa	2	thois	1	yo^{s}	1	þo^{o}	1
thoo	22	tha	3	þaas	2	thoos	1	yois	1	þoe	1
yos	22	thase	3	þaes	2	thot	1	yoose	1		

Figure 2: THOSE by frequency (a red dot signifies the PDE equivalent form: *those*)

Kretzschmar also points out the existence of parallel systems of ‘normal’ and ‘standard’ (2009: 4); a factor which determines the ‘standard’ form does not always correspond to the more frequent variation found in the A-curve. Looking at these figures, it seems that *these* was already shared with a majority of speakers, and its standardisation can be simply accounted for as a result of the scribal change in the letter <þ> to <th>.

On the other hand, in terms of frequency, *those* is on the borderline between 'normal' and 'different'. It is hard to claim that *those* was perceived as a 'normal' form within the English-speaking community up to the late fifteenth century[4]. From the fact that Chaucer also preferred to use *tho* and the forms without a plural marker *–se* were still predominant at that time, it is unlikely that *those*, or even more frequent *þose*, gained the higher position on its own. It would seem more reasonable to explain that the internal linguistic association between these two demonstratives operated for the rise of *those* over other forms, probably by adding the plural marker *–se* as discussed earlier.

These figures also illustrate the different stages which these two demonstratives had reached by this point. The smooth curve of THESE would indicate that the linguistic change which had started at the end of the Old English period had almost settled down. By contrast, the line of THOSE is uneven, seeming to suggest two things. One is that those variants of intermediate frequency were spoken with uncertainty between members of the community. The other one is that the linguistic change taking place within the system of THOSE was still under way. It seems the competition among variants is encouraging the apparent rise in use of these neither-normal-nor-different variants by speakers, making it possible for a minor variant to survive, as suggested above by Tridgill (1986).

These lists of variants make us aware of the fact that there are a huge number of diverse forms only for one referent but most of them are found only in a few texts and their frequency is extremely low. These 'sporadically-occurring forms' might be termed

[4] Except for in some border areas of the North West Midlands such as Cheshire, Lancashire and West Riding of Yorkshire (see Appendix 2).

'noise' (Smith 2006: 4). Although this 'noise' is an integral part of the history of the demonstratives, a more coherent picture of the forms will be obtained if this 'noise' is set aside. The most commonly used variants are given in Table 3 and Table 4 below.

Form	No. of texts
thes	140
these	185
this	48
thise	41
yes	74
yese	79
yis	60
yise	47
þes	184
þese	278
þis	101
þise	69

Form	No. of texts
þese	278
these	185
þes	184
thes	140
þis	101
yese	79
yes	74
þise	69
yis	60
this	48
yise	47
thise	41

Table 3: 12 most common forms for THESE in Middle English in alphabetical order (left) and in frequency (right).

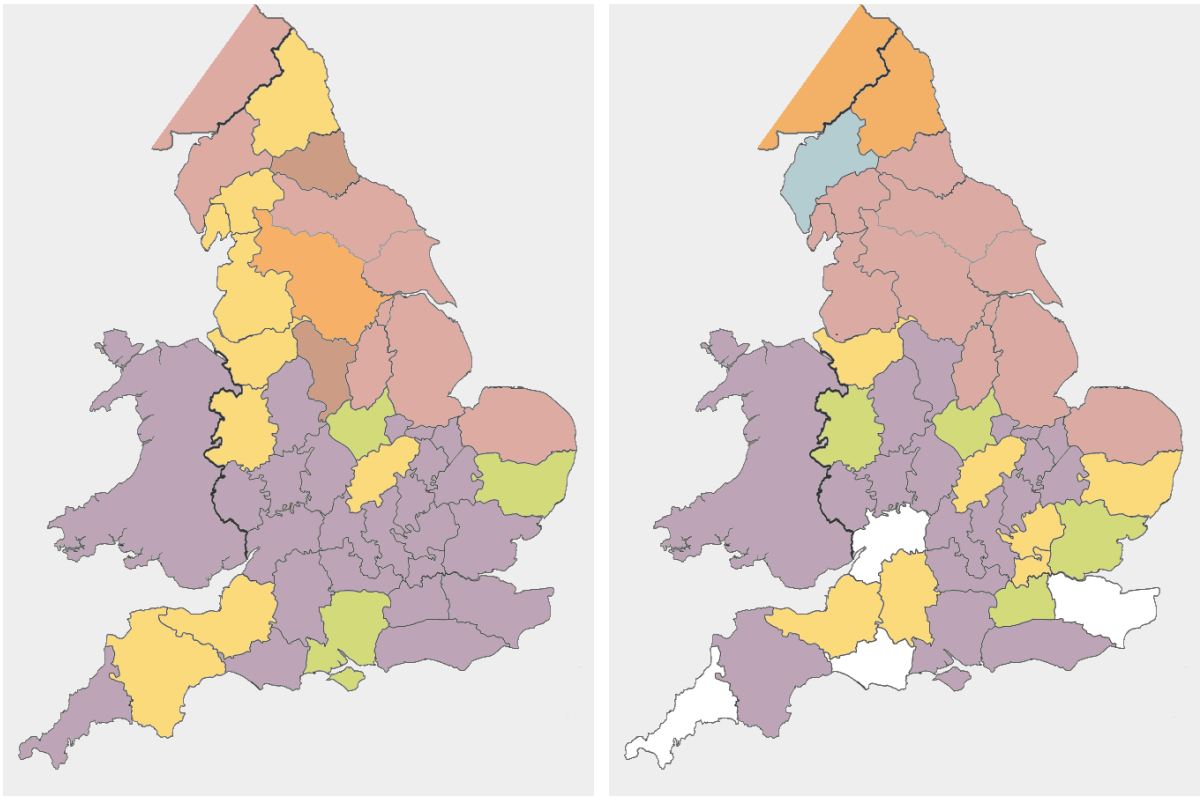
Form	No. of texts
tho	75
thoo	22
ya	29
yo	73
yoo	32
yos	22
yose	26
þo	121
þoo	46
þose	27
þoȝ	26

Form	No. of texts
þo	121
tho	75
yo	73
þoo	46
yoo	32
ya	29
þose	27
yose	26
þoȝ	
thoo	22
yos	

Table 4: 11 most common forms for THOSE in Middle English in alphabetical order (left) and in frequency (right).

Three candidates for the first grapheme for both THESE and THOSE are <th>, <y> and <þ>. As briefly mentioned earlier, <th> represents the revival of Latin usage by French scribes, and <þ> is the remains from Old English. The letter <y> is a symbol confused with the letter <þ> especially in the northern parts of England (Benskin 1982). It is obvious that *þ*-forms are still more frequent by the fifteenth century.

As illustrated in Maps 1-4 and Maps 6-7 in Chapter 1, the distribution of these three types clearly shows the geographical dependence. Whereas, for both of THESE and THOSE, *y*-forms can only be found in the limited areas of northern and some eastern parts of England, *þ*-forms are distributed only in the Southumbrian area. *Th*-forms are found throughout this area on the map. The most common initial letters of each county are illustrated in Map 15, presenting more practical pictures of the situation.



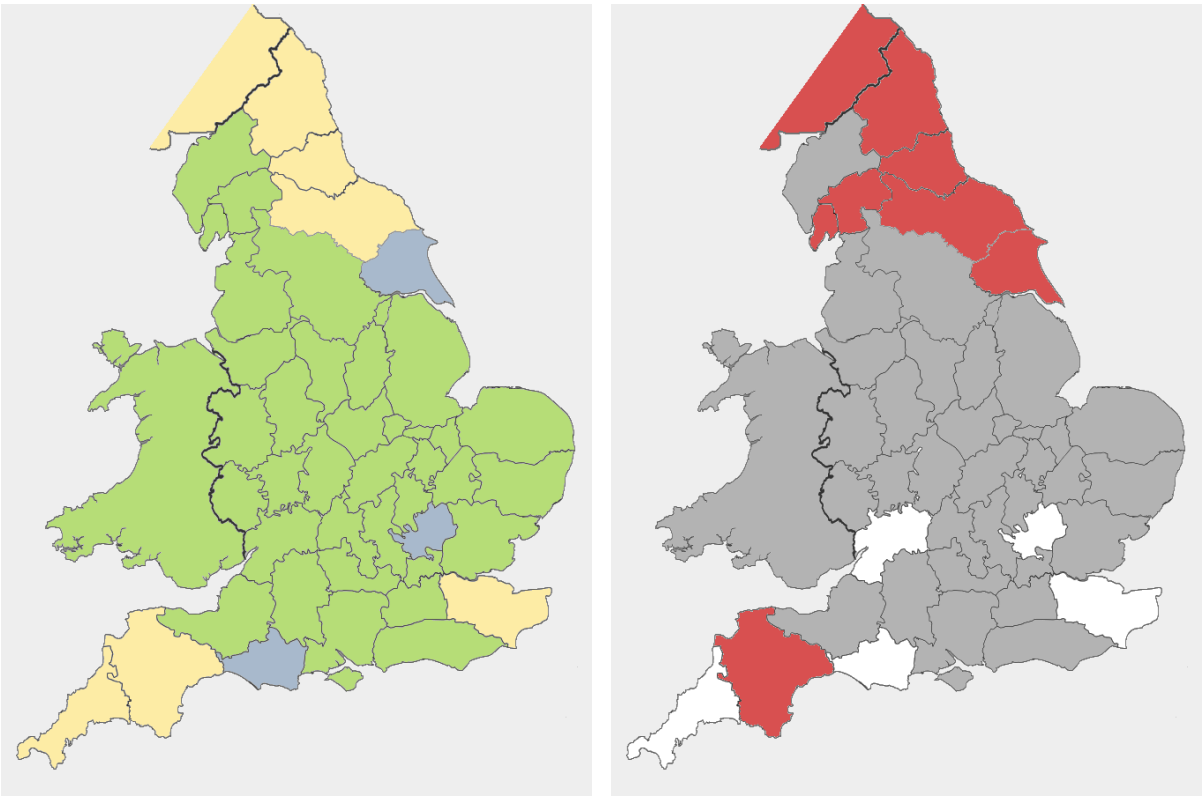
Map 15: The most common initial letters for THESE (left) and THOSE (right).
(1) Yellow: *th*, (2) Pink: *y*, (3) Purple: *þ*, (4) Green: *th*/*þ*, (5) Brown: *y*/*þ* (6) Orange: *th*/*y* (7) Blue: *th*/*y*/*þ*

What is shown by Map 15 above is the degree to which *th*-forms spread into the English-speaking area by the first half of the fifteenth century. The map illustrates the dominance of *y*-forms in the northern area and that of *þ*-forms in the Midlands and southern areas over *th*-forms. However, it seems that Northamptonshire, Cheshire, Lancashire and Somerset were already the established *th*-form areas at that time; with the surrounding areas i.e. Leicestershire, Shropshire, Devon, Suffolk and Essex also showing a large amount of *th*-forms. In particular, Cheshire displays the most marked tendency toward the use of *th*-forms for both THESE and THOSE, and Lancashire also demonstrates very strong evidence of *th*-forms for THESE. Table 5 lists those counties with the number of texts which contain *th*- and *þ*-forms. It is made clear from the map and the table that the development of *th*-forms is still in progress.

	THESE	THOSE
Northumpshire	26 (*th*) / 19 (*þ*)	14 (*th*) / 10 (*þ*)
Leicestershire	17 / 18	13 / 12
Shropshire	11 / 9	5 / 6
Cheshire	22 / 7	11 / 6
Lancastershire	28 / 3	9 / 5
Devon	14 / 10	5 / 7
Somerset	17 / 10	2 / 1
Hampshire	9 / 9	0 / 1
Suffolk	22 / 23	7 / 4
Essex	22 / 26	4 / 3

Table 5: The number of texts which contain *th*- and *þ*-forms.

In addition to the selection of initial letters, the rivalry of the *e*-type THESE and the *i*(*y*)-type THESE, and that of the *a*-type THOSE and the *o*-type THOSE are also an interesting issue to discuss. Here are Map 16 and Table 6, showing the distributions of each type and the frequency of each type, respectively.



Map 16: The most common medium letters for THESE (left) and THOSE (right).
(1) Yellow: *i*(*y*), (2) Green: *e*, (3) Blue: *i*(*y*)/*e*, (4) Red: *a*, (5) Grey: *o*, (6) White: N/A

	THESE	THOSE
Northumberland	8 (*i*) / 6 (*e*)	2 (*a*) / 0 (*o*)
Durham	9 / 4	3 / 0
NRY	16 / 11	12 / 4
ERY	9 / 9	9 / 2
WRY	-	23 / 50
Lincolnshire	-	8 / 39
Lancashire	-	4 / 24
Nottinghamshire	-	1 / 29
Hertfordshire	5 / 5	N/A
Kent	8 / 5	N/A
Dorset	1 / 1	N/A
Devon	12 / 9	6 / 4
Cornwall	1 / 0	N/A

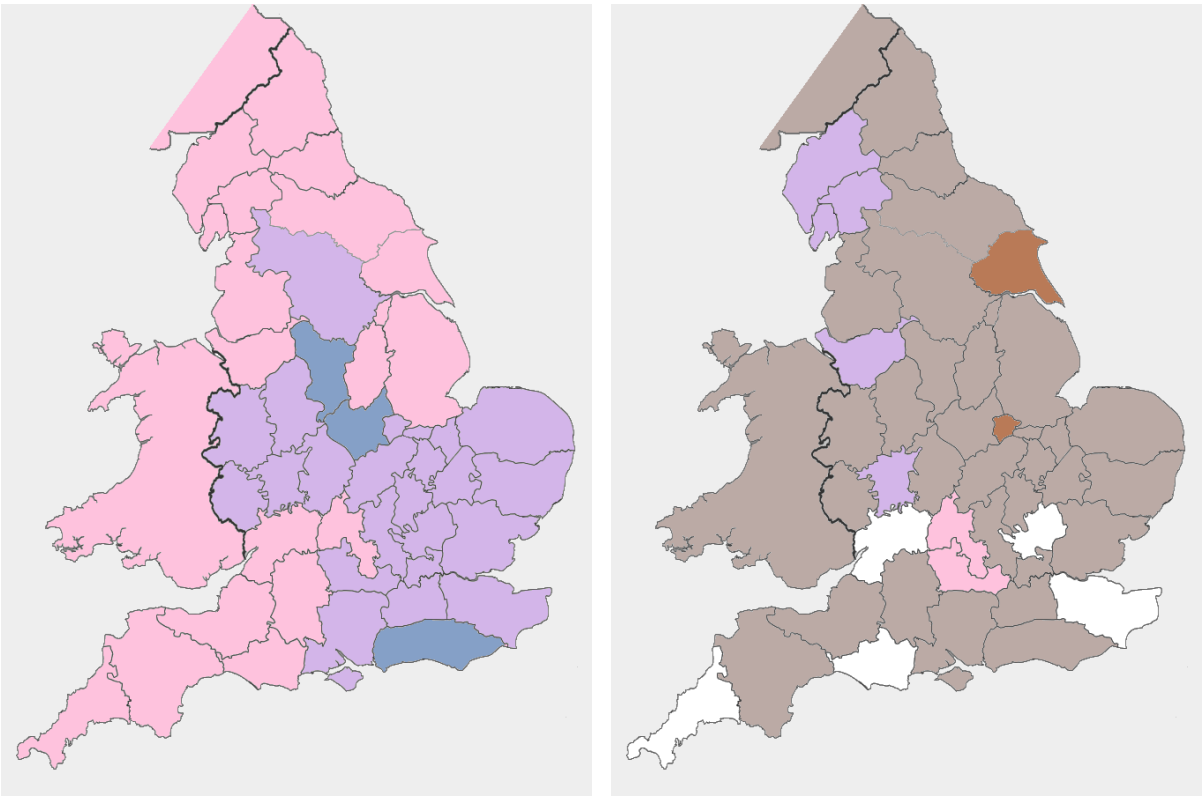
Table 6: The number of texts which contain *i*(*y*)- and *e*-forms for THESE; *a*- and *o*-forms for THOSE.

As discussed earlier, Heltveit (1953) argues that the form *þese* appeared in the thirteenth century but was restricted to the London area and was not frequent in Early Middle English. The map above seems to demonstrate the rapid spread of the form from London, but not up to the peripheral areas. In particular, the *i*-type is prevalent in Northumberland, Durham and North Riding of Yorkshire. In these areas *r*-forms such as *thir*, *yir* and *þir*, and *this* (*yis*) of plurality are still very common (*this* of plurality will be discussed in a later section). As discussed earlier, the map also shows the linguistic conservativeness of Kent despite its proximity to London.

The northern *a*-type and the southern *o*-type for THOSE are equally markedly distinguished on the map. Not only the counties in red on the map where *a*-type is predominant, but their surrounding areas i.e. West Riding of Yorkshire, Lincolnshire, the northern part of Lancashire and Nottinghamshire also show the trace of *a*-form; which suggests the borderline of the sound development of Old English *þā* lies around these regions.

Lastly, a clear pattern for the word ending can also be discerned seen in Map 17 and Table 7. From the map, it is revealed that the plural marker *–e* is recognised only in the South East, Midlands and West Riding of Yorkshire for THESE. As for THOSE, it seems neither *–s* nor –e is yet acknowledged as a plural marker, only sporadically observed in several counties in the West and the North. Moreover, since the number of samples is extremely small in Worcestershire, Northumberland and Westmorland, it is almost impossible to conclude that the *se*-type is characteristic of these areas. The only area where the plural marker *–e* is actively used for THOSE is Cheshire, where the number of texts which contain tokens with *–s*, *–se* and no marker (e.g. *þo*) is 5, 10

and 9, respectively. It is very intriguing that these areas are isolated from one another, as if giving a glimpse of the unpredictability of linguistic change. As illustrated in Map 5, the *r*-form is still very common in the northern areas.



Map 17: The most common final letters for THESE (left) and THOSE (right).
(1) Pink: *-s*, (2) Purple: *-se*, (3) Blue: *-s*/*-se*, (4) Grey: *neither of them*, (5) Brown: *s*/neither of them

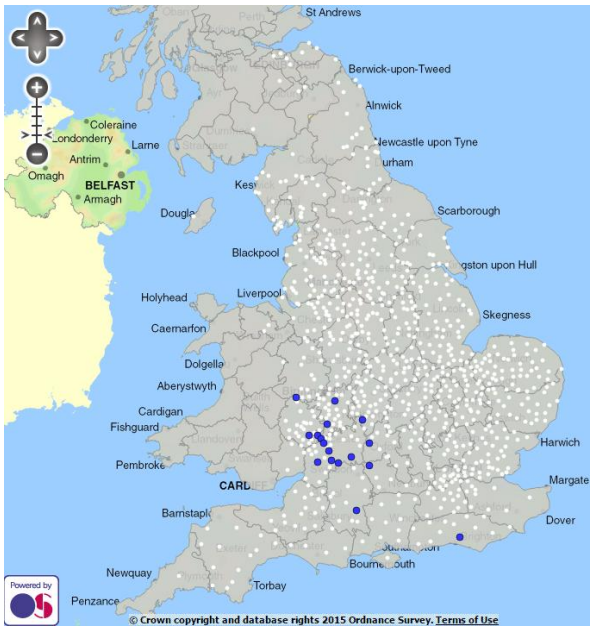
	THOSE
Cumberland	1 (*s*) / 2 (*se*) / 0 (*)
Westmorland	0 / 1 / 0
Cheshire	5 / 10 / 9
Worcestershire	0 / 1 / 0
Oxfordshire	1 / 0 / 0
Berkshire	1 / 0 / 0

Table 7: The number of texts which contain *a*- and *o*- forms (*: no marker).

In conclusion, a history of the demonstratives clearly demonstrates a history of choosing and adaptation of variations: that is, how one form comes to be felt as a more/less suitable one at one stage of the evolution. In other words, it may be taken as a typical example of the language evolution. A triumph of one variant describes a survival of the fittest one in language. In the following sections, several remarks on the development of particular forms of the demonstratives mainly discussed by Heltveit (1953) are to be made with the most up-to-date data gained from *eLALME*, when appropriate.

2.3.1 The plural *þeos*

The plural form *þeos* is investigated by Heltveit (1953). Heltveit considers that *þeos* is an analogical plural form of compound pronoun that emerged in dialects of the West Midlands and the South West (see Map 18 below).



Map 18: Distribution of *þeos*-type THESE (*Theos*, *theos*, *theose*, *þeos*, *þeose*)

As shown in Section 2.3, *þeos* consists of nominative singular feminine *þeo* + deictic marker *s*. As the nominative singular feminine form came to be used for the accusative as well when case distinctions were weakening, the form *þeos* was extended to the plural nominative and accusative forms because of the ancient association between the nominative singular feminine and nominative and accusative plural (Heltveit 1953: 75). The hereditary plural *þas* was then replaced by this new form *þeos* in the West Midlands and the South West areas (Heltveit 1953: 75).

As a result, the strong preference for the plural *þeos* can be noticed to *þes* and *þis* which were far more dominant forms in the Early Middle English period. As Heltveit notes (1953: 76), a single instance of *þise* cannot be observed in South Western texts up to the Early Middle English period, which was already a predominant form at that time. *eLALME* also presents no instance of *þise* in this area in the Late Middle English period. Table 8 provides a list of all the texts which contain *þeos*-type variants (i.e. *Theos*, *theos*, *theose*, *þeos*, *þeose*).

As shown in Table 8, *þeos*-type variants of THESE are found in 23 texts, in 16 (70%) of which *þeos*-type is the most major form. However, the solid preference for *þeos* did not last beyond the Late Middle English period, giving up its status to *þis*/*þise* and *þes*/*þese*. The only area in which *þeos* was still retained to some extent through the Late Middle English period is Gloucestershire; *þeos* appears in 9 out of 40 texts and for the most part, it is used as the most common variant, according to *eLALME*. It would seem, therefore, that *þeos*-type THESE only survived in the extremely limited areas of the West Midlands and the West South, mainly in Gloucestershire, and in these areas the Late Middle English period was a ‘transitional period’ from this unique

usage to the standardised one.

County	LP Number	*þeos*-type	others
Shropshire	704	*þeose, theose*	
	4037	*þeose*	*þis*
Herefordshire	7330	((*theos*))	*þese*, ((*þes*)), ((*thes*))
	7520	*þeose*, ((*þeos*))	*þese*, ((*þes*)), ((*þis*)), ((*þuse*))
	9260	((*þeose*))	*þis*, ((*þes*)), ((*þese*)), ((*þise*))
Worcestershire	7630	*þeos*, *þeose*	
	7640	*þeose*	*þese*, (*þes*), (*þis*)
	7780	*þeos*, *Theos*	*þes*, *þese*
Warwickshire	4684	(*theose*)	*þese*, *theis*, *thes*, (*these*)
	8040	*þeos*, *þeose*	*thes*, (*þese*), (*these*)
Gloucestershire	6970	*þeos*, *þeose*	*þis*
	6980	*þeos*	
	6990	*þeos*, *þeose*	
	7051	(*þeos*)	*þes*, *þues*, (*þeus*)
	7052	*þeose*	*þues*
	7070	*þeos*	*þis*
	7160	*þeos*	(*þes*)
	7180	*þeos*	
	7190	*þeos*, *theos*	(*thes*), (*thus*)
Oxfordshire	6920	*þeos*	
Berkshire	6810	(*þeos*)	*þis*
Wiltshire	5411	((*theos*))	*thes*, (*þis*), (*these*), (*this*), ((*þes*)), ((*þese*)), ((*þ*^{s}))
Sussex	9300	((*þeos*))	*þes*

Table 8: a list of texts which contain *þeos*-type variants (*Theos*, *theos*, *theose*, *þeos*, *þeose*).

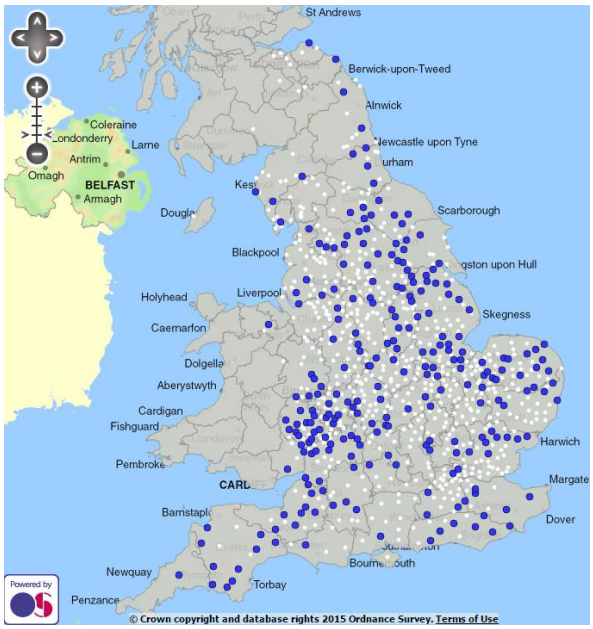
2.3.2 The singular/plural and the substantival/adjectival usage of *this*

Heltveit (1953: 77) also draws much attention to the use of *this* expressing the notion of plurality (e.g. *this three books*). It would appear that this usage emerged when the

compound pronoun lost all distinctions of gender and case in the singular form, and thus among all possibilities the neuter *þis* (*this*) came to be widely used as plural as well as singular by virtue of the weakened distinction of number. He also points out the fact that the compound pronouns in Middle English often occur as an equivalent to the definite article, with no distinction of number, which might have encouraged this usage.

Although Heltveit seems to have realised that instances of *this* of plurality is to be found in most Middle English texts, he would rather reject the notion that *this* for plurality is frequently used. However, as *eLALME* shows the high frequency of these related forms, as seen in Map 19 and Table 9, *this* of plurality became rapidly common by the Late Middle English period, and became a milestone in the history of demonstratives to establish the foundation of the modern form of the compound demonstrative. Carefully examining each item, *this* and *þis* are to be found throughout England whereas ones with initial <y> (*y*^{js} and *yis*) are only found in Northumbria as pointed out by Benskin (1982). It seems that *þys* is only found south of the Wash (see Appendix 1).

There are found some examples of *this* of plurality in *Cursor Mundi*, which are to be further discussed in a later chapter.



Map 19: Distribution of *this*-type THESE (*This*/this, *Thys*/*thys*, *y*^{is}, *yis*, *þis*, *þ*^{is}, *þys*)

This/*this*	50	10
Thys/*thys*	41	13
y^{is}	9	36
yis	60	9
þ^{is}	5	50
þis	101	5
þys	19	17

Table 9: a list of *this*-type THESE, with the number of texts (centre) and their rank (right).

2.3.3 Competition of *thise* and *these* (Chaucer and Caxton's use of *thise*)

It is interesting that, against the background of the fact that the *e*-type was already established as the ordinary form in the fifteenth century, Chaucer seems to have used the *i*-type constantly, and Caxton also still seems to have equally used both *thise* and *these* (Heltveit 1953).

Chaucer's conservative attitude toward the English language has been repeatedly pointed out (e.g. Samuels 1972, 1981; Millar 2000); for example, the preference of the *is*-form to *es*-form for the nominal plural ending (Benson 1992). It seems that Caxton's Kentish upbringing posed a difficult dilemma over making a definite choice between old/new forms (Samuels 1981: 88), in spite of the fact that the *i*-type THESE could not be found in texts issued from the press of other early printers (Heltveit 1953: 90).

The question raised here is the extent to which the influence of Chaucer's and Caxton's persistence on archaic usage had on the later scribal habits and linguistic norms, and, by extension, on the development of the modern forms. It should be safe to say that the prevalence of *e*-type throughout the English-speaking community at the time of their activities rendered their influence to a minimum. In any cases, the history of the demonstratives presents us the fact that the spread of one system is made possible by various complex reasons, from linguistic, functional needs to non-linguistic, social pressure.

2.3.4 The Northumbrian *r*-form

The Northumbrian plural form of the compound pronoun (e.g. *þi**r***, *þe**r***) is one of the biggest problems, or 'riddles' (Heltveit 1953: 92), whose origin has never been satisfactorily solved. According to Heltveit, there are a couple of views about its origin. The first one is that *þir* derives from Old Norse *þeir* (they). That it derives from the Scandinavian simple pronoun *þe* + the adverb *here* is the second, and the third one is that it is an analogical formation of the personal pronoun <*his*: singular – *hir(e)*: plural>. However, as Smith (2006: 5) remarks, the dominant view is that *þir* derives from Old Norse *þeir* because in Old Norse the paradigms for third person personal pronoun and the demonstrative simple pronoun were identical in the plural. In any case, the fact that *r*-form is restricted to the North is in itself would indicate a Scandinavian origin of this form, the inflective –*r* seeming to have raised the notion of plurality, especially for the compound pronoun only, in the area (Heltveit 1953: 95).

With regards to the period when the *r*-form emerged, because of the fact that there is

no trace of the *r*-form in Old Northumbrian, one of the Old English dialects, it is suggested that *r*-form became established after the eleventh century. Although Heltveit claims that it became a 'regular' form in the fourteenth and fifteenth century (1953: 93), the regularity of the form in the period is uncertain. Even though the *r*-form is obviously the distinctive variation found in the North area, as Figure 5 shows earlier, *eLALME* reveals that in no county is the *r*-form dominant during 1350 and 1450, apart from Cumberland (see Appendix 5 and Appendix 7). Instead, the *is*-type (e.g. *yis*, *this*) is more common. Therefore, it would be more plausible to say that the period when *r-form* was most commonly used is up to 1350, probably during the thirteenth century, and the *r*-form of plurality began to be obsolete from the fourteenth century onwards by another new marker of plurality *–s.*

2.3.5 The Plural *þeo*

Heltveit (1953: 102-106) makes a brief remark on the plural form *þeo*. He supposes that the plural *heo* of the personal pronoun (they) and the existence of *seo* (feminitive nominative singular) must have contributed to the establishment of *þeo*. He continues that *þeo* must have begun to replace the traditional form (*þa*, *þo*) in the South West Midlands by the late twelfth century, in the same way as the compound pronoun *þeos* did as discussed earlier. However, as he also admits, the unusual form having a comparatively limited range both geographically and chronologically seems to have failed to find its way to survive beyond the Early Middle English period, contrary to *þeos* which gained a reputation to some extent in the particular area. *eLALME* reveals no evidence of the form existing in the Late Middle English period. In short, *þeo* should be considered primarily as very much scribal, momentary usage.

2.4 A semantic history of English demonstratives

As briefly discussed earlier, the semantic history of the English demonstratives is also not as simple as it seems. What is known is that the main usage of demonstratives has shifted during the development in the history of English. Nevertheless, it must be admitted that much work on the semantic analysis of demonstratives has not been carried out enough to deepen the discussion. The definition cited in *OED* does not even seem to suffice to cover all the usage found in the historical contexts.

The challenge for analysis of the function of demonstratives in literary texts lie in the function of demonstratives itself. Pierce (1955) classifies linguistic signs into three categories, *icon*, *index* and *symbol*, and demonstratives can be included in *index*. A *symbol* is a word of relation between form and meaning which is learned, conventional and arbitrary, whereas the linguistic feature of *index* is mainly factual, existential contiguity between meaning and form; in other words, an *index* indicates relational concepts of time and space. Although there are a great number of resources which offer clues to decode the diachronic evolution of *symbolic* words, it can be more challenging to analyse *indexical* words because such words do not necessarily convey meaning nor represent some characteristics of a given referent.

Similarly, in modern linguistics, demonstratives are often categorised as ‘function words’, contrasted to ‘content words’, which depend on other words for their meaning, usually indicating some kind of relation (Attridge 1995: 27). Function words include preposition, article, demonstrative, conjunction, pronoun and auxiliaries. Most (monosyllabic) function words are unstressed in poetry, but they may be stressed if there is some reason to emphasise them. In addition, pronouns can vary whether or

not they are stressed because they are not purely function words (Attridge 1995: 32). As Novelli also points out (1957: 600), *this* is metrically ambiguous; a relatively weak accent is to be put in most cases, but it sometimes has a strong accent due to various reasons. It is this two-facedness of the function of demonstratives that makes it challenging to detect its real function in question, but there is no doubt that it is essential to synthesise various perspectives to approach to this semantic problem.

The recent study made by Epstein (2010) suggests that the function of the distal demonstrative determiner *se* in *Beowulf* cannot be limited to only a marker of deixis and definiteness, serving a variety of discourse-pragmatic functions. According to Epstein, in *Beowulf*, there are no examples where *se* is used for the stress pattern alone; therefore, every instance of the demonstrative determiner in *Beowulf* can be explained in semantic/pragmatic term. Firstly, there are several ways of the use of *se* to highlight characters. Major characters are more likely to occur with *se* than minor ones; and some less important characters may take *se* if the character is ‘discourse prominent’ and vice versa; and some character which will be important later on may also take *se*. Secondly, in connection with the first function, *se* conveys high topicality; that is, it serves as a subject-changing function in a clause. Thirdly, it is demonstrated that over 60% of noun phrases containing *se* occur within two clauses of the chapter boundary. Although this analysis is confined to the use of a simple demonstrative, especially the nominative masculine element, it gives us an interesting insight into the semantic history of demonstratives.

Novelli points out (1957) that there are two ways Chaucer’s uses the demonstrative adjective *this* in *Canterbury’s Tales*. The first effect is its colloquialism and informality.

He attributes the encouragement of the sense of colloquialism to the direct contact between a story-teller and the audience of Chaucer's time where stories are told 'spontaneously' to the listeners (1957: 246), leading to the repeated use of the demonstrative adjective as colloquial expressions. Comparison of this effect of the demonstrative adjective *this* is also made with sermons to strengthen his idea by noting that none of this colloquialism can be found in them although the speech is commonly given directly from a preacher to a congregation (Novelli 1957: 247).

Another effect of Chaucer's demonstrative adjective *this* presented by Novelli is the emphatic usage. He states that the repeated use of the demonstrative adjective tends to make the definition of any particular character sharper and stronger (1957: 247). Chaucer's use of an epithet following *this* i.e. <this + epithet + noun>, seen in 'this *noble* merchant' and 'this *honurable* knyght', also helps to distinguish a character in a peculiar but effective way (1957: 247). It is also pointed out that Chaucer uses *this* most frequently at a point of transition, when the narrator shifts from one person to another (1957: 246). This function indicates an excellent means of focusing attention to the newly referred character.

It seems that Novelli recognises the emphatic effect of the demonstrative adjective to great extent, but he does not surrender its function as a colloquial device to remind the listeners of someone telling the story because colloquialism calls their attention more. Even if the validity of his argument is considered to some degree, however, the fact should not be forgotten that *Canterbury Tales* is written with a 'syllable-stress verse'. Chaucer is the one who brought syllable-stress verse into the English poetry, whose basic metrical structure is different from 'stress verse' seen in the Old and

Middle English prose before Chaucer. According to Attridge (1995: 63), the main feature of stress meter is that it allows the number of syllables between beats to vary within certain limits, encouraging the rhythm of ordinary spoken language. On the other hand, syllable-stress does not allow the number of syllables between beats to vary, so it is less likely to develop any colloquial rhythm of the language. Therefore, it is improbable that Chaucer's use of *this* carries conversational tone. It is natural to understand that the use of *this* in a narrative is simply to attract attention to the character or the thing which is going to be focused on.

His argument is criticised by Smith (2012), as said earlier, claiming that it does not explain the 'semantic change' of the English demonstrative system; that is, the shift from a more/less emphatic distinction in the sense of this/that in Old English and Middle English, to the present-day proximal/distal distinction. Its emphatic function is well-known and has been widely discussed (e.g. Heltveit 1953; Millar 2000). The sense of emphasis apparently missed by Novelli can be noticed, according to Smith, by looking at the metrical pattern of Chaucer's use of demonstratives. He also demonstrates that the more/less emphatic distinction in Middle English can be traced back to the early Germanic languages, such as Gothic, which shows the same semantic framework as the Old English and Middle English demonstrative system.

In the next chapter, the issues raised in the preceding will be illustrated through the investigation of four versions of the same text, viz. *Cursor Mundi*, with a view to showing how diverse systems were deployed across the range of Middle English dialects.

Chapter 3: Text Analysis of various *Cursor Mundi* texts

Cursor Mundi is a Middle English historical and religious poem of nearly 30,000 lines written around the fourteenth century. According to the preface of Morris's *Early English Text Society* edition (*EETS*, hereafter), the poem was written by an anonymous cleric, who regards himself as the one being given a talent by God to write this poem (1893: xix). As certain morphological features of the Northumbrian English dialect of the period are clearly shown, the poem is thought to have been originally written somewhere in northern England. It is mostly written in eight-syllable couplets. The poem is divided in accordance to the seven ages from Creation until Doomsday, and summarises the history of the world as described in the Christian Bible with other additional legendary material.

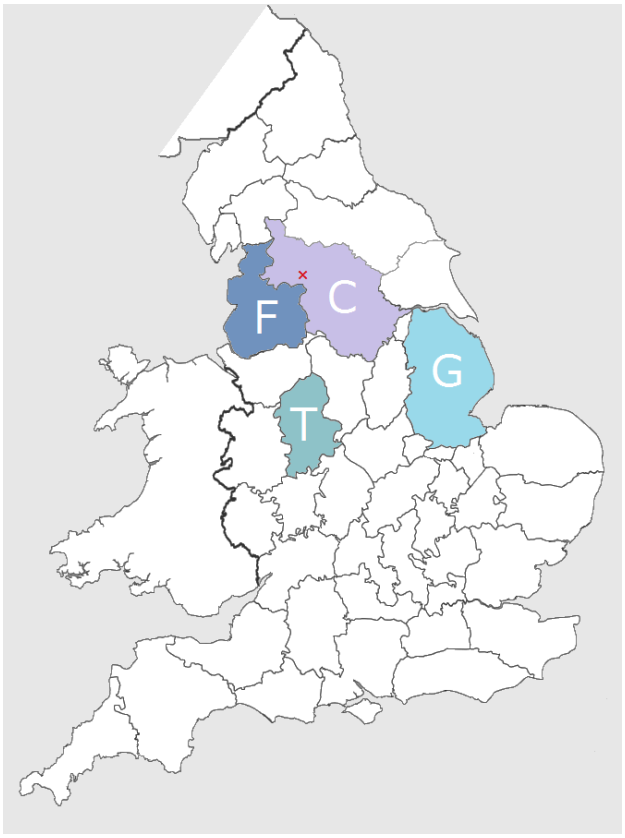
Cursor Mundi was written in the time when the status of the English language was downgraded in Middle English literary culture because of the authority of Latin and French. The author of the *Cursor Mundi* claims that, given such situations, he is writing his work for an English audience that includes those who encounter French verse texts yet simply do not understand them (Thompson 1998: 5). His intention of writing the narrative for this 'educational' purpose was eventually hugely accepted by such French-illiterate people, and as a result the poem continued to be read during a period of about 150 years since the first publication (Thompson 1998: 18-19); that is why a large number of manuscripts have been copied and survived.

There are currently ten extant manuscripts preserving *Cursor Mundi* texts (Morris 1893; Thompson 1998).

1. **E** Edinburgh, Royal College of Physicians
2. **C** Cotton, Vespasian A iii, British Library
3. **G** Göttingen, Theol. 107 r., Göttingen University Library
4. **T** Trinity College Library, Cambridge, MS. R. 3.8
5. **H** London, College of Arms (Herald's College), Arundel Press LVII
6. **F** Fairfax MS. 14, Bodleian Library, Oxford
7. **Add** London, British Library, Additional 31042
8. **B** London, British Library, Additional 36983
9. **L** Laud MS. 416, Bodleian Library
10. **CG** Cotton Galba E.9

The three earliest manuscripts are **E**, **C** and **G**; and all these three were probably written in the first quarter of the fourteenth century; among them, **E** is the oldest text. **E** and **C** are written in a dialect that can be located to the West Riding of Yorkshire, and **G** shows features associated with Lincolnshire usage. None of these manuscripts is a holograph copy. **T**, **H** and **F** date from the fourteenth century. **T** and **H** may have been copied in Staffordshire, and **F** was copied in a dialect characteristic of Lancashire. **Add**, **B** and **L** were produced in the fifteenth century. The usage seen in **Add** can be located to the North Riding of Yorkshire; **B** is derived from the Bedfordshire/Warwickshire area; and **L** is of unknown origin but was probably later owned by a nun at Syon (near London).

The bulk of the *EETS* edition is comprised of four almost complete manuscripts: **C**, **F**, **G** and **T**. Of these four texts, the Cotton MS is the most complete manuscript (*EETS* 1893, xxi). Map 20 below illustrates the distribution of each manuscript. *eLALME* detects the exact place in which the original manuscripts was written (a red cross).



Map 20: The distribution of manuscripts:
Cotton (C), Fairfax (F), Göttingen (G) and Trinity (T).

As the map shows, **C**, **F** and **G** are Northern in contrast to **T**, which is from West Midlands. **C** and **G** exhibit the chief characteristics of the Northumbrian dialect, while **F** also contains many peculiarities of the West Midland dialects as well as the Northern features (*EETS* 1893: xxi). Above all, the morphological difference between **C** and **T** is the most outstanding, showing considerable regional differences in spite of their geographical proximity.

Cursor Mundi is of major importance for the history of the English language in the way it offers a vast amount of linguistic data. The variety of manuscripts also provides us with numerous comparative data for regional differences. In the following sections, linguistic systems found in each manuscript will be discussed.

3.1 Forms in the four *Cursor Mundi* manuscripts

3.1.1 THESE

Variants of THESE seen in the Cotton, Fairfax, Göttingen and Trinity MSS are listed in Tables 10.

Cotton	
þir	161
þire	8
þese	8
þeir	5
thees	1
these	1
þier	1
this	1

Fairfax	
þer/þe*r*	75
thise	21
þes	7
þese	4
þere	2
þi	2
þis	2
þir	1
these	1
þise	1
thyse	1

Göttingen	
[þ]ir/þir	100
þis	7
þise	3
þie	2
þiese	1

Trinity	
þese	82
þir	4
þis	4
þes	2
þus	2
þie	1
þeese	1

Table 10: Forms found in the Cotton, Fairfax, Göttingen and Trinity MSS.

The main form seen in Cotton is *þir*, appearing 162 times (87%). The minor forms are *þire*, *þese*, *þeir*, *these*, *these*, *þier* and *this* although it may be better not to include *þese* since it appears only in lines in Cotton which are written in a different hand (Hand B) and do not correspond to the other three manuscripts; besides, *þese* is a typical southern form not seen in the area where this text was written. A great majority of the variants favoured in this manuscript are *þ*-forms, but it should be noted that *eLALME* takes these forms for *y*-type. It is also interesting that, unlike the more frequent ones, the three rarest forms are *th*-forms even though each of them appears only once; there is the possibility that, considering the fact that the text is derived from the area very close to the border of Lancashire where *th*-forms were dominant, the scribe was influenced by the usage of such areas in Lancashire. The medial letter

favoured in this text is *i*-type, and plurality is mostly expressed by the inflexion *–r*, both demonstrating northern characteristics.

The most common form seen in Fairfax is *þer*, appearing 75 times (64%), which is a very rare form with only two texts found in *eLALME*. The second most frequent form is *thise*, appearing 21 times (18%), but this form may not be of the original scribe of Fairfax but of a different exemplar considering the fact that *th*-form and *i*-type are not his regular forms and appears only partially (between lines 9500-11500 and 16600-18500). The occasional forms in Fairfax are *þes*, *þese*, *þere*, *þi*, *þir* and *þis*. A great majority of the variants favoured in this manuscript are *þe*-forms whereas minor forms tend to be *i*-type. Whether *thise* is the type of the scribe or not, it would appear that he was aware of both *–r* and *–se* expressing plurality. As a whole, the scribe's habits seem to reflect both features of the North and the Midlands, as is suggested in *EETS*. The abundance of variations found in Fairfax is worth noting; nearly double the amount of Göttingen and Trinity.

Similarly to Cotton, the main form used in Göttingen is *þir*, appearing 100 times (85%), and sporadic forms are *þis*, *þise*, *þie* and *þiese*. As for the word ending, his regular form is *–r*, but more minor forms are likely to be *–s*(*e*), a newer plural marker. There is no sign of a different hand or a different exemplar in this manuscript, and it is worth mentioning that the scribe of this manuscript has the fewest variations for THESE of all the four manuscripts, maintaining his scribal habits through the texts. The least common forms *þie* and *þiese* are extremely rare; *eLALME* does not even record either form (though one text using *þies* is found in Lincolnshire).

The dominant form in Trinity is *þese*, appearing 82 times (86%). The minor forms are *þes*, *þir*, *þis*, *þus* and *þie*. Similarly to Göttingen, this manuscript was probably written by a single hand. A great majority of the variants favoured in this manuscript are *þe*-forms, and the less common forms are *þi*-forms. Both *þ*-form and *e*-type are typical of Staffordshire whereas *i*-type is extremely rare in that region; therefore, it can be said that this text is written in a typical dialect of Staffordshire, and probably of the Midlands, as *þese* is the most common variant in the Midlands. It is interesting nevertheless that there are found four examples of the northern *þir*. It is also intriguing that *þus* is used in the manuscript, which is very rare and found only in the South West (see Appendix 1). The scribe who copied this manuscript might have been influenced by such areas e.g. adjacent Worcestershire. Similarly to Göttingen, the number of variations in Trinity is relatively small, nearly half of Fairfax.

Figure 3 shows the frequency distributions of these four manuscripts for THESE, distinguishing more and less common variants. It is revealed that variations owned by even one scribe clearly demonstrate the A-curve even though the sample size is not very large.

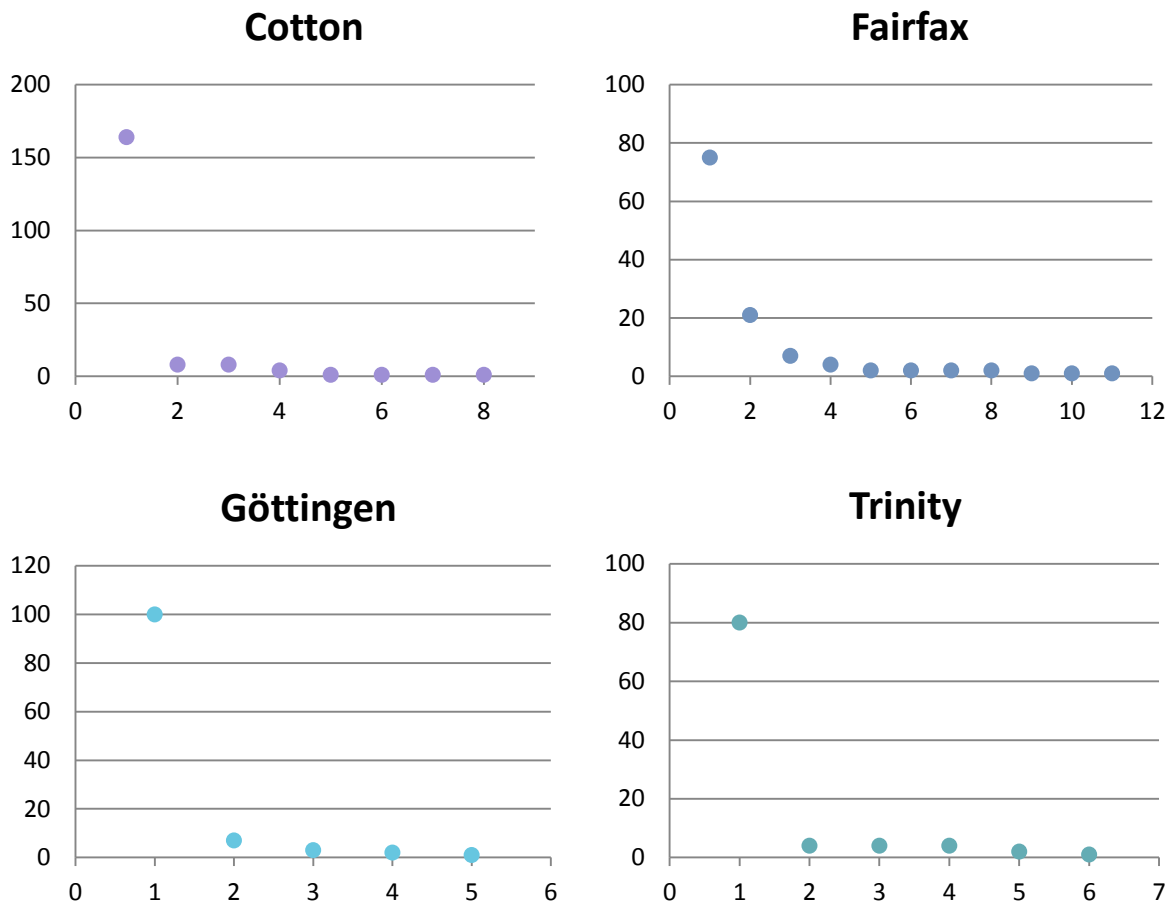


Figure 3: Frequency distribution of variants in the Cotton, Fairfax, Göttingen and Trinity MSS

3.1.2 THOSE

Variants of THOSE seen in the Cotton, Fairfax, Göttingen and Trinity MSS are listed in Table 11.

Cotton	
þaa	237
þas	27
þo	18
þaas	7
þoo	5
þais	1
þos	1

Fairfax	
þa	111
tho	35
þas	5
þo	3
þase	2
ȝa	1
þ^{o}	1
þaa	1
þos	1
thow	1

Göttingen	
[þ]a/þa	89
þaa	45
þas	23
þo	1
þos	1
þou	1

Trinity	
þo	112
þa	8
þar	1
þoh	1
þoo	1

Table 11: Forms appearing in the Cotton, Fairfax, Göttingen and Trinity MSS.

The main form seen in Cotton is *þaa*, appearing 237 times (80%). It is an extremely rare form elsewhere, with only one instance in Yorkshire (presumably NRY) according to *eLALME*. *þas* and *þo* are occasionally used (11% and 7.6%, respectively), and the minor forms are *þaas*, *þoo*, *þais* and *þos*. The form *þo* might be of a different exemplar as *o*-type forms are not the original scribe's and it appears successively between 200 lines from lines 17800 to 18000. All the variants found in this manuscript take *þ*-forms, but, similarly to the cases for THESE, *eLALME* regards these forms as *y*-type. The medial letter favoured in this text is *a*-type, showing the northern characteristic. By far most variants do not use an extra inflexion to express plurality, but more minor ones tend to take *–s* as a means of flagging plurality.

The most common form seen in Fairfax is *þa*, appearing 111 times (69%). *Tho* is fairly frequent, appearing 35 times (22%), but similarly to THESE, this form may not be of

the original scribe of Fairfax but of a different exemplar given the fact that *th*-form is not his usual form and appears only partially (between lines 10700-12000 and 16400 and 18200). A great majority of the variants favoured in this manuscript are *þ*-forms, but again these might be *y*-forms according to *eLALME*. Rare forms in Fairfax are *þas*, *þo*, *þase*, *ȝa*, *þ*^{o}, *þaa*, *þos* and *thow*. Above all, it is very interesting that there are found two exceedingly rare forms which *eLALME* does not even cover: i.e. *ȝa* and *thow*. *þ*^{o} is also rare in view of the fact that it is mainly observed in the East Midlands though Fairfax is thought to be written in Lancashire. The scribe's regular type for the medial letter is *a*-type, and for the most part a marker for plurality is not used. The number of variations found in Fairfax is the largest of all four manuscripts similarly to THESE; twice as much as Trinity.

Similarly to Fairfax, the main form in Göttingen is *þa*, appearing 89 times (56%). The second most frequent form *þaa* is relatively frequent, appearing 45 times (28%), and the third most frequent *þas* appears 22 times (14%), which is also not very rare if not recurring. Sporadic forms are [*þ*]*as*, *þo*, *þos* and *þou*. All the forms seen in Göttingen are *þ*-forms, and a great majority of the variants are *a*-type. Göttingen shows the clear tendency to take *a*-type for the more common tokens and *o*-type for the less common tokens. Most of the variants do not use a plural marker to express plurality similarly to Cotton and Fairfax.

The dominant form in Trinity is *þo*, appearing 109 times (95%), which is exceptionally high frequency. The minor forms are *þa*, *þar*, *þoh* and *þoo*. All the variations found in this text are *þ*-forms, and a great majority of variants are *o*-type, showing the southern feature which derives from the northern *a*-type *þa*, the second most frequent form, by

the sound change. There is no sign of a different hand or a different exemplar in this manuscript, and this manuscript has the fewest variations for THOSE of all the four manuscripts. The least common forms *þar* and *þoh* are quite rare; *eLALME* does not even record either forms (*thar* is found in Selkirkshire, in Scotland, but there hardly seems to share a common origin).

Figure 4 shows the frequency distributions of these four manuscripts for THOSE, distinguishing more and less common variants. Similarly to THESE, it demonstrates the clear A-curve. As shown in the figure, the high frequency of *þaa* in Göttingen might indicate that it was also the scribe's regular form considering the fact that its origin may be very close to *þa*.

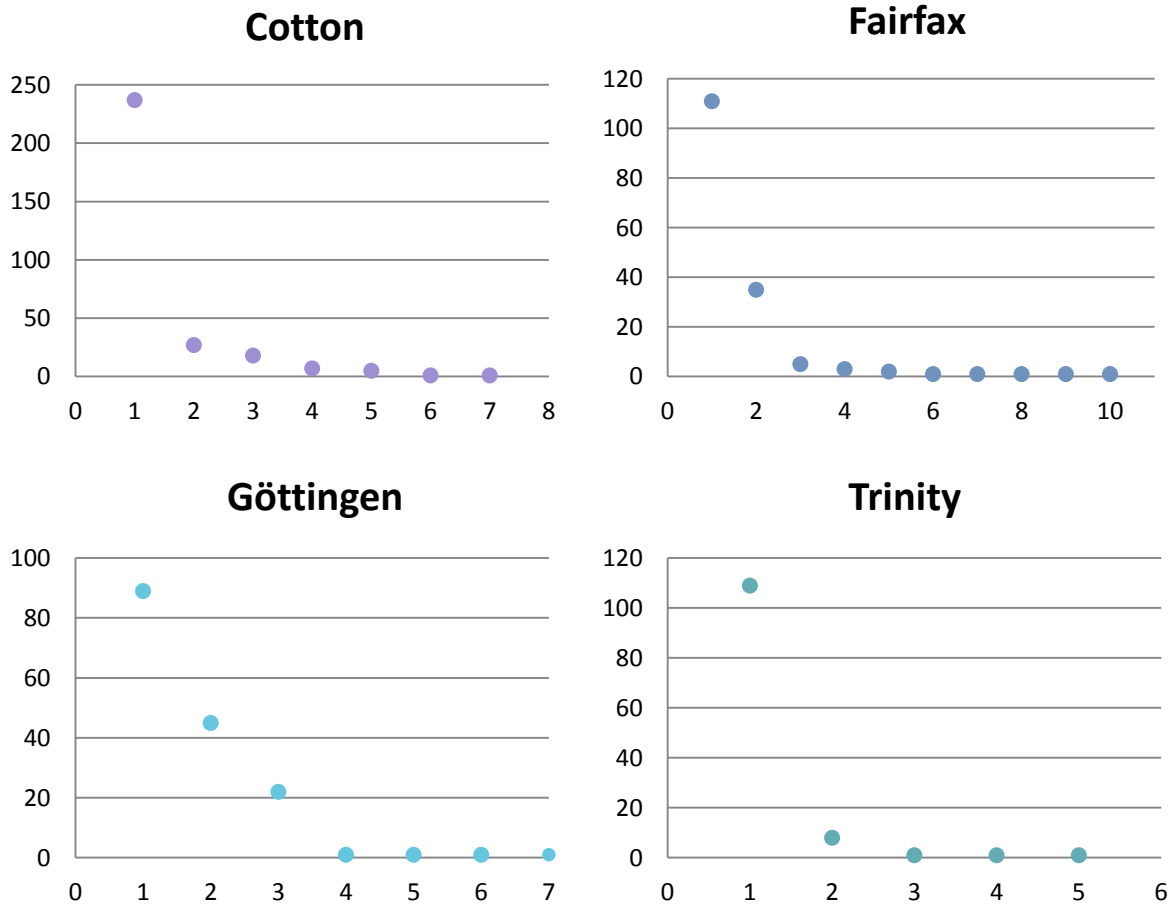


Figure 4: Frequency distribution of variants in the Cotton, Fairfax, Göttingen and Trinity MSS

3.2 Functions

3.2.1 THESE

Functions of THESE seen in the Cotton, Fairfax, Göttingen and Trinity MSS are listed in Table 12.

Cotton		
	adj	pron
þir	122 (76)	39 (24)
þire	3 (38)	5 (63)
þese	7 (88)	1 (13)
þeir	4 (80)	1 (20)
thees	0 (0)	1 (100)
these	1 (100)	0 (0)
þier	1 (100)	0 (0)
this	1 (100)	0 (0)
TOTAL	139 (75)	47 (25)

Fairfax		
	adj	pron
þer/þer	60 (81)	14 (19)
thise	21 (88)	3 (13)
þes	5 (71)	2 (29)
þese	2 (50)	2 (50)
þere	1 (50)	1 (50)
þi	1 (50)	1 (50)
þis	2 (100)	0 (0)
þir	1 (100)	0 (0)
these	0 (0)	1 (100)
þise	1 (100)	0 (0)
thyse	0 (0)	1 (100)
TOTAL	94 (79)	25 (21)

Göttingen		
	adj	pron
[þ]ir/þir	74 (74)	26 (26)
þis	4 (57)	3 (43)
þise	1 (33)	2 (67)
þie	2 (100)	0 (0)
þiese	0 (0)	1 (100)
TOTAL	81 (72)	32 (28)

Trinity		
	adj	pron
þese	64 (80)	16 (20)
þir	4 (100)	0 (0)
þis	3 (75)	1 (25)
þes	1 (67)	1 (33)
þus	1 (50)	1 (50)
þie	0 (0)	1 (100)
TOTAL	75 (79)	20 (21)

Table 12: Functions of THESE found in the Cotton, Fairfax, Göttingen and Trinity MSS. (Numbers in brackets signify percentage)

On the whole, variants for THESE appear much more frequently as demonstrative adjectives than demonstrative pronouns; approximately 70-80% is used adjectivally. It

would appear that the minor variants are slightly more likely to be used as pronouns than the more common ones (approx. 55-75% as adjectives), but whether the form is more common or less is not a decisive factor on their functions; they are dependent more on each token. For example, *þire* in the Cotton MS is used more as a pronoun in spite of its etymological closeness to *þir*, which is used as an adjective for the most part in the four manuscripts. This form is found only in the last 6000 lines (see Appendix 9), and since there cannot be found any formal change similar to this, i.e. from *þir* to *þire*, made by other three scribes, it might be suggested that a scribal crossover was taking place in the Cotton MS from the Northumbrian archaic form *–r* to the word formation using the new plural suffix *–e.*

Cotton		
23097	þan sal **þire** felauscip tua	(adj)
23124	**þire** to deme sal be na nede,	(pron)
23460	All **þire** in þe sal be plente,	(pron)
23475	**þire** ar þe ilk blesced-hedes,	(pron)
23639	Wit alkin thing sal **þire** acorde,	(pron)
28505	Desird o **þire** wymmen scen,	(adj)
28624	þat **þire** maners of beting thrin	(adj)
29006	And **þire** we ask here ilk a day	(pron)

The form *þese* and *these* in the Cotton MS appear only in lines which are written in a different hand (*Hand B*) and do not correspond to the other three MSS. According to the notes in *eLALME*, these additional lines by *Hand B* are written in the fifteenth century. It can be suggested, from the fact that the Cotton MS was written in the first quarter of the fourteenth century and this form is not characteristic of WRY *r*-form of the period, that the scribe who inserted the additional lines is from an area in which *þese* was a common variant, i.e. south of the Humber. Of course it is not improbable

that this part was added by a scribe from the nearby area in WRY even though these two forms were rare in this area. In any case, there is not sufficient information so as to analyse the origin of this token in this manuscript any further. As for the usage, this form almost always appears as a demonstrative adjective.

Cotton		
16762-109	os **þese** clerkez witen,	(adj)
17288-69	**þese** thre thinges a-bod our lord,	(adj)
17288-97	**þese** thre, makand ter mone,	(adj)
17288-124	**þese** thre maries come tiderward,	(adj)
17288-140	When **þese** wymmen come	(adj)
17288-290	With **þese** thre maries os ta went,	(adj)
17288-304	**þese** wymmen told amang hom all	(adj)
17288-460	"**þese** are te wordez, I-wis,	(pron)

It can be speculated that *these* is a second variant of Hand B.

Cotton		
16762-65	***These*** ilk wordez said he.	(adj)

In the Cotton MS, *þeir* is used in two ways: that is, as a third-person personal genitive plural pronoun (*their* in PDE) and as a compound plural demonstrative pronoun (*these* in PDE). There are 5 instances out of 22, shown below, functioning as the latter. Middle English *þeir* is thought to be derived from Old Norse *þeir*, which is originally a third-person nominative plural personal pronoun (*they* in PDE). As mentioned earlier, in Old Norse there was a semantic overlap in the usage of the personal pronoun and the demonstrative pronoun in the plural, and therefore *þeir* is also used as the plural demonstrative pronoun in Old Norse. It became *they* as it is in Present-Day English and retains its form as the third-person genitive plural personal pronoun *their*.

Cotton

1536	þai put tam in **þeir** pilers tuin;	(adj)
4083	Bi **þeir** storis mai man wel se	(adj)
4085	**þeir** breþer, þat i said of are,	(adj)
5831	If þai tru noter o **þeir** tua,	(adj)
27146	**þeir** ar þe pointes þat scriftes lattes,	(pron)

In the first 4 instances, *þeir* is used adjectivally, that is, as a demonstrative determiner. For example, *þeir pilers tuin* in line 1536 indicates 'two pillars' which is mentioned in line 1533; therefore, *þeir* here functions as a demonstrative determiner which is used to refer to a thing which has already been mentioned. *beir broþer* in line 4085 suggests Joseph's brothers appearing in his dream. There is no doubt that this *þeir* is used as a demonstrative pronoun because the thing referred to in this context is Joseph's (= *his*, not *their*) brothers. It is also unlikely that *þeir* of *þeir tua* in line 5831 signifies *their* because a genitive personal pronoun is not usually followed by a single number word (e.g. *their two*). The usage in line 27147 as a demonstrative pronoun is very rare, appearing only once, but it does refer to the contents mentioned just before the line.

The form *þier* appears once. It can be thought to be metathesis of either *þeir* or *þire*. It is interesting that this form follows *þeir* in line 5831, suggesting a formal influence on the next token (see Appendix 9).

Cotton

5938	Wald do **þier** frosses a-wai fra me.	(adj)

There is one instance of *this* being used as a plural demonstrative pronoun as follows.

Cotton

12624 Soght þe abute **<u>this</u>** thre dais (adj)

As briefly remarked earlier, Heltveit (1953: 77) states that in Middle English texts *this* is used to indicate ‘time’ which may logically, ‘psychologically’ (Millar 2000: 269) be conceived of as singular (e.g. *this three weeks*), although such concept should be grammatically plural from the point of view of Standard Modern English, and even in Middle English. The usage shown here is the example of this meaning. There are more examples of this usage in other manuscripts which expresses this ‘collective’ concept (Millar 2000: 270),

Göttingen/Trinity

7323 Amo*n*g **<u>þis folk</u>** shal þ*o*u fynde on

Göttingen

10399 **<u>þis hundrid</u>** þat war boune

Trinity

20178 Þ*a*t bringest me **<u>þis tiþing</u>** here

as well as the usage with normal plural nouns such as:

Göttingen

363 þe world, and tyme, **<u>þis thinges thre</u>**

18903 þu scheu vs queder of **<u>þis tua sere</u>**,

3.2.2 THOSE

Functions of THOSE seen in the Cotton, Fairfax, Göttingen and Trinity MSS are listed in Table 13.

Cotton		
	adj	pron
þaa	172 (73)	65 (27)
þas	24 (89)	3 (11)
þo	12 (67)	6 (33)
þaas	7 (100)	0 (0)
þoo	3 (60)	2 (40)
þais	1 (100)	0 (0)
þos	0 (0)	1 (100)
TOTAL	220 (74)	76 (26)

Fairfax		
	adj	pron
þa	77 (70)	33 (30)
tho	30 (86)	5 (14)
þas	5 (100)	0 (0)
þo	2 (67)	1 (33)
þase	2 (100)	0 (0)
ȝa	0 (0)	1 (100)
þ^{o}	1 (100)	0 (0)
þaa	1 (100)	0 (0)
þos	1 (100)	0 (0)
thow	1 (100)	0 (0)
TOTAL	121 (75)	40 (25)

Göttingen		
	adj	pron
[þ]a/þa	76 (85)	13 (15)
þaa	32 (71)	13 (29)
[þ]as /þas	20 (87)	3 (13)
þo	0 (0)	1 (100)
þos	1 (100)	0 (0)
þou	1 (100)	0 (0)
TOTAL	130 (81)	30 (19)

Trinity		
	adj	pron
þo	92 (82)	20 (18)
þa	4 (50)	4 (50)
þar	1 (100)	0 (0)
þoh	1 (100)	0 (0)
þoo	1 (100)	0 (0)
TOTAL	99 (80)	24 (20)

Table 13: Functions of THOSE found in the Cotton, Fairfax, Göttingen and Trinity MSS (Numbers in brackets signify percentage).

Similarly to THESE, the items for THOSE generally appear more frequently as demonstrative adjectives than demonstrative pronouns; approximately 75-80% is used adjectivally. However, their usage seems to depend much more on each token than THESE. For example, the usage of *þas* in the Cotton MS is striking; it appears 24

times out of 27 as an adjective. There seems to be four usages as follows; namely, (1) *þas*: as a simple plural demonstrative adjective; (2) *þas*: as a simple plural demonstrative pronoun; (3) *þas oþer*: functioning as a demonstrative adjective; and (4) *þas oþer*: functioning as a demonstrative pronoun.

There is one instance of type (1).

Cotton
20965 Ful mani war **þas** ilk torfere,

In this case, there is a possibility that *þas* is used as a singular demonstrative adjective as *ilk torfere* means *each difficulty*. The equivalent tokens in Fairfax and Göttingen are *þat* and *tat*, both expressing singularity. There is no equivalent word for *þas* in Trinity.

Fairfax
20965 ful mani hit was þat ilk torfere
Göttingen
20965 Ful mani it was tat ilk torfer,

The second type can be seen in the following two lines as follows:

Cotton
10007 **þas** er four vertus principals,
10037 **þas** er þe seuen virtus to tell,

Both are used as demonstrative pronouns, as subjects and appear line-initially. The similarity in their grammatical structure and their back-to-back occurrence might suggest the scribe's whimsical behaviour. Interestingly, it is replaced with singular pronouns in other three manuscripts.

Fairfax
10007 That are iiij vertuis princypals
10037 That arn vij vertuis to telle,
Göttingen
10007 þat er four v*er*tus principalys,
10037 þat er þe seuen v*er*tus to tell,
Trinity
10007 þat are foure v*er*tus p*ri*ncipales,
10037 þat are þe seuen v*ir*tues to telle

Type (3) is found eight times:

Cotton
491 **þas oþer** gastes þat fell him wiht
4567 **þas oþer** seuen yede i to see
4597 **þas oþer** seuen nede nett
6059 Mare þan all **þas oþer** smert
7134 Til **þas oþer** sco it vndidd,---
12351 **þas oþer** leons þat war ald,
17491 þe prestes and **þas oþer** ald
21292 **þas oþer** thre wrat in gru,

Various tokens are used differently in context in the other three manuscripts, but mostly *þas* in Cotton is substituted for *þa* in Fairfax, *þas* in Göttingen and *þese* in Trinity, as shown below.

Fairfax
6059 mare þen alle þa oþ*er* smert
12351 þa oþ*er* leonis þat ware alde,
Göttingen
6059 Mare þan all þas oþ*er* smert,
12351 þas oþir leonis þat war ald,
Trinity
6059 More þen alle þese oþ*er*e smert
12351 þese oþ*er*e leou*n*s þat were olde,

The fourth type is seen most frequently, appearing 16 times out of 27 total instances.

Cotton

2295 þis for-bisening he hild **þas oþer**
2376 Mi bliscing sal **þas oþer** misse.
4580 Bot o **þas oþer** me thoght vncuth,
4948 Stilli menand til **þas oþer**,---
5902 **þas oþer** it wired al bidene;
6074 **þas oþer** all þan sal þam find;
10000 Of al **þas oþer** es nan sa god,---
10616 **þas oþer** þat in þe temple was.
11258 **þas oþer** lighted dun thic-fald,
14306 For he wepe sarer þan **þas oþer**
16709 **þas oþer** said, "if he be crist,
19486 þai stod, bot all **þas oþer** fledd:
25059 **þas oþer** þat his lagh forsok,
26431 þat o **þas oþer** for-giuen es nan,
26492 For þis an of **þas oþer** all.
26902 Or for **þas oþer** þe quilk he bette

Roughly speaking, *þas oþer* is equally used as subject and object. If the function of this type is to be considered solely as a pronoun, *þas* is to be used much more frequently as a pronoun.

Tho in Fairfax specialises in the function as an adjective. As discussed earlier, this form is probably not of the original scribe considering the difference in northern and southern features and their limited appearance in the text. The form is nearly always used as a demonstrative adjective with only 5 out of 35 instances as a pronoun shown in the following.

Fairfax

10741 And **tho** that theder come wold
17563 Therfor seid Nicodeme **tho**
17913 **Tho** I haue that sakeles of pligħt,
17967 And to alle **tho** that bene baptyst
17975 Mornyng among hem **tho** was gon

In line 17563 *tho* is used as an object of the verb *seid*. In lines 10741, 17913 and 17959 it is used as an antecedent (*those who* in PDE). In line 17967 it functions as a subject of the line. The rest of the instances show adjectival functions as follows. *Tho* as a demonstrative adjective appears roughly double the frequency as an object.

Fairfax

11264 That come of **tho** angils brigħt,
11374 And offird to hym **tho** kyng*es* hend
11510 Byheld **tho** yeft*es* seith som boke
11520 **Tho** kyngg*es* iij^{e} wer*e* brogħt in bed
11537 **Tho** kyng*es* yedyn a-nothir wey
11542 yf they migħt **tho** kyngg*es* mete
11613 And stode vpon **tho** best*es* grym
11939 Amon **tho** childryn oon þ*er* was
16429 To **tho** wilfułł wode he toke
16636 **tho** hound*es* alle of helle
16912 **tho** caytifs ar but shent
17375 To **tho** women that I*es*u sougħt
17391 what wer **tho** wy*m*men that hym sougħt
17475 Wo was hem **tho** wrecchis wyk
17501 To councele **tho** iij^{e} men they led
17524 Seid so to **tho** knyght*es* alle
17565 ye seke **tho** fellis alle to-gethir
17595 Thus wore **tho** iewis alle mysled
17639 he kyste **tho** messagers alle
17742 Afft*er* **tho** ilke prest*es* iij^{e}
17752 And did **tho** iij^{e} men in sondre
17881 **Tho** folk in dedly derkenes stad
17897 **Tho** sennt*es* alle ther þan stad

17932 To **tho** yat*es* of paradice
17960 And many othir by **tho** weyse
17973 **Tho** patryark*es* that th*is* herd
17974 W*ith* moche ioy **tho** they ferd
17977 Whe[n] satan sie **tho** seint*es* dere
18155 **Tho** wofułł were so dedly dym
18197 Thus seid alle **tho** legions there

Based on the presumption that a simple demonstrative is likely to appear both as demonstrative adjective and demonstrative pronoun, this highly specialised function of *tho* is distinctive. This usage might be relevant to that of *thise* in Fairfax, which also shows higher frequency as an adjective (88%) since it is assumed that both tokens are of a different exemplar.

Although it can be considered that *þa* and *þaa* are etymologically close, developed from the Old English *þā*, it would appear they function differently in Göttingen; *þa* is used more frequently as an adjective and *þaa* more as a pronoun. There is apparently a formal correspondence of the usage of *þas* in Göttingen to that of *þas* in Cotton, and the usage of *þa* and *þaa* in Göttingen to that of *þa* in Cotton (see Appendix 10). It is probable, considering these facts, that *þa* and *þaa* are the main variants of the scribe of Göttingen, but there are not sufficient clues to distinguish these two tokens.

Finally, the functional difference of *a*-type variants and *o*-type variants is very suggestive. It seems that *þo* and *þoo* are used more frequently as pronouns (37%) when they are used as minor forms, i.e. in Cotton, Fairfax and Göttingen, compared to the average usage as a common form in Trinity (18%). On the contrary, *a*-type minor variants are much more frequently used as adjectives (91%); which might suggest that the scribes were conscious of the plural inflexion even though it is not their

regular usage, and deliberately employed them differently.

Chapter 4: Conclusion

The key to understanding Middle English is an understanding of linguistic variation (Samuel & Smith 1988: 2). This thesis has attempted to gain a better insight into two of these enigmatic but thought-provoking linguistic variations, namely, demonstratives *these* and *those*. The main focus is on the geographical distribution and use of various forms in texts in the critical period during 1350-1450, which has not yet been explored enough.

The most mysterious issue on the history of the English demonstratives is that these two words have followed separate paths over the course of their development. The Old English compound plural demonstrative *þās* came to be replaced with *þise* and *þese* which originate in the singular nominative neuter and masculine forms + a plural marker *–e*. The latter form *þese*, which appeared in the London area in the thirteenth century, gradually gained ground and spread across the English-speaking community, mainly the Midlands, and as a result *these* was established as a standard form by the fifteenth century. As for *those*, the root of the present standard form stems from the southern form of the Old English simple plural demonstrative *þo*, developed from the northern *þā*. Contrary to the case of *these*, the archaic form *þo* was still in use in the fifteenth century without a plural marker, but it developed hastily to catch up with its established counterpart with a plural marker *–se* in the first half of the sixteenth century.

Although it has long been discussed through bibliographic survey by many scholars that *these* and *those* underwent different processes of the development as mentioned above, no study has demonstrated this difference of the development in terms of the

time scale numerically and quantitatively. This study has successfully shown this by using the prototypal concept proposed by Kretzschmar (2009) i.e. A-curve. It has been revealed from the resulting evidence that the development of *these* had attained stability whereas the development of *those* was on its way by the second half of the fifteenth century. Further research on the finishing stage of the development of *those* is needed, especially from 1450 to 1550, in order to clarify the process of how one of the mere intermediate variants pushed its way to become the standard form.

This study has also attempted to grasp a whole picture of geographical distributions of numerous variants using *eLALME*. It goes without saying that these findings have been enabled by a large amount of surviving data made available by working in the *Linguistic Atlas* tradition over the last half century including the most notable *eLALME*. Even though it has a certain amount of deficiency in data collection, this database has allowed us not only to identify the characteristics of a particular area but also to analyse the data focusing on one particular linguistic feature, such as affix which implies different functions for the development of the variations.

Finally, this study has conducted a brief qualitative analysis on the functions of demonstratives by referring to the four different MSS of *Cursor Mundi*. There is found a slight difference in functions between *these* and *those*, as expected; both words are mostly (i.e. approx. 70-80%) used as demonstrative adjectives, but there are more tokens in THOSE whose functions are specialised in either way or used equally. This result might suggest the unstable and unfinished development of THOSE.

Although the findings in this thesis are relatively limited in scale, it can be said that the

attempt to develop a new methodology for analysis of variations has been successful. The method employed in this study will be applicable to any items or linguistic features of any stage of the development. It should be kept in mind, however, that much more analysis on semantic history is needed alongside this kind of quantitative survey, for which this study is inadequate overall.

Bibliography

An Electronic Version of A Linguistic Atlas of Late Mediaeval English (eLALME) (2013) (by Benskin, M., Laing, M., Karaiskos, V., Williamson, K.) available at: [http://www.lel.ed.ac.uk/ihd/elalme/elalme.html]

Anttila, R. (1977) *Analogy*. Mouton: The Hague.

Attridge, D. (1995) *Poetic Rhythm: An Introduction*. Cambridge, New York: Cambridge University Press.

Benskin, M. (1982) 'The letter <þ> and <y> in later middle English, and some related matters'. *Society of Archivists*, 7(1): 13-30.

Fell, C. (1982-83) '*Unfrið*: An Approach to a Definition', *Saga-Book of the Viking Society,* 21.

Firth, J. R. (1964) *Speech*. Oxford: Oxford University Press.

Horobin, S. and Smith, J. J. (2002) *An Introduction to Middle English*. Edinburgh: Edinburgh University Press.

Jesperson, O. (1956) *Growth and Structure of the English Language* (9th edn.). Oxford: Blackwell.

Krause, W. (1958) *Handbuch des Gotischen*. Munich: C.H. Beck'sche Verlagsbuchhandlung.

Kretzschmar, William A. (2009) *The linguistics of speech*. Cambridge: Cambridge University Press.

McIntyre, D. (2009) *History of English – A resource book for students*. London, New York: Routledge.

Millar, Robert M. (2000) *System collapse, system rebirth: the demonstrative pronouns of English, 900-1350 and the birth of the definite article*. Oxford: Peter Lang.

Morris, R. (ed) (1893) '*Cursor mundi* (The cursur of the world): a Northumbrian poem of the 14th Century in four versions', *The Early English Text Society* 57 59 62 66 68 99 101, K. Paul, Trench, London: Trübner.

Novelli, C. (1957) 'The Demonstrative Adjective This: Chaucer's Use of a Colloquial Narrative Device.' *Medieval Studies* 19: 246–9.

Oxford English Dictionary (OED) (2013). available at: [http://www.oed.com].

Pinker, S. (1994) *The Language Instinct*. Harmondsworth: Penguin.

Samuel, M. L. (1972) *Linguistic Evolution*. Cambridge: Cambridge University Press.

Samuel, M. L. & Smith, J. J. (1988) *The English of Chaucer and his contemporaries* (ed. By Smith, J. J.). Aberdeen: Aberdeen University Press.

Smith, J. J. (2006) 'Phonaesthesia, Ablaut and the History of the English Demonstratives', *Medieval English and its Heritage* (ed. by Ritt, N. et al), 16: 1-17, Frankfurt: Lang.

Smith, J. J. (2009) *Old English: a linguistic introduction*. Cambridge, New York, N.Y: Cambridge University Press.

Smith, J. J. (2012) ‘Chaucer’s Use of the Demonstrative’. *English Studies*, 93(5): 593-603.

Thompson, J. J. (1998) ‘The *Cursor Mundi*: Poem, Texts and Contexts’. Society for the Study of Mediaeval Languages and Literature: Oxford.

Townend, M. (2002) *Language and history in Viking age England: linguistic relations between speakers of Old Norse and Old English*. Turnhout: Brepols.

Trudgill, P. (1986) *Dialects in Contact*. Oxford: Basil Blackwell.

Wright, J., with O. L. Sayce (1954) ‘Grammar of the Gothic language: and the Gospel of St. Mark, selections from the other Gospels and the Second Epistle to Timothy’, Oxford: Clarendon Press.

Yokota, Y. (2006) ‘Form and function of demonstratives in the Middle English southern texts and speculation on the origin of *th*-type third person plural pronouns in the north and south’. *Notes and Queries*, 09/2006, 53(3): 300-303.

Appendix 1 (THESE)

01_Thees

02_Theos

03_Thes

04_These

05_Theys

06_This

07_Thise

08_Thuse

09_Thys

10_Thyse

11_dese

12_dyse

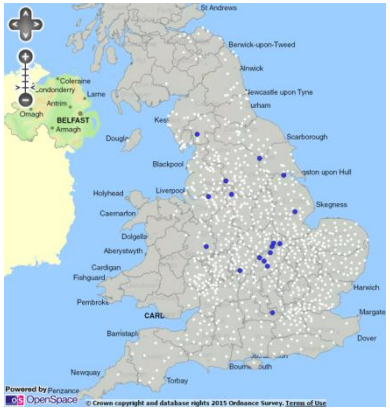
13_theer

14_thees

15_theese

16_theeɜ

17_theis



18_theise

19_theiɜ

20_theos

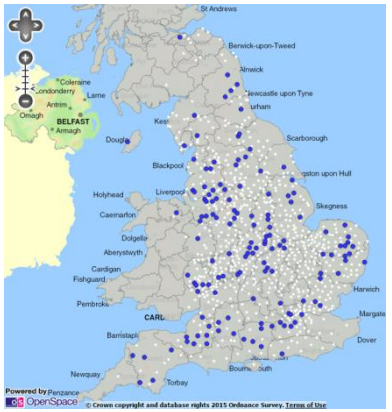
21_theose

22_ther

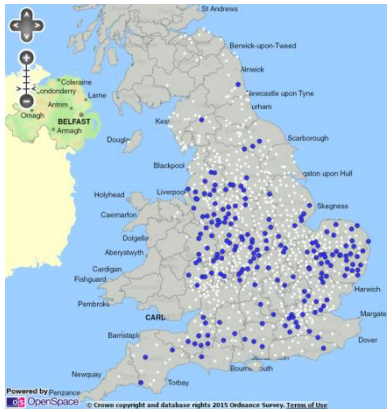
23_there

24_there

25_thes



26_these



27_theses

28_theseɜ

29_thesse

30_thesɜ

31_theyes

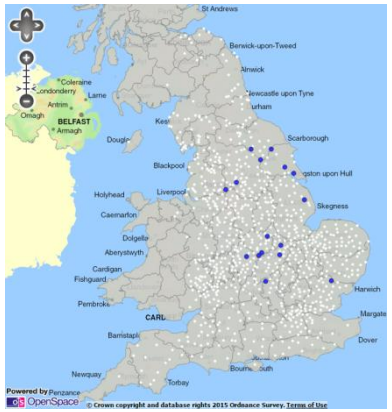
32_theys

33_theyse

34_the3

35_thier

36_thies



37_thiese

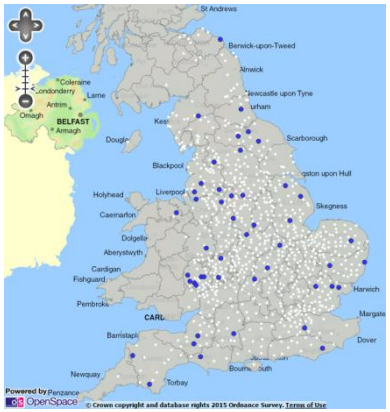
38_thie3

39_thir

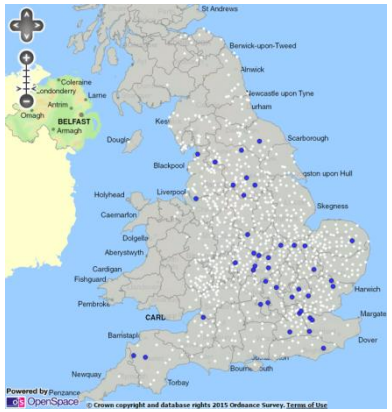
40_thire

41_thire

42_this



43_thise



44_thues

45_thus

Appendix 1 (THESE)

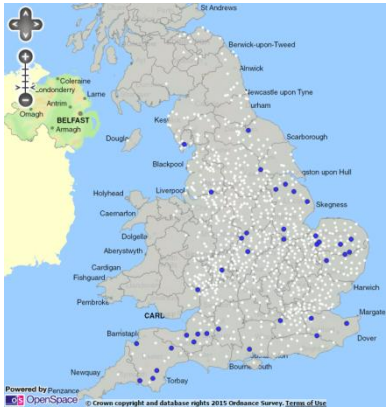
46_thuse

47_thyes

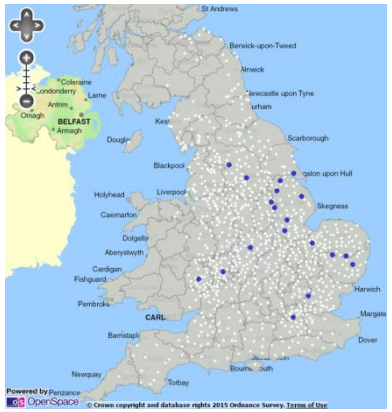
48_thyese

49_thyr

50_thys



51_thyse



52_tys

53_*yese*

54_*yer*

55_*yere*

56_y^{es}

57_y^{e}se

58_y^{i}se

59_y^{eis}

60_y^{es}

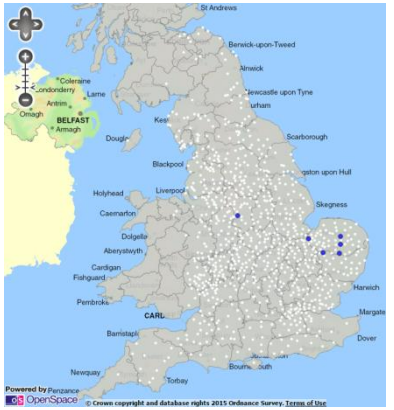
61_y^{e}es

62_y^{e}is

63_y^{e}ise

64_y^{e}s

65_y^{e}se



66_y^{e}ɜ

67_y^{ies}

68_y^{is}

69_y^{i}es

70_y^{i}r

71_y^{i}se

72_y^{s}

73_yair

74_ye*us*

75_yece

Appendix 1 (THESE)

76_yees

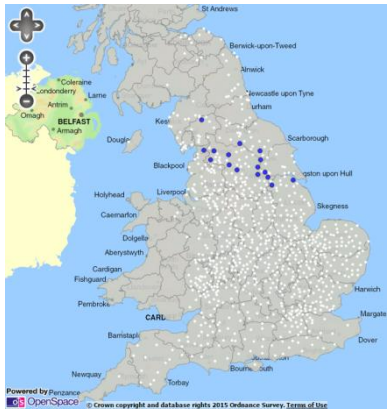
77_yeese

78_yeir

79_yeis

80_yeise

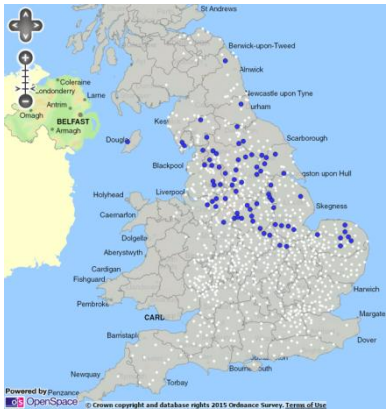
81_yer



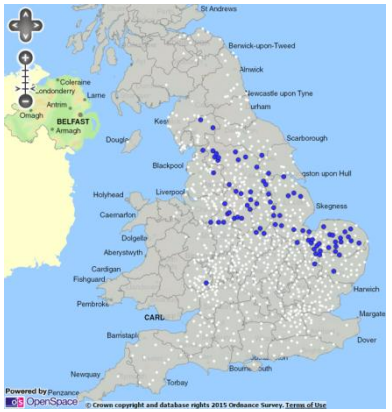
82_yere

83_yere

84_yes



85_yese



86_yess

87_yesse

88_yeys

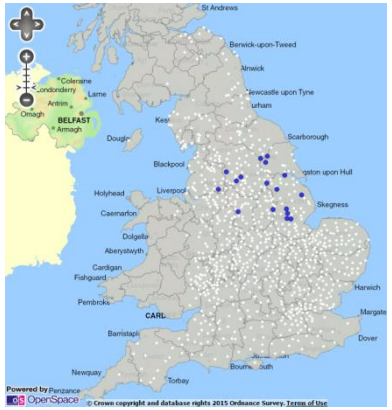
89_yeyse

90_yeɜ

91_yhese

92_yier

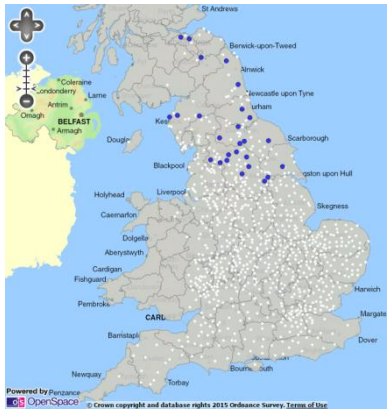
93_yies



94_yiese

95_yieȝ

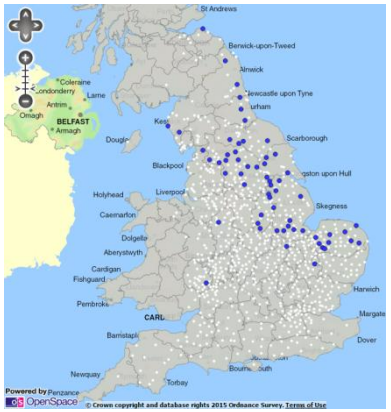
96_yir



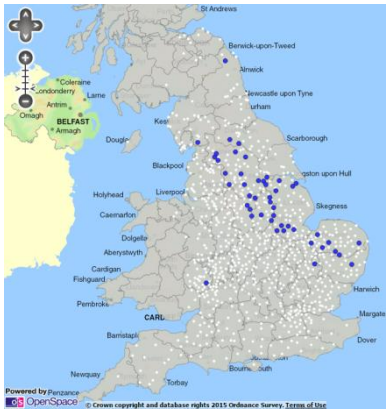
97_yire

98_yire

99_yis



100_yise



101_yises

102_yisse

103_yyes

104_*þer*

105_*þere*

106_þ*us*

107_þ^{eis}

108_þ^{es}

109_þ^{e}s

110_þ^{e}se

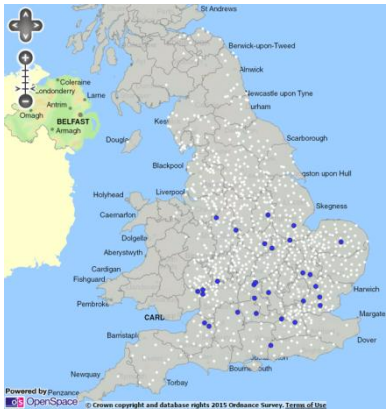
111_þ^{is}

112_þ^{s}

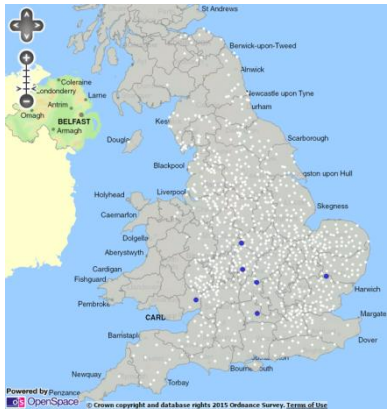
113_þe^{s}

114_þece

115_þees



116_þeese

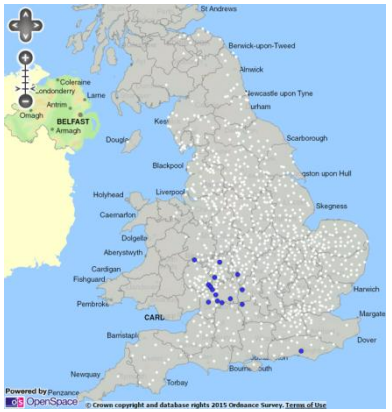


117_þeeʒ

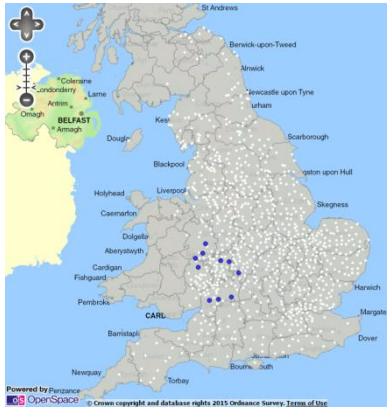
118_þeis

119_þeise

120_þeos



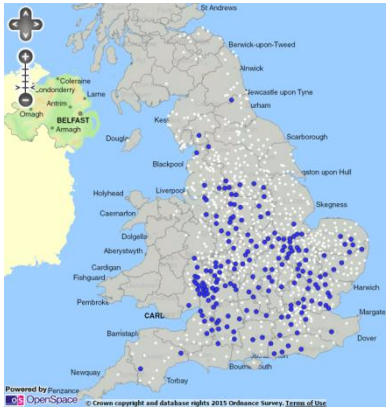
121_þeose



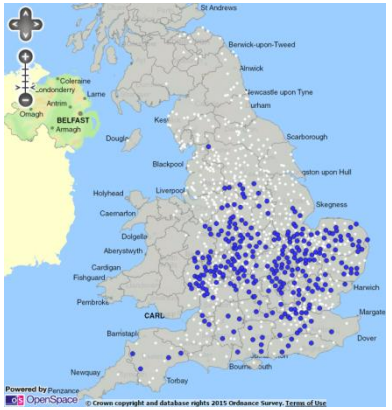
122_þer

123_þere

124_þes



125_þese



126_þesen

127_þesse

128_þeus

129_þeys

130_þeyse

131_þez

132_þeze

133_þeȝ

134_þeȝe

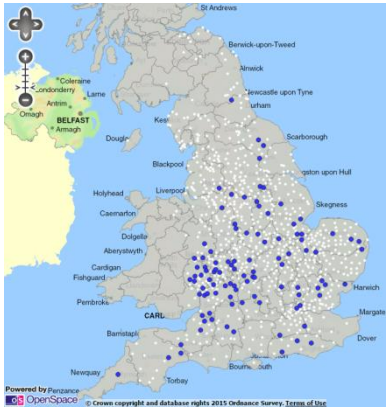
135_þies

136_þiis

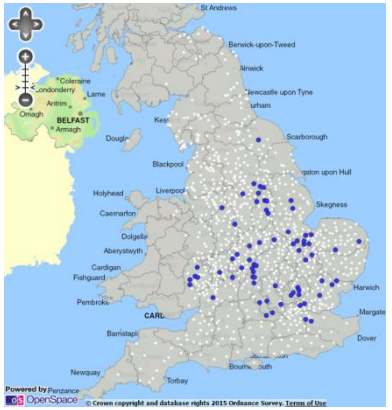
137_þir

138_þire

139_þis



140_þise



141_þisse

142_þiȝe

143_þoes

144_þos

145_þs

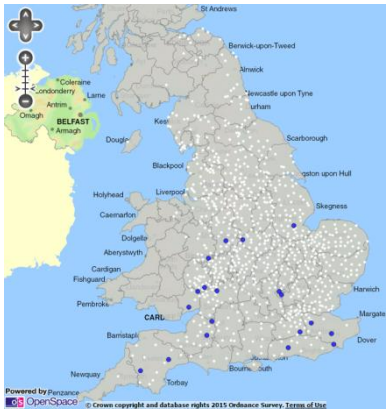
146_þues

147_þuese

148_þus

149_þuse

150_þys



Appendix 1 (THESE)

151_þyse

152_þy3e

153_3es

154_3ese

155_3eyse

156_3ise

01_tha

02_thai

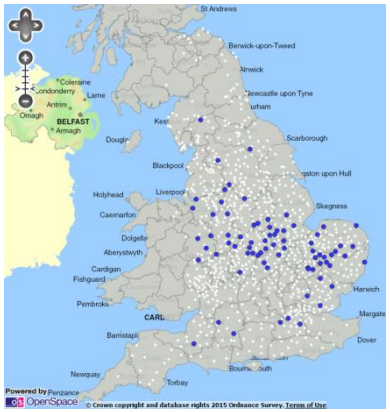
03_thar

04_thas

05_thase

06_the

07_tho

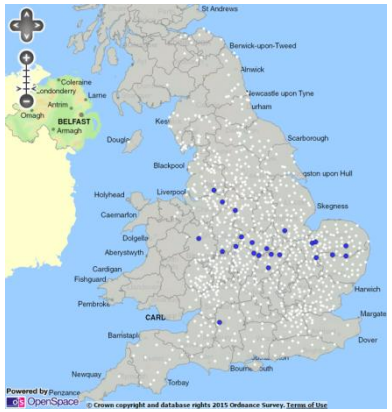


08_thoe

09_thoes

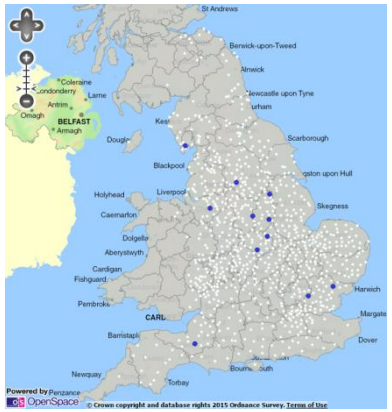
10_thois

11_thoo

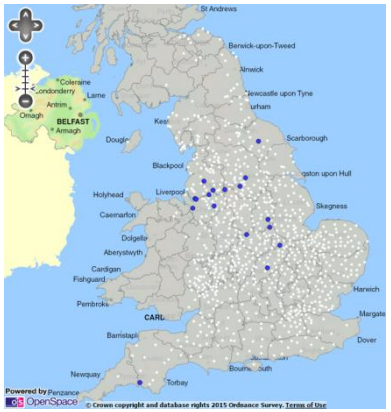


12_thoos

13_thos



14_those



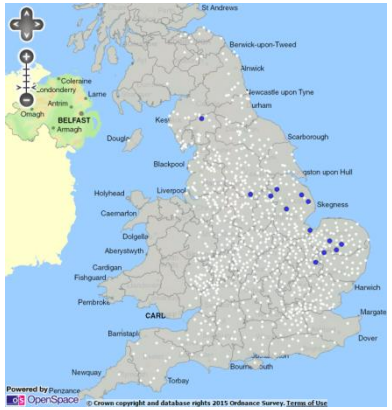
15_thot

Appendix 2 (THOSE)

16_thow

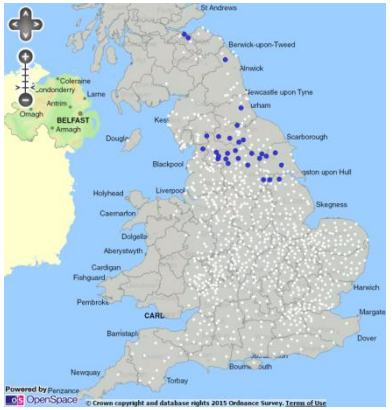
17_thoys

18_y^{o}



19_y^{os}

20_ya



21_yaa

22_yaas

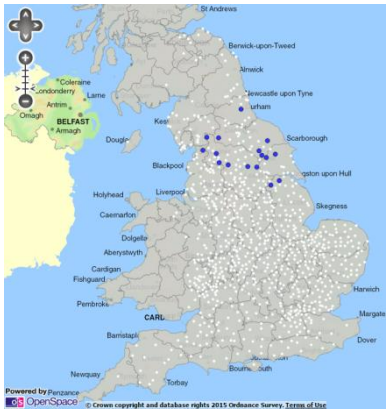
23_yai

24_yais

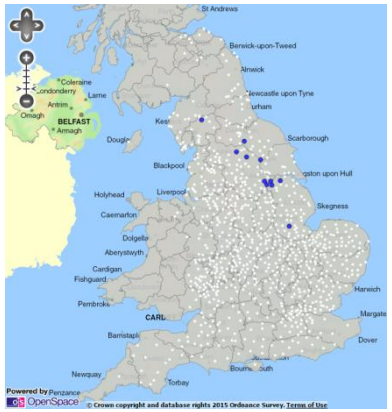
25_yaise

26_yaisse

27_yas



28_yase



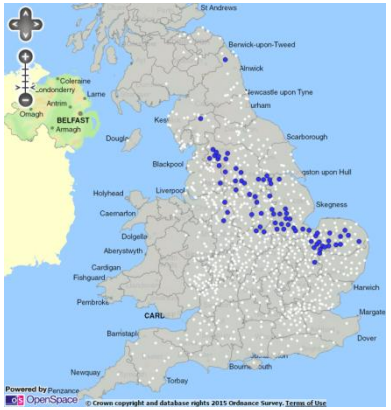
29_yay

30_yayes

Appendix 2 (THOSE)

31_yho

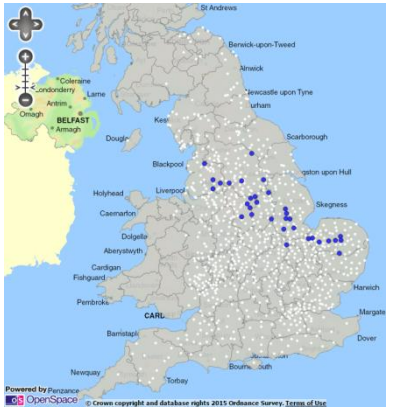
32_yo



33_yo^{s}

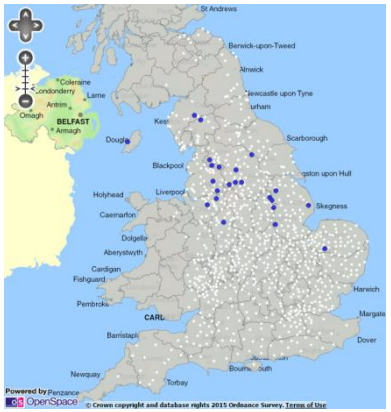
34_yois

35_yoo

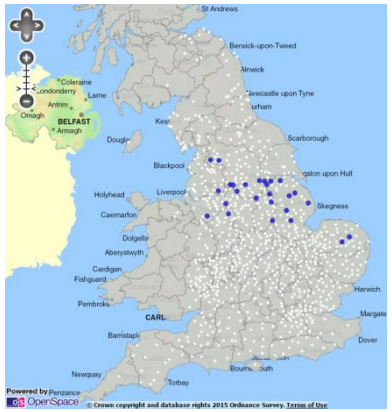


36_yoose

37_yos

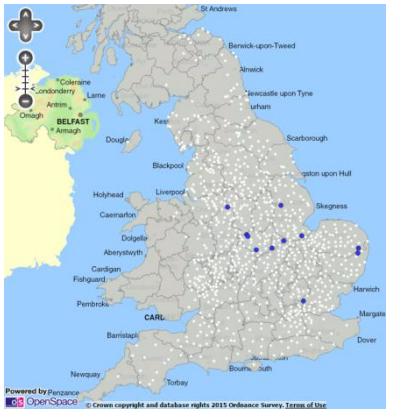


38_yose



39_yoʒ

40_þ^{o}



41_þ^{o}se

42_þa

43_þaa

44_þaas

45_þaes

Appendix 2 (THOSE)

46_þai

47_þais

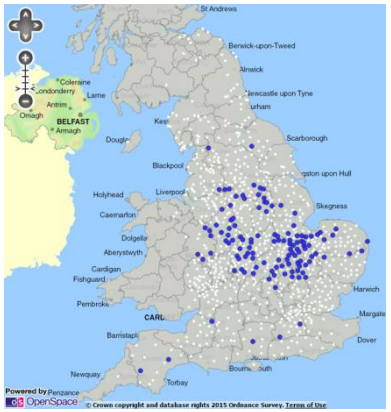
48_þas

49_þase

50_þay

51_þhose

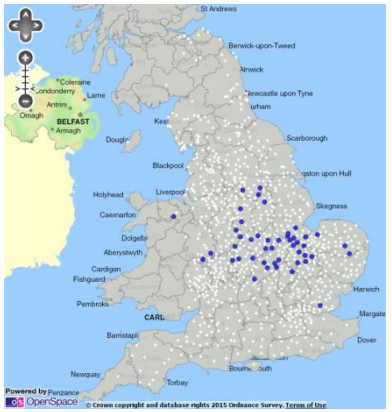
52_þo



53_þo^{o}

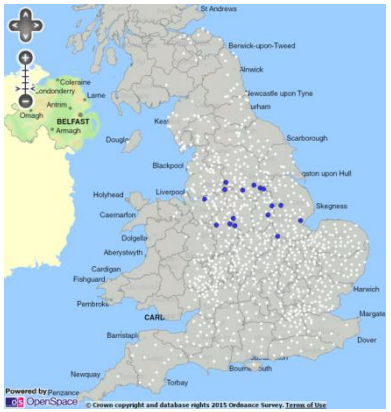
54_þoe

55_þoo

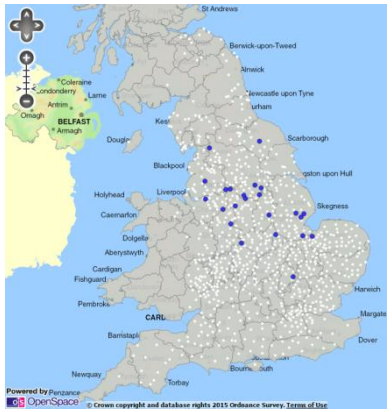


55_þoo

57_þos



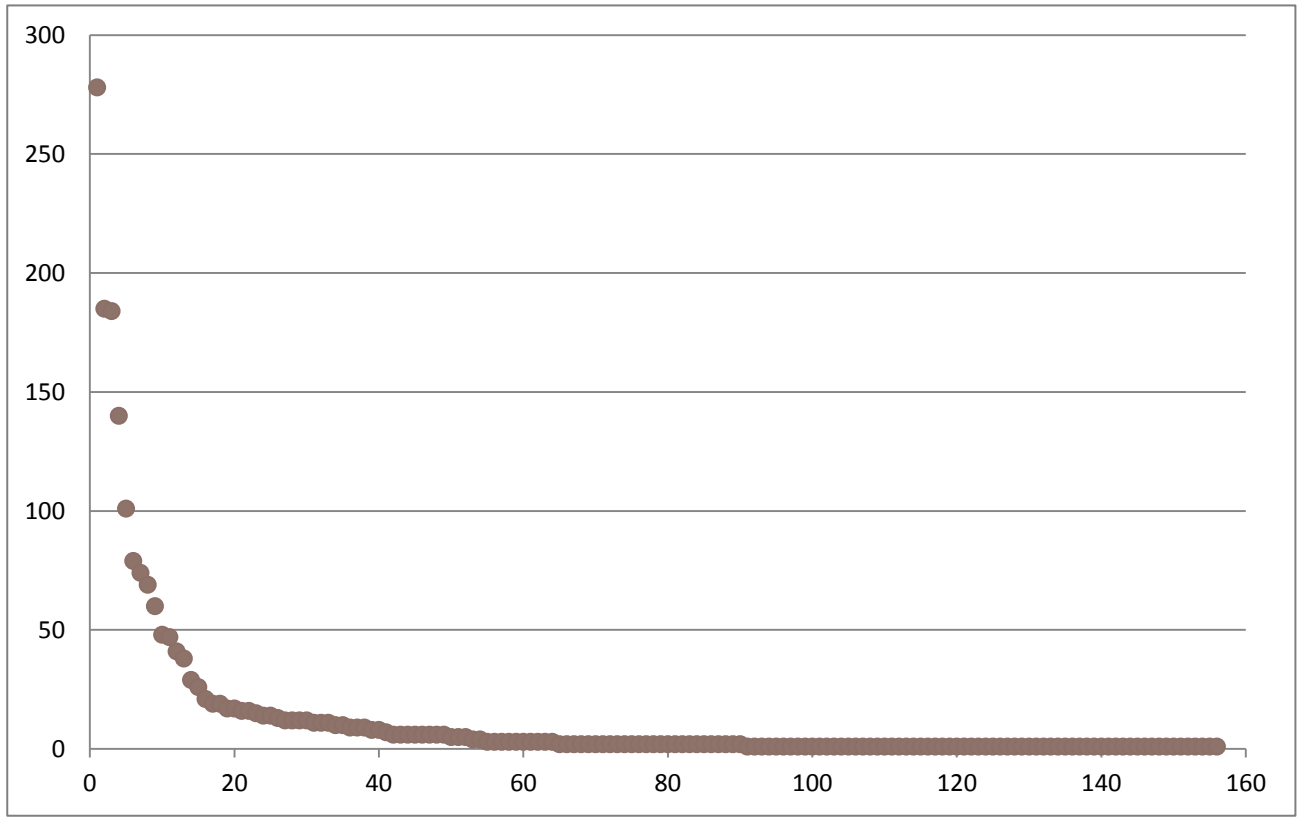
58_þose



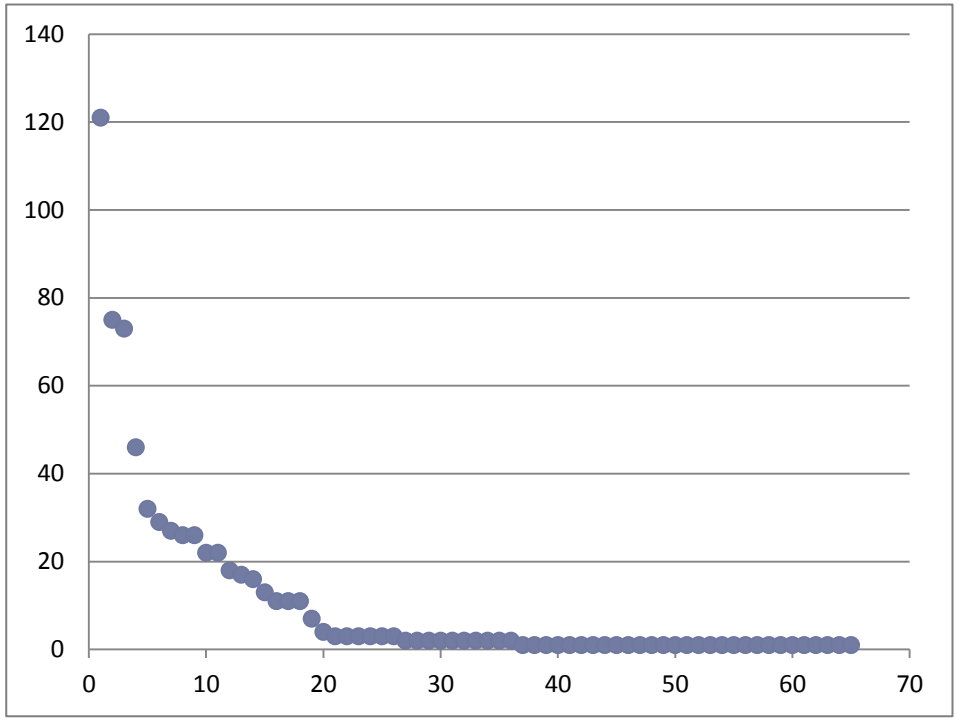
No.	Word	Count	Ranking
125	þese	278	1
26	these	185	2
124	þes	184	3
25	thes	140	4
139	þis	101	5
85	yese	79	6
84	yes	74	7
140	þise	69	8
99	yis	60	9
42	this	48	10
100	yise	47	11
43	thise	41	12
50	thys	38	13
115	þees	29	14
96	yir	26	15
14	thees	21	16
51	thyse	19	17
150	þys	19	17
17	theis	17	19
93	yies	17	19
36	thies	16	21
81	yer	16	21
149	þuse	15	23
32	theys	14	24
120	þeos	14	24
110	þ^(e)se	13	26
47	thyes	12	27
60	y^es	12	27
119	þeise	12	27
151	þyse	12	27
4	These	11	31
22	ther	11	31
39	thir	11	31
121	þeose	10	34
148	þus	10	34
54	y*er*	9	36
68	y^is	9	36
118	þeis	9	36
18	theise	8	39
76	yees	8	39
72	y^s	7	41
3	Thes	6	42
65	y^(e)se	6	42
79	yeis	6	42
80	yeise	6	42
112	þ^s	6	42
116	þeese	6	42
135	þies	6	42
137	þir	6	42
29	thesse	5	50
83	yere	5	50
111	þ^is	5	50
34	the$_{3}$	4	53
108	þ^es	4	53
9	Thys	3	55
20	theos	3	55
24	there	3	55
45	thus	3	55
78	yeir	3	55
90	ye$_{3}$	3	55
91	yhese	3	55
92	yier	3	55
127	þesse	3	55
134	þe$_{3}$e	3	55
6	This	2	65
11	dese	2	65
15	theese	2	65
21	theose	2	65
33	theyse	2	65
41	thire	2	65
46	thuse	2	65
62	y^(e)is	2	65
64	y^(e)s	2	65
66	y^(e)$_{3}$	2	65
67	y^ies	2	65
77	yeese	2	65
82	yer*e*	2	65
88	yeys	2	65
94	yiese	2	65
97	yir*e*	2	65
102	yisse	2	65
104	þ*er*	2	65
106	þ*us*	2	65
122	þer	2	65
126	þesen	2	65
129	þeys	2	65
130	þeyse	2	65
133	þe$_{3}$	2	65
146	þues	2	65
153	$_{3}$es	2	65
1	Thees	1	91
2	Theos	1	91
5	Theys	1	91
7	Thise	1	91
8	Thuse	1	91
10	Thyse	1	91
12	dyse	1	91
13	theer	1	91
16	thee$_{3}$	1	91
19	thei$_{3}$	1	91
23	there	1	91
27	theses	1	91
28	these$_{3}$	1	91
30	thes$_{3}$	1	91
31	theyes	1	91

35	thier	1	91
37	thiese	1	91
38	thie$_{3}$	1	91
40	thire	1	91
44	thues	1	91
48	thyese	1	91
49	thyr	1	91
52	tys	1	91
53	ye se	1	91
55	ye*r* e	1	91
56	y^{e}^s	1	91
57	y^{e}se	1	91
58	y^{i}se	1	91
59	y^eis	1	91
61	y^(e)es	1	91
63	y^(e)ise	1	91
69	y^(i)es	1	91
70	y^(i)r	1	91
71	y^(i)se	1	91
73	yair	1	91
74	ye*us*	1	91
75	yece	1	91
86	yess	1	91
87	yesse	1	91
89	yeyse	1	91
95	yie$_{3}$	1	91
98	yire	1	91
101	yises	1	91
103	yyes	1	91
105	þe*r* e	1	91
107	þ^eis	1	91
109	þ^(e)s	1	91
113	þe^s	1	91
114	þece	1	91
117	þee$_{3}$	1	91
123	þere	1	91

128	þeus	1	91
131	þez	1	91
132	þeze	1	91
136	þiis	1	91
138	þire	1	91
141	þisse	1	91
142	þi$_{3}$e	1	91
143	þoes	1	91
144	þos	1	91
145	þs	1	91
147	þuese	1	91
152	þy$_{3}$e	1	91
154	$_{3}$ese	1	91
155	$_{3}$eyse	1	91
156	$_{3}$ise	1	91

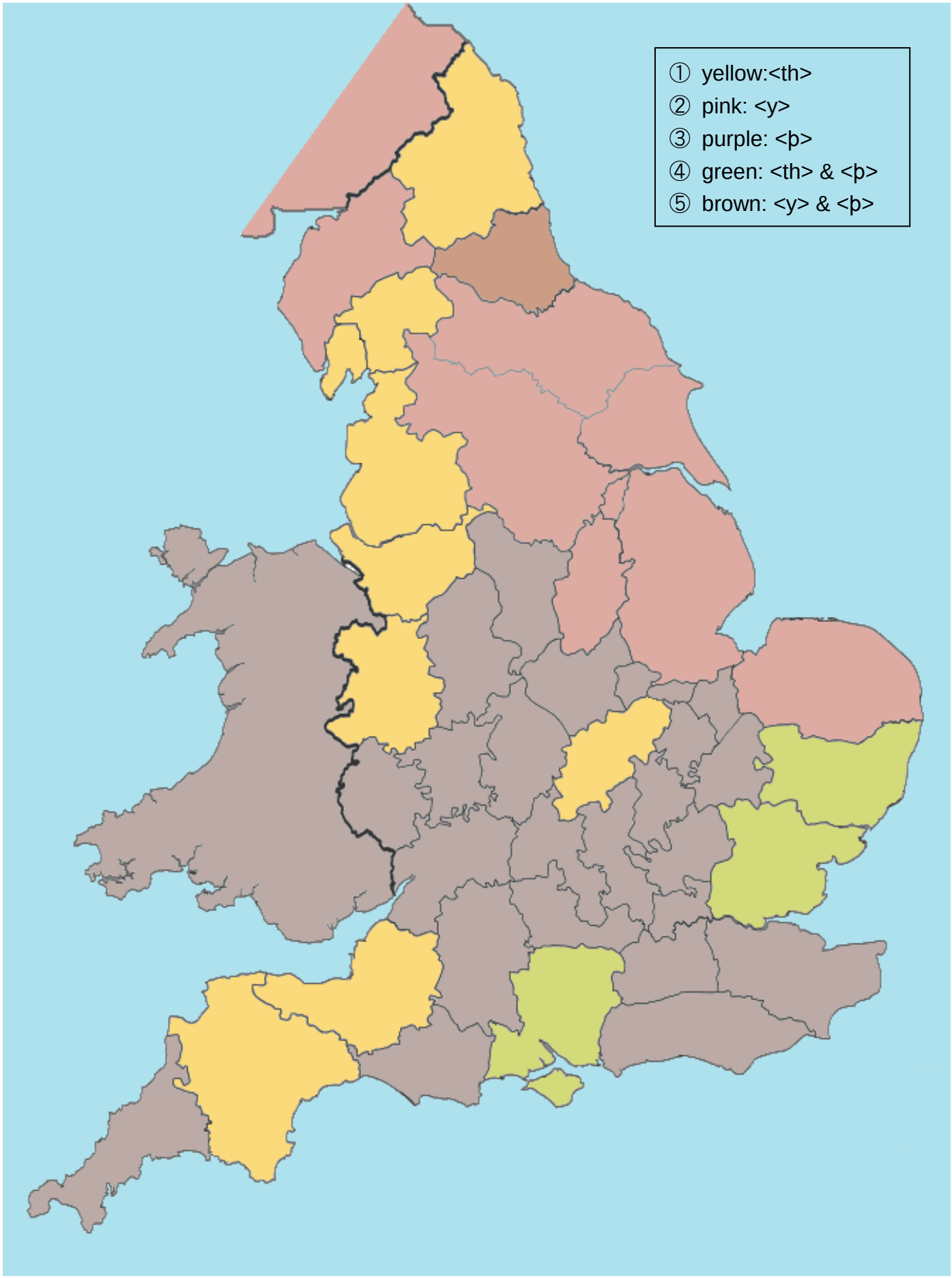


No.	Word	Count	Ranking
58	þo	121	1
10	tho	75	2
35	yo	73	3
61	þoo	46	4
38	yoo	32	5
23	ya	29	6
64	þose	27	7
41	yose	26	8
65	þoȝ	26	8
14	thoo	22	10
40	yos	22	10
63	þos	18	12
17	those	17	13
30	yas	16	14
21	y^o	13	15
16	thos	11	16
31	yase	11	16
43	þ^o	11	16
27	yais	7	19
46	þa	4	20
1	tha	3	21
5	thase	3	21
9	they	3	21
26	yai	3	21
52	þas	3	21
53	þase	3	21
2	thai	2	27
6	thay	2	27
12	thoes	2	27
22	y^os	2	27
24	yaa	2	27
48	þaas	2	27
49	þaes	2	27
54	þay	2	27
56	þei	2	27
62	þoos	2	27
3	thar	1	37
4	thas	1	37
7	thaye	1	37
8	the	1	37
11	thoe	1	37
13	thois	1	37
15	thoos	1	37
18	thot	1	37
19	thow	1	37
20	thoys	1	37
25	yaas	1	37
28	yaise	1	37
29	yaisse	1	37
32	yay	1	37
33	yayes	1	37
34	yho	1	37
36	yo^s	1	37
37	yois	1	37
39	yoose	1	37
42	yoȝ	1	37
44	þ^(o)se	1	37
45	þ^(t)	1	37
47	þaa	1	37
50	þai	1	37
51	þais	1	37
55	þaye	1	37
57	þhose	1	37
59	þo^o	1	37
60	þoe	1	37

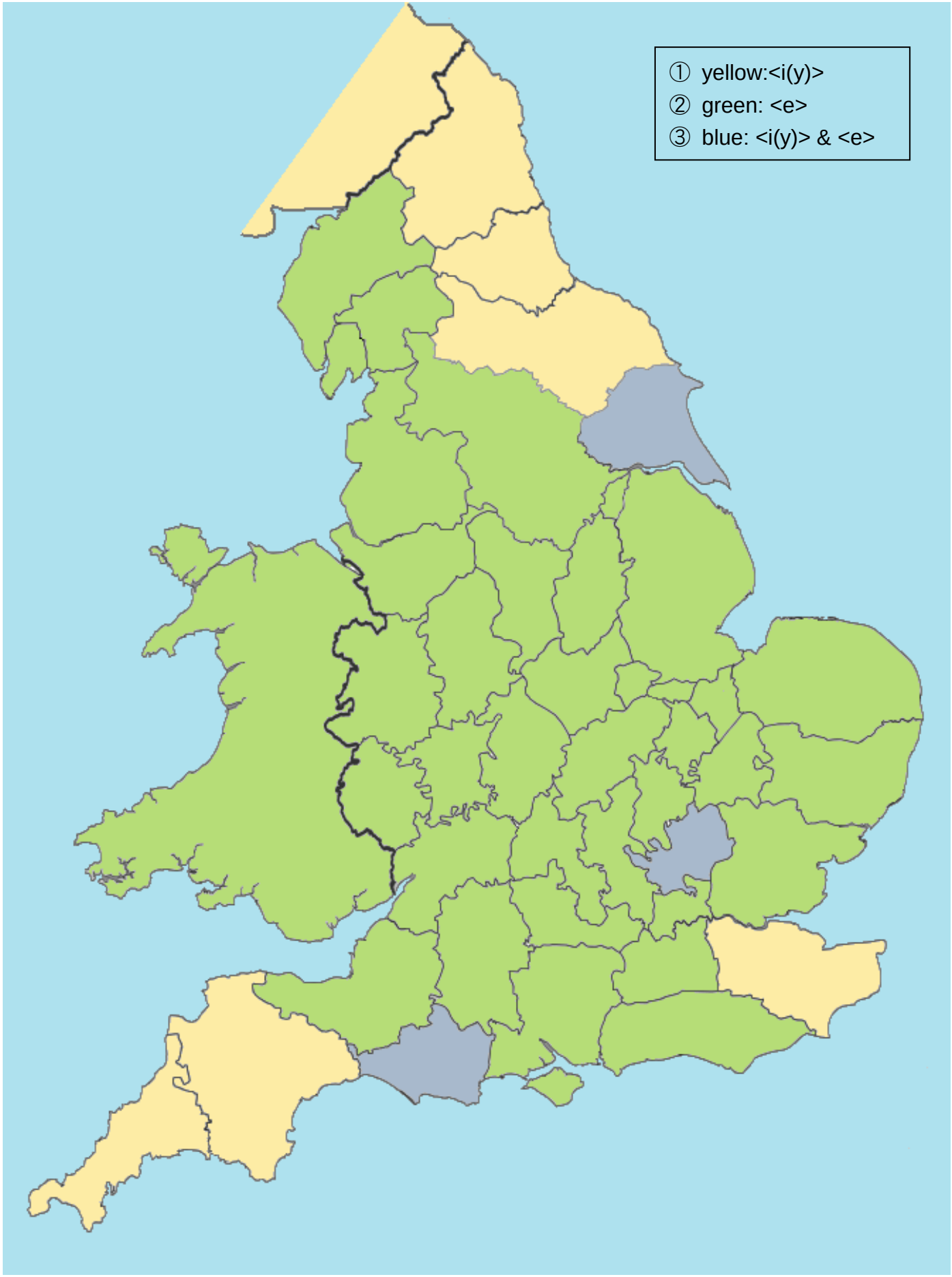


The most common token (THESE)

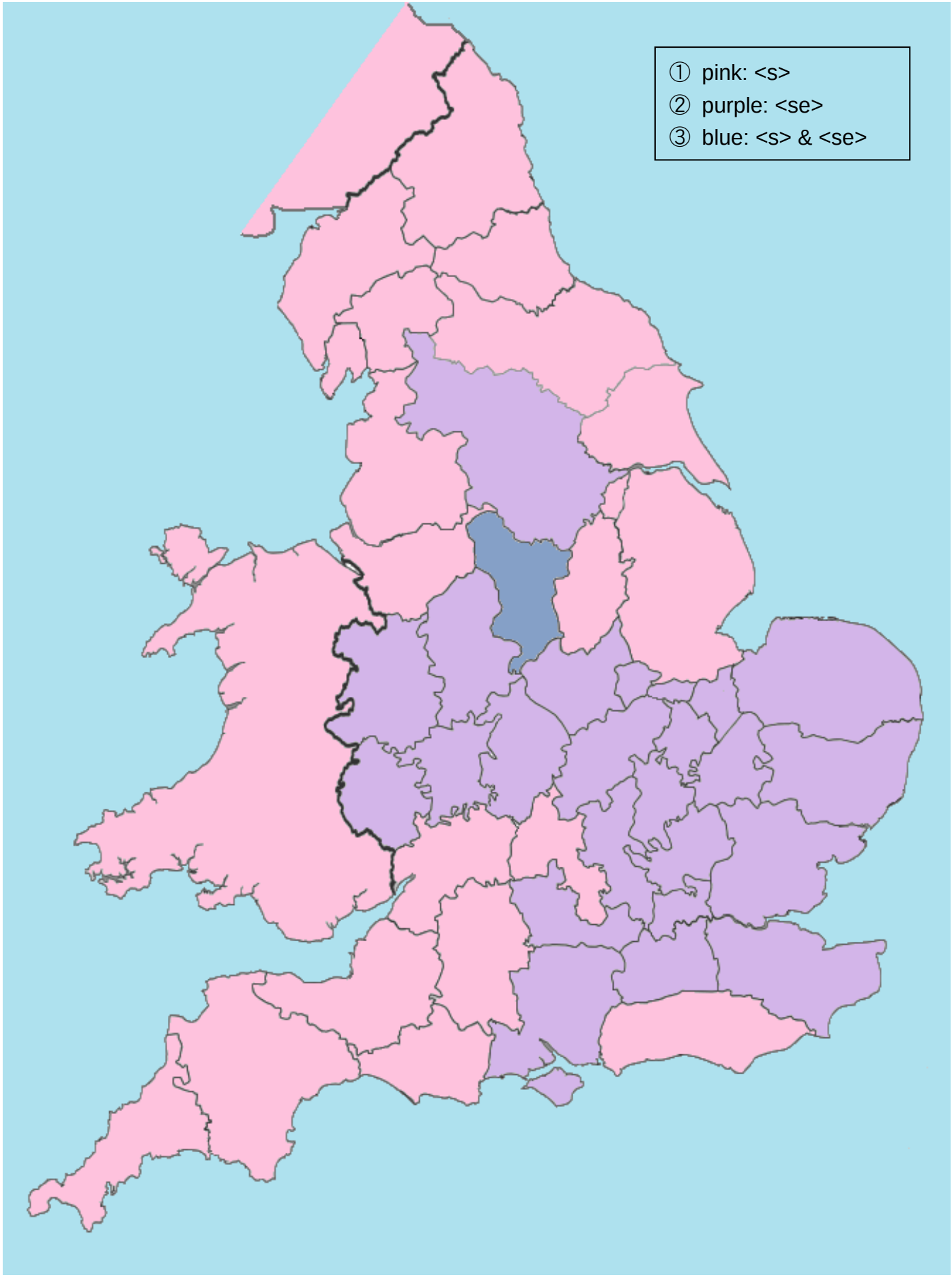
The most common initial letters (THESE)



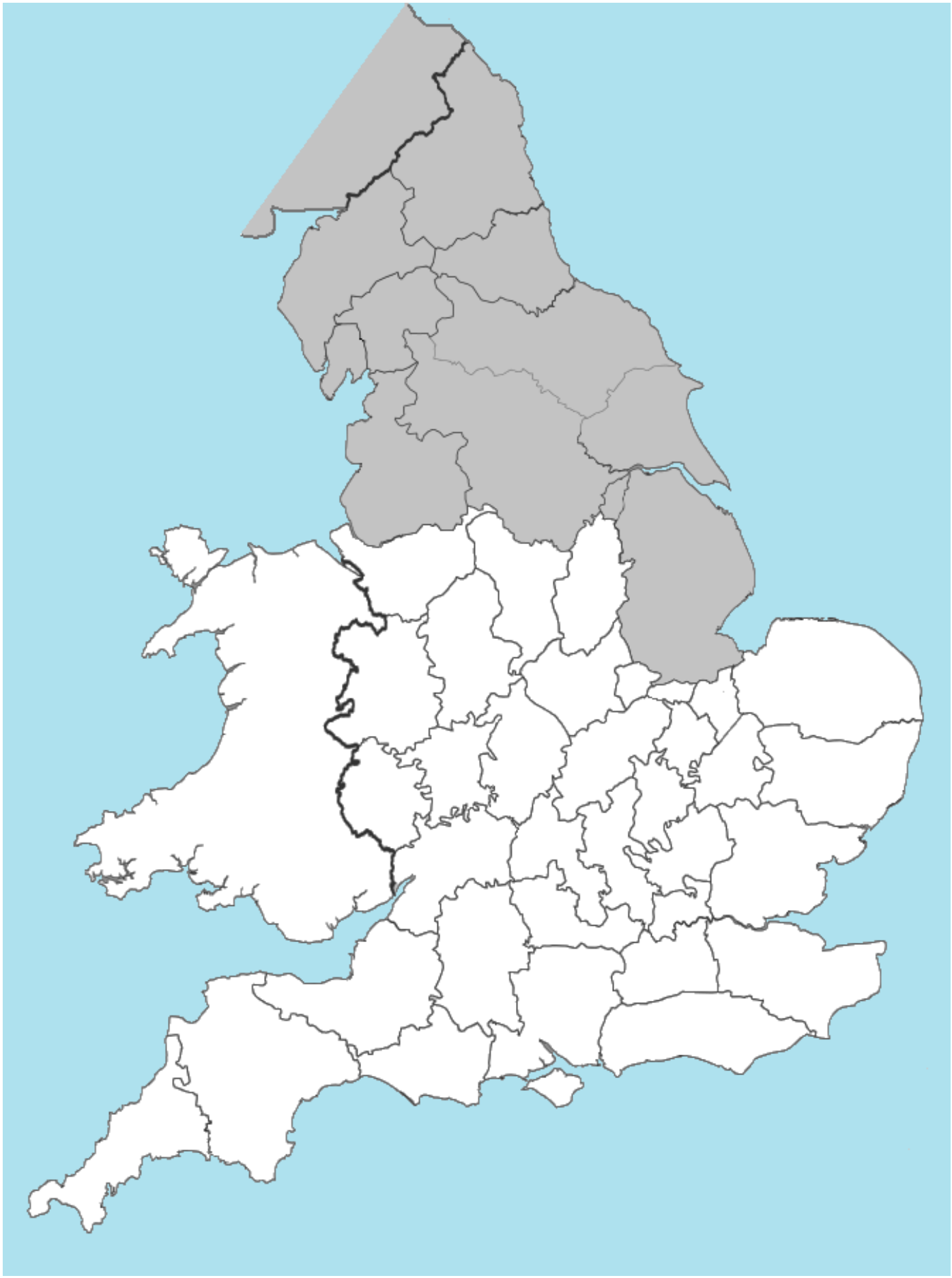
The most common medial letters (THESE)



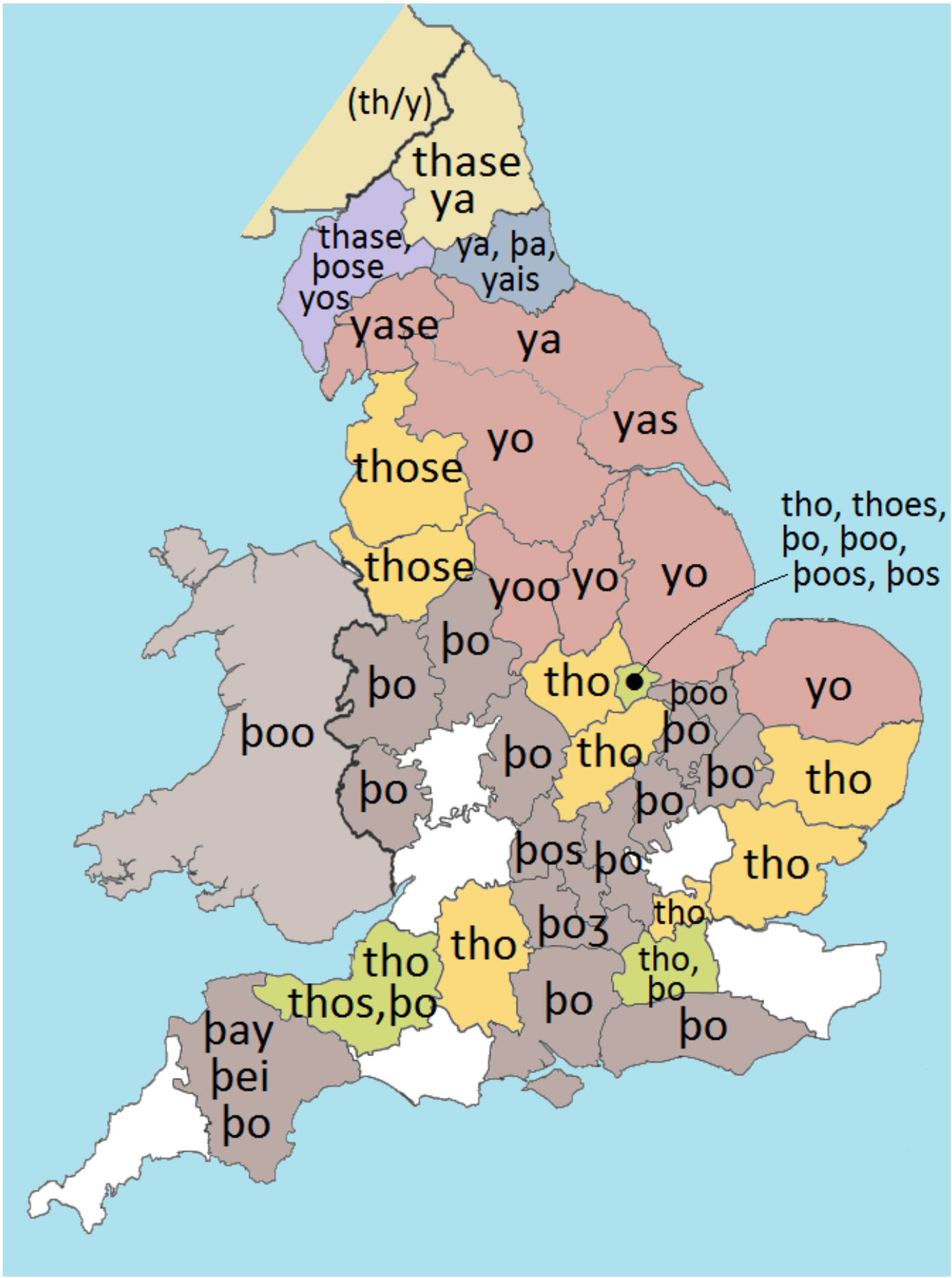
The most common final letters (THESE)



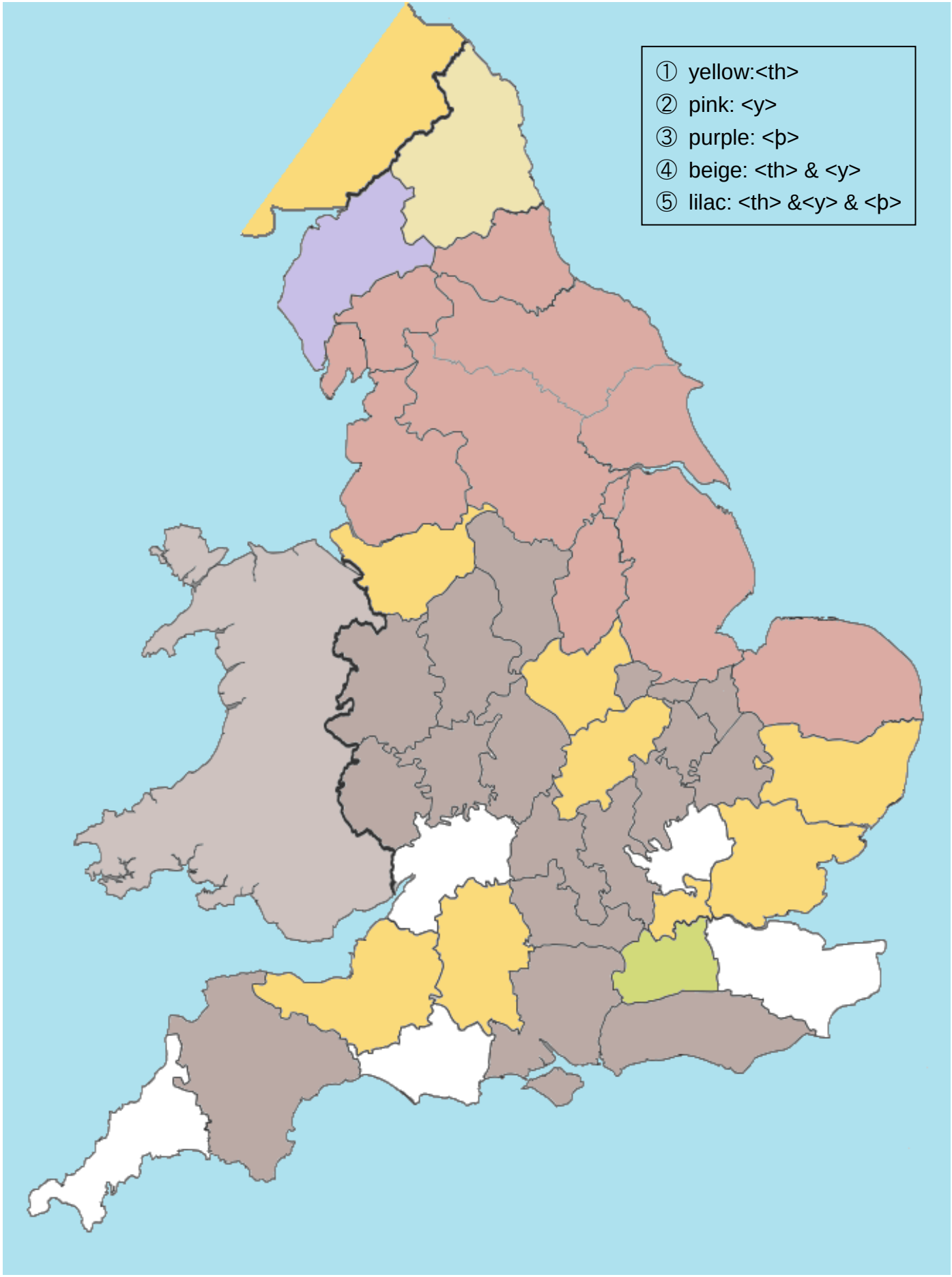
The final letter <r(e)> (THESE)



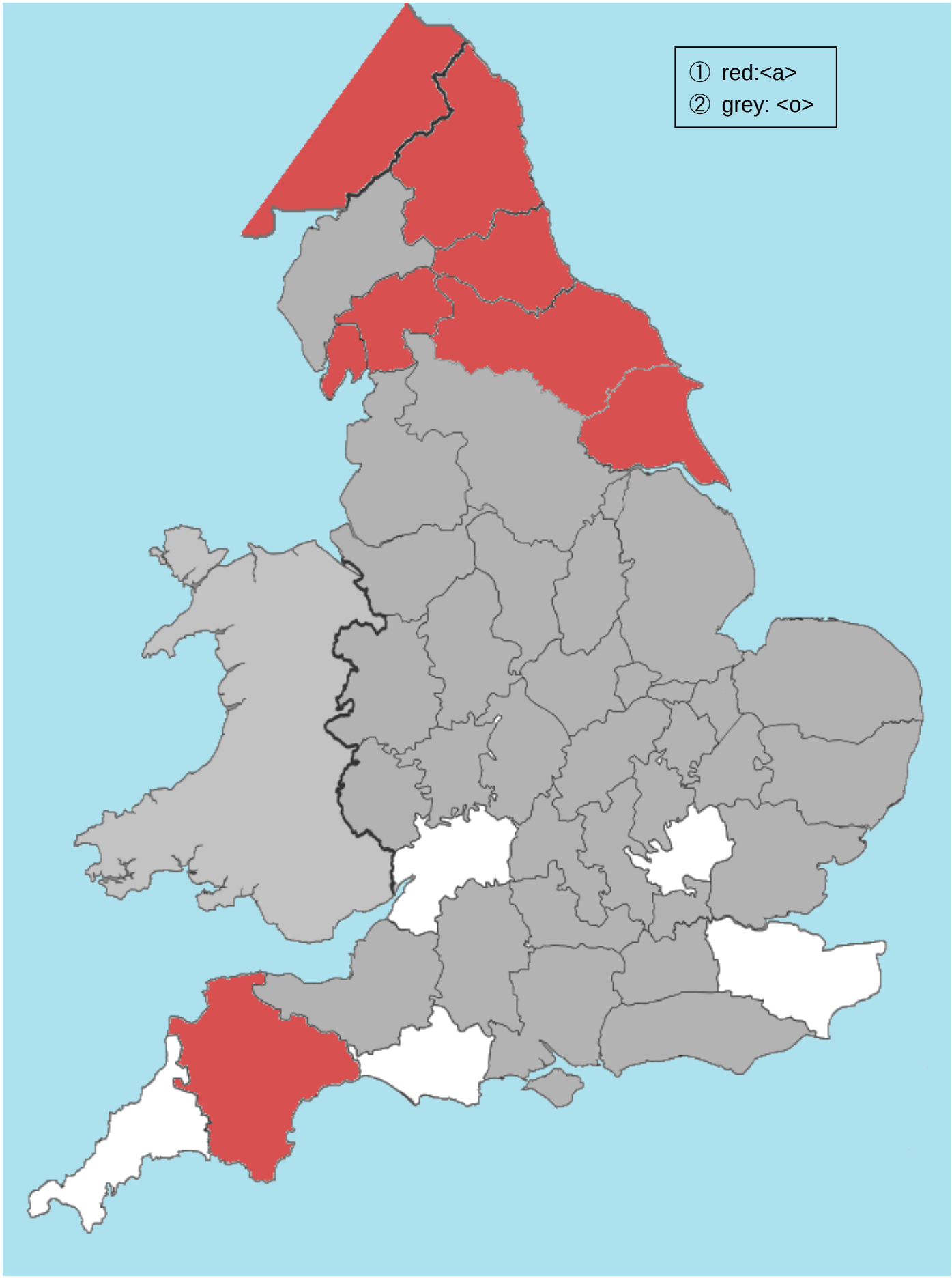
The most common token (THOSE)



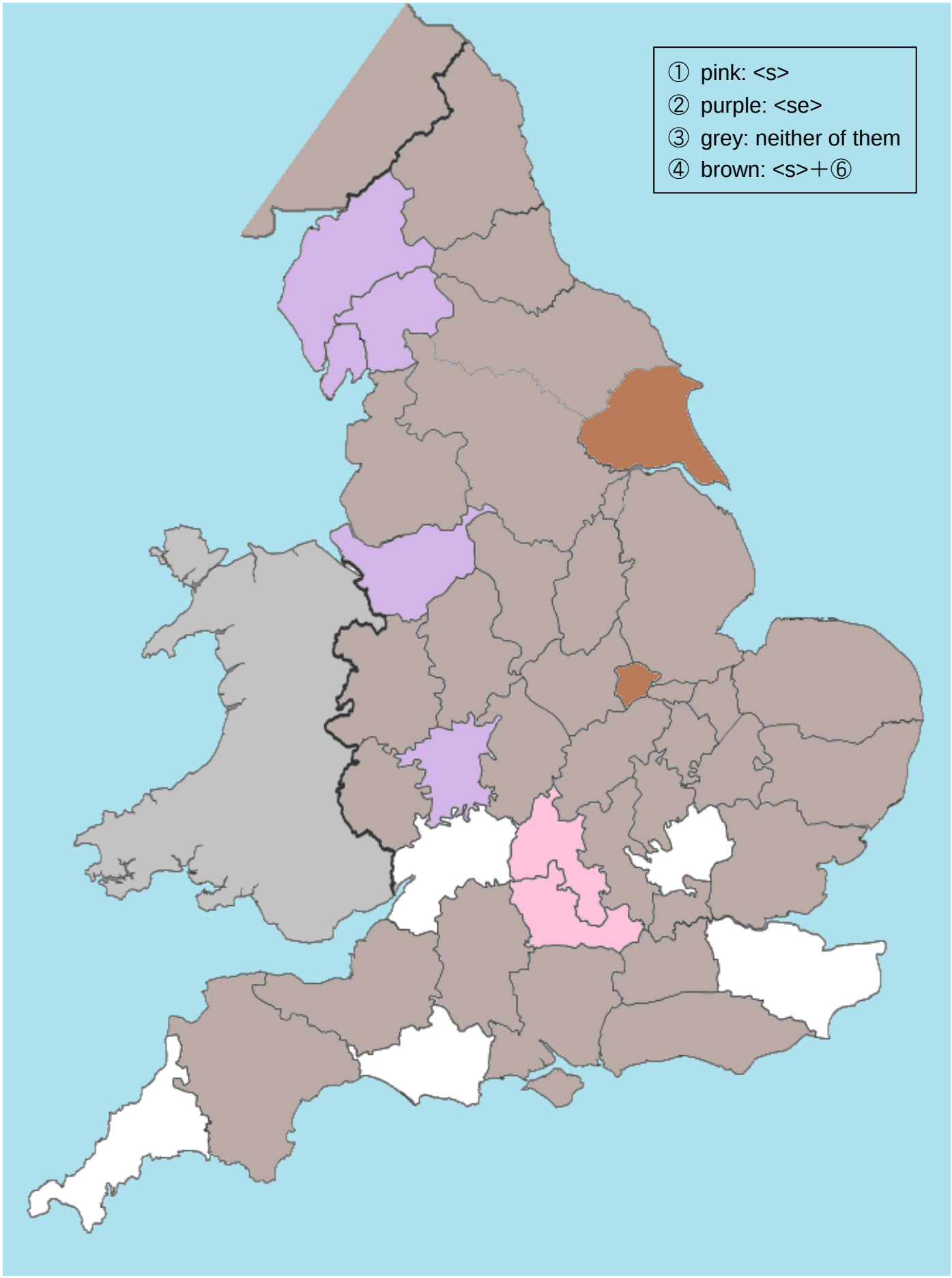
The most common initial letters (THOSE)



The most common medial letters (THOSE)



The most common final letters (THOSE)



County			most common token	initial						medial				final					
				variation			citation			variation		citation		variation			citation		
				th	y	þ	th	y	þ	i(y)	e	i(y)	e	s	se	r(e)	s	se	r(e)
England	SE	Kent (12)	thise	3	0	6	5	0	8	5	4	8	5	3	6	0	4	9	0
		Surrey (20)	þese	6	0	8	11	0	15	6	8	6	19	7	7	0	10	15	0
		Sussex (14)	þes	4	0	8	5	0	18	5	7	8	15	7	7	0	13	12	0
		Total					21	0	41			22	39				27	36	0
	SW	Berkshire (7)	þese	1	0	6	1	0	11	2	4	2	9	3	4	0	5	7	0
		Cornwall (1)	(þis)	0	0	1	0	0	1	1	0	1	0	1	0	0	1	0	0
		Devon (17)	thys	7	0	6	14	0	10	6	4	12	9	8	5	0	17	7	0
		Dorset (5)	(this, þes, þu	1	0	2	1	0	2	1	1	1	1	2	1	0	2	1	0
		Gloucestershire (40)	þes	5	0	10	8	0	44	4	6	11	27	10	6	0	35	13	0
		Hampshire (17)	þese	3	0	2	9	0	9	1	4	1	17	3	2	0	8	10	0
		Oxfordshire (10)	þese	3	0	9	3	0	14	3	4	4	9	9	3	0	10	8	0
		Somerset (19)	thes, these	5	0	6	17	0	10	6	4	9	17	6	4	0	16	11	0
		Wiltshire (19)	þes	2	0	6	6	0	15	1	5	1	18	5	1	0	16	3	0
		Total					59	0	116			42	107				110	60	0
	EM	Bedfordshire (8)	þese	1	0	5	1	0	11	1	3	1	9	3	3	0	3	9	0
		Buckinghamshire (12)	þese	5	0	5	6	0	12	3	5	3	12	5	6	0	6	12	0
		Cambridgeshire (16)	þes, þese	2	1	4	5	1	12	3	4	3	15	2	5	0	6	12	0
		Essex (35)	þese	9	0	5	26	0	26	4	10	11	37	9	8	0	23	29	0
		Hertfordshire (10)	þese	3	0	6	3	0	9	4	3	5	5	2	7	0	2	10	0
		Huntingdonshire (12)	þese	0	2	3	0	2	11	1	3	1	11	2	3	0	2	9	0
		Leicetstershire (27)	þese	8	1	8	17	1	18	2	9	2	28	10	7	0	17	18	0
		Lincolnshire (50)	yis	8	14	6	15	34	13	8	7	21	23	15	8	2	30	25	2
		London (6)	þese	3	0	3	5	0	5	2	5	3	7	3	4	0	3	7	0
		Middlesex (11)	þese	2	0	5	5	0	11	2	5	2	13	3	4	0	6	9	0
		Norfolk (67)	þese	8	16	9	35	62	39	8	9	24	88	11	15	0	34	77	0
		Northamptonshire (41)	thiese, þese	9	0	4	27	0	19	3	6	6	22	7	7	0	19	28	0
		Nottinghamshire (26)	yes, þise	2	7	5	2	15	10	4	7	10	14	8	5	0	14	12	0
		Soke of Peterborough (3)	þese	0	0	2	0	0	3	1	1	1	2	0	2	0	0	3	0
		Rutland (5)	þese	3	0	4	3	0	5	2	3	2	4	3	4	0	3	5	0
		Suffolk (31)	þese	6	4	8	23	4	23	5	9	6	38	7	9	0	11	35	0
		Yorkshire, West Riding (72)	yese	8	15	6	18	58	14	12	11	37	52	18	11	7	40	52	20
		Total					191	177	241			120	307				219	352	22
	WM	**Cheshire (35)**	thes	7	3	5	22	7	7	6	8	7	29	9	5	0	24	13	0

	Derbyshire (21)	yese, þes, þe	3	3	2	5	7	8	0	7	0	19	5	3	0	10	10	0
	Herefordshire (27)	þese	3	1	8	4	1	28	2	8	5	25	7	6	0	14	19	0
	Lancashire (58)	these	9	6	2	28	16	3	3	11	4	41	13	4	3	31	15	3
	Shropshire (17)	these	4	1	3	11	1	9	1	5	1	17	3	5	0	4	17	0
	Staffordshire (32)	these	3	7	3	14	10	18	1	8	1	37	8	3	0	17	23	0
	Warwickshire (34)	þese	5	0	8	17	0	32	3	5	9	30	6	7	0	13	31	0
	Worcestershire (25)	þese	4	0	8	12	0	30	2	5	5	30	8	4	0	18	24	0
	Total					113	42	135			32	228				131	152	3
N	**Cumberland (32)**	yer	5	6	1	7	11	1	5	5	8	9	8	3	6	14	4	11
	Durham (18)	yis	5	4	1	7	7	1	5	4	9	4	8	1	4	11	1	7
	Northern (15)	yir	2	10	0	2	21	0	5	4	12	8	6	4	6	15	6	12
	Northumberland (19)	thes	5	3	0	9	5	0	5	3	8	6	7	1	3	13	1	5
	Westmorland (17)	thes	5	2	2	6	3	2	2	5	2	6	5	1	0	6	3	0
	York (5)	(theis, thies,	2	1	2	2	1	2	1	1	1	1	5	0	1	5	0	1
	Yorkshire (2)	yes	0	2	0	0	3	0	1	1	1	2	2	0	0	2	0	0
	Yorkshire, North West (4)	yes	0	3	0	0	4	0	1	2	1	3	3	0	1	4	0	1
	Yorkshire, East Riding (23)	yes, yis	5	9	0	8	14	0	4	7	9	9	9	4	5	15	6	7
	Yorkshire, North Riding (30)	this	7	9	3	13	14	3	7	9	16	11	14	5	7	26	5	10
	Total					54	83	9			67	59				111	26	54
	Unlocalised (3)	(these, thise)	2	0	0	2	0	0	1	1	1	1	1	1	0	1	1	0
Wales	Carmarthen (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Denbighshire (3)	(thes, this)	2	0	0	2	0	0	1	1	1	1	2	0	0	2	0	0
	Monmouth (4)	þes	0	0	4	0	0	5	1	3	1	4	2	2	0	3	2	0
	Montgomery (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total					2	0	5			2	5				5	2	0
Scotland	Ayrshire (2)	(yer , yir, ȝyr	0	2	0	0	2	0	2	1	2	1	3	0	3	3	0	3
	Berwickshire (1)	(thir)	1	0	0	1	0	0	1	0	1	0	1	0	1	1	0	1
	Dumfriesshire (1)	(thir)	1	0	0	1	0	0	1	0	1	0	1	0	1	1	0	1
	East Lothian (3)	(yis)	0	1	0	0	1	0	1	0	1	0	1	0	0	1	0	0
	Lanarkshire (1)	(thir)	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0
	Midlothian (5)	yir	2	2	0	2	4	0	2	2	4	2	3	1	3	5	1	5
	Perthshire (1)	(this)	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0
	Roxburghshire (1)	(yir)	0	1	0	0	1	0	1	0	1	0	1	0	1	1	0	1
	Selkirkshrie (1)	(thir)	1	0	0	1	0	0	1	0	1	0	1	0	1	1	0	1
	Wigtownshire (1)	(yir)	0	1	0	0	1	0	1	0	1	0	1	0	1	1	0	1
	Total					7	9	0			14	3				16	1	13

County			most common token	initial						medial				final					
				variation			citation			variation		citation		variation			citation		
				th	y	þ	th	y	þ	a	o	a	o	*	s	se	*	s	se
England	SE	Kent (12)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Surrey (20)	tho, þo	1	0	1	2	0	2	0	2	0	4	2	0	0	4	0	0
		Sussex (14)	þo	0	0	1	0	0	1	0	1	0	1	1	0	0	1	0	0
		Total					2	0	3			0	5				5	0	0
	SW	Berkshire (7)	þo3	0	0	1	0	0	1	0	1	0	1	0	1	0	0	1	0
		Cornwall (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Devon (17)	þay, þei, þo	5	0	4	5	0	7	4	3	6	4	6	1	2	9	1	2
		Dorset (5)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Gloucestershire (40)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Hampshire (17)	þo	0	0	1	0	0	1	0	1	0	1	1	0	0	1	0	0
		Oxfordshire (10)	þos	0	0	1	0	0	1	0	1	0	1	0	1	0	0	1	0
		Somerset (19)	(tho, thos, þo	2	0	1	2	0	1	0	3	0	3	2	1	0	2	1	0
		Wiltshire (19)	tho	2	0	1	2	0	1	0	3	0	5	3	0	0	5	0	0
		Total					9	0	12			6	15				17	4	2
	EM	Bedfordshire (8)	þo	0	0	1	0	0	2	0	1	0	2	1	0	0	2	0	0
		Buckinghamshire (12)	þo	0	0	1	0	0	2	0	1	0	2	1	0	0	2	0	0
		Cambridgeshire (16)	þo	1	1	2	5	1	8	0	4	0	14	4	0	0	14	0	0
		Essex (35)	tho	2	0	2	4	0	3	0	3	0	6	2	1	0	5	1	0
		Hertfordshire (10)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Huntingdonshire (12)	þo	0	1	1	0	1	10	0	2	0	11	2	0	0	11	0	0
		Leicetstershire (27)	tho	3	1	3	13	1	12	0	7	0	26	4	1	1	20	2	2
		Lincolnshire (50)	yo	1	13	7	1	37	15	6	11	8	39	6	6	5	28	9	6
		London (6)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
		Middlesex (11)	(tho)	1	0	0	1	0	0	0	1	0	1	1	0	0	1	0	0
		Norfolk (67)	yo	3	4	4	9	35	15	0	9	0	52	7	1	1	50	1	1
		Northamptonshire (41)	tho	3	0	2	14	0	10	0	5	0	24	4	0	1	22	0	2
		Nottinghamshire (26)	yo	3	7	4	5	18	8	1	12	1	29	6	4	3	17	8	5
		Soke of Peterborough (3)	þoo	0	0	2	0	0	3	0	2	0	3	2	0	0	3	0	0
		Rutland (5)	(tho, thoes, þ	2	0	4	2	0	4	0	6	0	6	3	3	0	3	3	0
		Suffolk (31)	tho	3	2	2	7	3	4	0	4	0	11	4	0	1	11	0	1
		Yorkshire, West Riding (72)	yo	5	10	5	8	50	16	9	11	23	50	11	4	5	44	13	16
		Total					69	146	112			32	276				233	37	33
	WM	**Cheshire (35)**	those	5	4	3	11	8	6	0	11	0	23	4	3	3	9	5	10

	Derbyshire (21)	yoo	2	1	4	2	4	8	0	7	0	14	4	2	1	11	2	1
	Herefordshire (27)	(þo)	0	0	1	0	0	1	0	1	0	1	1	0	0	1	0	0
	Lancashire (58)	those	4	6	4	9	14	5	3	11	4	24	7	4	3	14	7	7
	Shropshire (17)	þo	2	0	1	5	0	6	0	3	0	11	3	0	0	11	0	0
	Staffordshire (32)	þo	2	1	5	4	1	14	0	7	0	18	5	1	1	14	3	1
	Warwickshire (34)	þo	2	0	4	4	0	14	1	5	1	17	4	1	1	15	1	2
	Worcestershire (25)		0	0	1	0	0	1	0	1	0	1	0	0	1	0	0	1
	Total					35	27	55			5	109				75	18	22
N	**Cumberland (32)**	(thase, yos, þ	1	1	1	1	1	1	1	2	1	2	0	1	2	0	1	2
	Durham (18)	(ya, yais, þa)	0	2	1	0	2	1	3	0	3	0	2	1	0	2	1	0
	Northern (15)	ya	0	9	0	0	21	0	6	3	17	4	4	3	2	10	5	6
	Northumberland (19)	(thase, ya)	1	1	0	1	1	0	2	0	2	0	1	0	1	1	0	1
	Westmorland (17)	(yase)	0	1	0	0	1	0	1	0	1	0	0	0	1	0	0	1
	York (5)	þase	0	2	1	0	2	1	3	0	4	0	1	1	1	1	1	2
	Yorkshire (2)	(yase, yo, yo	0	3	0	0	3	0	1	2	1	2	1	1	1	1	1	1
	Yorkshire, North West (4)	yase	0	3	0	0	7	0	3	0	7	0	1	1	1	2	2	3
	Yorkshire, East Riding (23)	yas	1	5	2	1	8	2	6	2	9	2	3	2	3	4	4	3
	Yorkshire, North Riding (30)	ya	2	3	3	2	11	3	4	4	12	4	4	1	3	9	3	4
	Total					5	57	8			57	14				30	18	23
	Unlocalised (3)	(y^o, þ^o)	0	1	1	0	1	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Wales	Carmarthen (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Denbighshire (3)	(þoo)	0	0	1	0	0	1	0	1	0	1	1	0	0	1	0	0
	Monmouth (4)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Montgomery (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total					0	0	1			0	1				1	0	0
Scotland	Ayrshire (2)	(yai, yha)	0	2	0	0	2	0	2	0	2	0	2	0	0	2	0	0
	Berwickshire (1)	(tha, thai)	2	0	0	2	0	0	2	0	2	0	2	0	0	2	0	0
	Dumfriesshire (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	East Lothian (3)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Lanarkshire (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Midlothian (5)	ya	1	1	0	1	2	0	2	0	3	0	2	0	0	3	0	0
	Perthshire (1)	(tha)	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0
	Roxburghshire (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Selkirkshrie (1)	(thar)	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0
	Wigtownshire (1)	N/A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Total					5	4	0			9	0				9	0	0