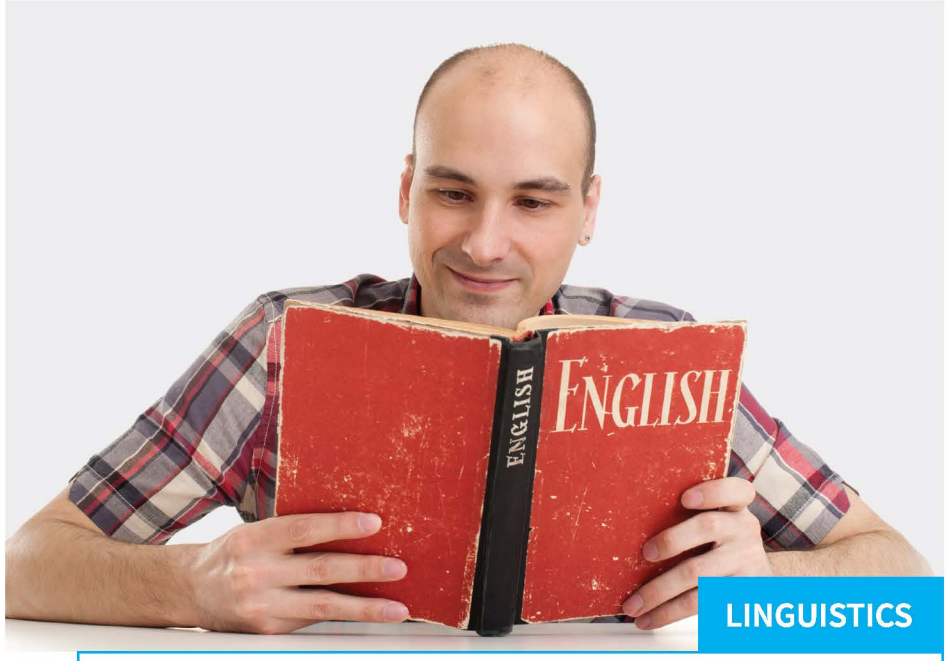
editors
Ewa Waniek-Klimczak
Anna Cichosz

Variability in English across time and space



LINGUISTICS

PHONETICS, DIALECTOLOGY, HISTORICAL LINGUISTICS

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Variability in English across time and space

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

editors
Ewa Waniek-Klimczak
Anna Cichosz

Variability in English across time and space

WYDAWNICTWO
UNIWERSYTETU
ŁÓDZKIEGO

Łódź 2016

Ewa Waniek-Klimczak, Anna Cichosz – University of Łódź, Faculty of Philology
Institute of English Studies, Department of English and Applied Linguistics
90-236 Łódź, 171/173 Pomorska St.

REVIEWER
Włodzimierz Sobkowiak

INITIATING EDITOR
Urszula Dzieciątkowska

TYPESETTING
Munda – Maciej Torz

TECHNICAL EDITOR
Leonora Wojciechowska

COVER DESIGN
Studio 7A

Cover Image: © Depositphotos.com/spaxiax

Published by Łódź University Press
First Edition. W.06965.15.0.K

ISBN 978-83-8088-065-8
e-ISBN 978-83-8088-066-5

https://doi.org/10.18778/8088-065-8

Publisher's sheets 7.2; printing sheets 10.0

Łódź University Press
90-131 Łódź, 8 Lindleya St
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl

tel. (42) 665 58 63

Preface

http://dx.doi.org/10.18778/8088-065-8.01

Variability in language invariably attracts attention of linguists, particularly those for whom the use of language remains the core of linguistic enquiry. Whether viewed from a diachronic or synchronic perspective, the variable use of language provides insights into the formation, evolution and dynamism of the language system. In the case of English, the observed variability seems to be inexhaustible. The language with so many native and non-native varieties, spoken by millions of people as a modern Lingua Franca offers endless possibilities for research into its numerous manifestations. The studies collected in the present volume bear witness to the wealth of problems and approaches that can be used in the search for variability patterns in the development and present-day usage of English. Adopting various methodological approaches, the young researchers from the Gavagai Student Society, whose studies have been collected here, take the challenge of describing the variability in the use of English across space and time, from Middle English to present-day English, used by native and non-native speakers.

In the first contribution Adamczyk, presents an in-depth analysis of two competing spelling variants of dental fricatives in the Northern dialect of Late Middle English. The study shows that the so-called Northern system, in which the choice between the two variants was supposed to reflect difference in voicing, can be traced in the analysed data but it was not consistently applied. The study focuses on function words, because this is the area in which spelling variability has been found, and it is based on a large corpus of Middle English legal documents. Thanks to a well-thought-out methodology and scope, the results of the study are an interesting contribution to the study of Middle English dialectology and spelling.

The next study investigates the same period in the history of English, but the variability examined there is syntactic in nature. In his study, Grabski aims to check whether the choice between single and multiple negation patterns in The Canterbury Tales is – at least partly – related to sociolinguistic factors. The study shows that Chaucer seems to have used negation as marker of the social status of his characters, with highly-educated speakers preferring single negation and speakers with lower status using negative concord relatively more often. This in itself is an interesting observation since single negation is supposed to have developed in the (Early) Modern English period during the process of language standardisation, and multiple negation became generally stigmatised long after the Middle English period.

The study by Matysiak analyses variability in English from a modern perspective, focusing on Polish users of English. The most interesting aspect of the study is that the informants are Polish immigrants living in London, surrounded by people using English natively and communicating in English on a daily basis. Recently, this group has become very numerous and its linguistic behaviours are an extremely interesting area of study. The aim of this investigation is to check to what extent the level of proficiency in English on arrival in the UK and the quality of English input afterwards influence the use of English by Polish speakers. The linguistic variable selected for the purpose of the study is aspiration of voiceless plosives. The study shows that previous language experience is of crucial importance and determines success in the acquisition of L2 pronunciation in the investigated group.

In the next study, Rajtar examines the degree of formulaicity of native and non-native English used by Polish learners. The aim of the investigation is to determine the most frequent two- and three-word sequences from both samples and compare the occurrence of silent pauses within them. The study is based on the British National Corpus and PLEC corpus (PELCRA Learner English Corpus) created at the University of Łódź. The study shows that native and non-native speakers use almost completely different sets of formulaic phrases (the only exception is I don't know) while the distribution of silent pauses is very similar in both samples, which proves that formulaic expressions demonstrate similar behaviour regardless of language (variety).

With the study by Rybińska, the volume moves back to the Middle English period. This contribution focuses on lexical variation and it aims to establish regional differences in the distribution of Late Middle English lexicon. The study is limited to verbs and the material selected for the purpose of the analysis are two distinct versions of Mandeville's Travels: one manuscript was produced in the Northern dialect while the other comes from the south of England. The regional differences

established on the basis of this text are verified by the author on the basis of a larger sample of Northern and Southern Middle English texts. The study shows that the differences in ME dialects are not only phonological or syntactic but also lexical, which is often disregarded in historical studies of this period.

The investigation carried out by Szczytko is also historical, but based on a collection of letters coming from the Late Middle and Early Modern English period. The investigated variable are forms of address, with special attention paid to the choice between two competing second person pronouns thou and you in private correspondence. The aim of the author is to establish which sociolinguistic factors are decisive in the selection of forms of address. Quite surprisingly, no instances of thou have been identified in the analysed sample, which shows that letters are clearly different from drama from the same period, where variability in second person pronouns does exist. The author suggests that this result may be related to the epistolary conventions of the period. The investigation proves that the choice of nominal forms of address was determined by two main factors: social position and family relations.

The last contribution in this volume is a study of phonetic imitation by Zając who analyses the linguistic behaviour of Polish learners of English. The study aims to show to what degree the informants imitate the native or non-native model. The linguistic variable analysed in the study is the duration and quality of vowels. The analysis suggests that phonetic imitation does have an impact on L2 pronunciation because the informants showed a tendency for convergence towards the native English speaker and divergence from the Polish speaker. This result seems important from the point of view of phonetic training and shows that attitude towards native and foreign-accented speech have an impact on L2 pronunciation.

The studies presented in this volume reflect but a fraction of variability present in historical and contemporary English. However, the very fact that they are so varied shows how stimulating the wide range of possible venues to be taken within this area of research can be. This collection of papers bears witness to the effect of cooperation and discussion within a student society environment. We hope that it will be both interesting and stimulating for other young scholars interested in English linguistics. Finally, we hope that our contributors will never stop being curious about language and out of this curiosity they will keep asking questions and will continue to explore various aspects of variability in English and other languages.

Ewa Waniek-Klimczak
Anna Cichosz

Contents

Realisations of the Word-initial Variable (th) in Selected Late Middle English Northern Legal Documents

http://dx.doi.org/10.18778/8088-065-8.02

Michał Adamczyk

University of Lodz

Abstract

This paper is a study in Late Middle English orthography and its relationship with the phonological system. The study was conducted on a representative sample of legal documents from all core northern counties. The analysis concerned the variable (th) that stands for a systemic distinction between /ð/ and /θ/ by means of two graphemes: <þ/y> and <th> in the north of England. The results of the quantitative analysis confirmed the existence of the Northern System, however, in its decline. The analysis of discrete grammatical words proved that *the*, *that* and *they* were the most conservative words showing a significantly higher preference for <þ/y> than the remaining grammatical words examined in the present study.

1. Preliminary remarks

The study was conducted on a representative sample of texts from the *Middle English Grammar Corpus* (Stenroos et al. 2011) (later MEG-C). This recently developed corpus has proven to be a commendable source of data in the area of historical dialectology, and due to its accessibility it enabled the author of the present paper to conduct a modest study of the variable (th). According to Stenroos (2004: 257), the term variable (th) combines all Middle English spelling conventions of representing a dental fricative in writing, with the exclusion of those cases in which a graph corresponds to a word-medial /ð/ developed from earlier /d/.

Word-initially, the variable (th) has two possible spelling variants found in the Middle English period: <þ/y> and <th>. On the basis of this orthographic variance, a system distinguishing between the voiced dental fricative and its voiceless counterpart is thought to have developed in the north of England. The presence of the systemic distinction known as the Northern System is worth investigating on a representative sample of documents from all core northern counties and areas of transition between dialects: Lancashire and West Riding of Yorkshire. For this reason, the aim of the paper was to investigate the presumed relationship between the spelling variants and the voicing of word-initial fricative on a representative sample of 126 Late Middle English legal documents from Cumberland, Durham, Lancashire, Northumberland, Westmorland and Yorkshire. Legal documents used in the study were selected from the subset of the MEG-C. Documents were searched for grammatical words with word-initial /ð/ and lexical words with /θ/ in the onset. The results of the quantitative analysis of two variants of the variable (th) were subjected to statistical analysis, plotted on a map in order to account for a possible pattern of spatial distribution, arranged chronologically and divided into particular words in order to gain some insight into differences between separate grammatical words distinguished in the design of the study. Multidimensional analysis of the data was then used to support the view suggesting the existence of the separate Northern System, which distinguishes between /ð/ and /θ/ using two variants of the variable (th): <þ/y> and <th>.

2. Variable (th)

The Middle English phonetic inventory included two phonemic dental fricatives: /θ/ and /ð/, which in most cases, descended directly from the Old English /θ/. However, in the Old English phonological system, the distinction between voiced and voiceless fricatives was allophonic, hence, the allophones appeared in complementary distribution with the voiced sound occurring word-medially between vowels or voiced consonants and its voiceless reflex in the remaining positions: initially, finally and when double medially. The latter developments occurring in the Middle English period resulted in the formation of pairs of fricatives distributed contrastively.

Although dental fricatives in Middle English developed the phonemic contrast of voice similarly to other pairs of fricatives (aside from /ʃ/ and /x/), the reason be-

hind the change differed significantly. In pairs /s/ : /z/ and /f/ : /v/ the development of the phonemic contrast of voice around the year 1250 occurred as a result of the large amount of borrowings with word-initial /z/ and /v/ from French, degemination of word-medial /s:/ and /f:/, the loss of final /ə/, and finally voicing of word-initial fricatives in many southern dialects of Old English (Lass 1999: 59). Though in the case of /θ/ : /ð/, voicing of word-initial dental fricative appeared in low sentence stress words such as deictic expressions *the*, *this*, *that*, *these*, *there*, *then* and some conjunctions like *through*. A parallel process altered some other weakly stressed words: *is*, *of* and *was* by voicing word-final consonants (ibid.: 59–60). Because of the difference in the development of /θ/ : /ð/ compared to other pairs of fricatives, the phonemic distinction of voice between the two fricatives in the onset, however at no time discriminated totally, was reduced to a limited number of minimal pairs such as *thy* and *thigh*.

Old English, having [z], [v] and [ð] only as allophones of /s/, /f/ and /θ/, did not distinguish between voiced and voiceless sounds by the means of different graphs; <s> was used for both [s z], <f> for [f v] and <þ ð> for [θ ð]. The emergence of the final two graphs used interchangeably for a single dental fricative, however, is worth exploring for the purpose of the study. According to Quirk & Wrenn (1957: 8), in the earliest surviving Old English texts, bilateral <th> borrowed from Irish (Hogg 1992: 77) was used for [θ] and [ð]. In texts from the later eighth century, one may also encounter <d> used for a dental fricative. Some scholars claim that the use of <d>, similarly to <th>, may be a result of a borrowing from Irish scribal tradition since in Irish <d> was sometimes used to signify a voiced fricative (Quirk & Wrenn 1957: 8). As Christianity became firmly established, the previously used graphs were replaced by <þ>. It is argued that the signs of runic alphabet, especially <þ>, started to be widely employed at the time when the elements of the pagan Germanic culture stopped being viewed as a potential threat to the position of the Christian Church (ibid.). The use of the three graphs mentioned so far might be clearly seen in different spellings of the word *thought* in *Cædmon's Hymn*: *modgidanc* (Moore Bede), *modgithanc* (Leningrad Bede) and *modgeþanc* (West Saxon version, first half of the tenth century) (Hogg 1992: 76–77). By the beginning of the ninth century, the graph <þ> known by its runic mnemonic name *thorn* from *futhorc* was being used alongside with the new graph <ð>. Although it is considered uncertain (ibid.: 75), some scholars claim that <ð> is once again a borrowing from the Irish-Latin alphabet formed by drawing a line through the upper part of <d> (Quirk & Wrenn 1957: 8). The name of the graph, *eth* or *edh* is thought to

be a nineteenth-century coinage originating in the name of the corresponding Modern Icelandic letter *eð* (ibid.: 8; Fulk 2012: 23). Originally, in the Old English period, the graph was known as ðæt (Hogg 1992: 75). Although the interchange between <þ> and <ð> in some Old English texts seems to be regular, for example, in some of the most carefully written MSS of Ælfric where <þ> is used word-initially and <ð> word-medially or word-finally (Quirk & Wrenn 1957: 9; Upward & Davidson 2011: 56), in a broader perspective, these two graphs appeared in free orthographic variation.

While <ð> had been lost in the beginning of the fourteenth century (Jensen 2012), the graph <þ> continued to be widely used to represent both /θ/ and /ð/ throughout the Middle English period. However well-established the graph <þ> was, reintroduced <th> started to appear in writing from the beginning of the twelfth century onwards. <th> took over the role of <þ> altogether by the end of the fifteenth century (Lass 1999: 36). The modified version of <þ>, which virtually merged with <y>, was, however, still in use even in Early Modern English in forms such as *ye* or abbreviated y^t.

According to Jensen (2012), the usual explanation for the loss of the graph <þ> that is the introduction of print may, actually, prove to be quite inadequate. It is argued that the replacement of <þ> with <th> was a gradual process that had begun long before the arrival of the first printing press to England. The emergence of <th> in the twelfth century in the southern part of England and its steady grow in use argued by Lass (1999: 36) seem to confirm this view. Furthermore, it is claimed that printing facilities and different scribal practices existed in isolation even as late as in the end of the Early Middle English period (Scragg 1974). Finally, Jensen (2012) argues that first of all *thorn* was included in a number of types used in England, and secondly the graph <y> may have been easily employed by the printing industry in order to avoid the introduction of <th>.

In the course of time, <þ> and <th> were established as the two Late Middle English variants of the variable (th). In the South, the two variants appeared in free variation with <th> getting the upper hand over <þ>. In the North, however, the distribution was constrained by the systemic factors addressed later in this section. The distribution of <þ> and <th> was further confused by the fact that <þ> and <y> merged into a single *y*-shaped letter in many scribal hands well before the end of the fourteenth century (Fulk 2012: 23). As it was suggested (Benskin 1982: 21 ff), the merger originated in textura scripts and was graphic in nature. What is more, the merger may have been subjected to a spatial distribution,

with the coalesced <y> occurring, especially, in the North and in large parts of Northeast Midlands and East Anglia (Fulk 2012: 23). As a result of the change, the northern orthographic system contained one graph less when compared to its southern equivalent.

Apart from the sole number of letters available, some scholars noted that another systemic dissimilarity differentiated the two systems. Stenroos (2004: 267) argues that some Middle English texts exhibit a tendency to distinguish between voiced and voiceless fricatives by means of different graphemes. This tendency was addressed by Benskin (1977: 506–507) in a noteworthy footnote:

> There thus arises a system whereby (1) words like *think*, *through*, *thousand* are spelled *th*-, but (2) words like *they*, *them*, *there* are spelled *þ*- or *y*-. The use of *þ* (or *y* for *þ*) is hence phonetically conditioned in the orthographies of a great many scribes, an observation which seems to have eluded most scholars.

Benskin's (1977) perspective on a diachronic change of the system assumes four consecutive stages in the spread of <th> in the varieties of the Northern dialect: (1) final position occupied by a voiceless fricative only, (2) initial position when occupied by a voiceless fricative, (3) in word-medial position, (4) word-initially when voiced. Stenroos (2004: 267) points out that "between the second and third stages, there arises a system where the voiceless dental fricative is spelled *th* and the voiced one þ or *y*." Hence, one may assume that the system no longer contained a set of graphs, but rather two separate graphemes: <th> and <þ>, through which the contrast of voice between, for example, *thin* 'thin' and þin 'thine' may have been orthographically maintained. However, as it was noted by Jensen (2012), due to the unique development of dental fricatives, with the voiced dental fricative present usually in grammatical words and the voiceless one in lexical words, the distinction may be, as well, interpreted as purely lexical.

With the implementation of standard writing conventions, <th> grew more common; however, it did not replace <þ> and <y> instantly. As it was stated (Benskin qtd. in Jensen 2012), scribes from the north of England, while adopting a new standard, had to, first of all, reincorporate the graph lost in the merger with <y>: widely used <y> had to be replaced with the earlier <þ> when it represented a consonant. Secondly, they had to abandon the distinction in writing made between voiced and voiceless dental fricatives. In fact, whereas in the South the change from <þ> to <th> was a matter of simple graphic replacement, in the North, it

equalled a systemic change, in which the loss of one of two graphemes corresponding separately to /θ/ or /ð/ resulted in the loss of phonologically conditioned division. It can be easily viewed as a merger.

Although the language of legal documents is often thought to be the one that is the most receptive to the standard, it might be interesting to gain some insight into some other possible factors, apart from genre, influencing the reception of the standardised forms and the retention of the original system with word-initial <th> being used for /θ/ and initial <þ/y> used for /ð/. Geographical distribution, frequencies in particular documents or lexical diffusion may be particularly interesting to explore from the point of view of historical dialectology.

3. Sources

The sample of Late Middle English texts used in the study consisted of 126 documents from the MEG-C included in the *Linguistic Atlas of Late Medieval English* (Benskin et al. 2013) (later LALME). The part of the MEG-C consisted of 76055 words. The texts were located in the counties in the area of the northern dialect: Northumberland, Durham, Cumberland, Westmorland, Yorkshire and Lancashire. The selection of texts from every county was a prerequisite for providing an all-embracing analysis of spatial diffusion of the variable (th). As for the genre of the texts selected, these were legal documents only, primarily because the study aimed at the description of the variable (th) only in this particular genre. Secondly, the analysis of differences between genres, as it was done in the research conducted by Jensen (2012), in which legal documents were set against religious prose texts, would be hardly possible for counties such as Cumberland since the MEG-C provides only documents, making this type of research anything but congruous. In the Appendix, all legal documents used in the study are listed in accordance with the codes used in the Catalogue of Sources - version 2011.1 (Stenroos 2011) accompanying the MEG-C. Figure 1 on the following page shows localisation of the legal documents, however, without initial *L* and *0*. The MEG-C codes consist of a capital *L* followed by the LALME Linguistic Profile code made into a four-digit code by adding initial zeros as necessary, for example, L0147 used in the MEG-C corresponds to the LALME LP147. Sometimes when a complex LP had been split into smaller units by the authors of the MEG-C, for the purpose of this study, it was merged back into an original entity. Similarly, when more than one legal document was located in a single locality

from the LALME, they were structured into a single entity in the part of the research devoted to spatial distribution of the variable (th). The MEG-C included few legal documents that were not placed on maps during the process of compilation of the LALME. The design of the present study required the texts to be tied to a particular locality, be it real space or a localisation based on an assemblage of linguistic features. Hence, all the legal documents that had been used in the compilation of the LALME but had not been placed on maps were excluded from the scope of the research. Finally, in regard to spatial distribution, L1348 was the only document out of the whole material which has been moved slightly southwest on the basis of Jensen's study (2012).

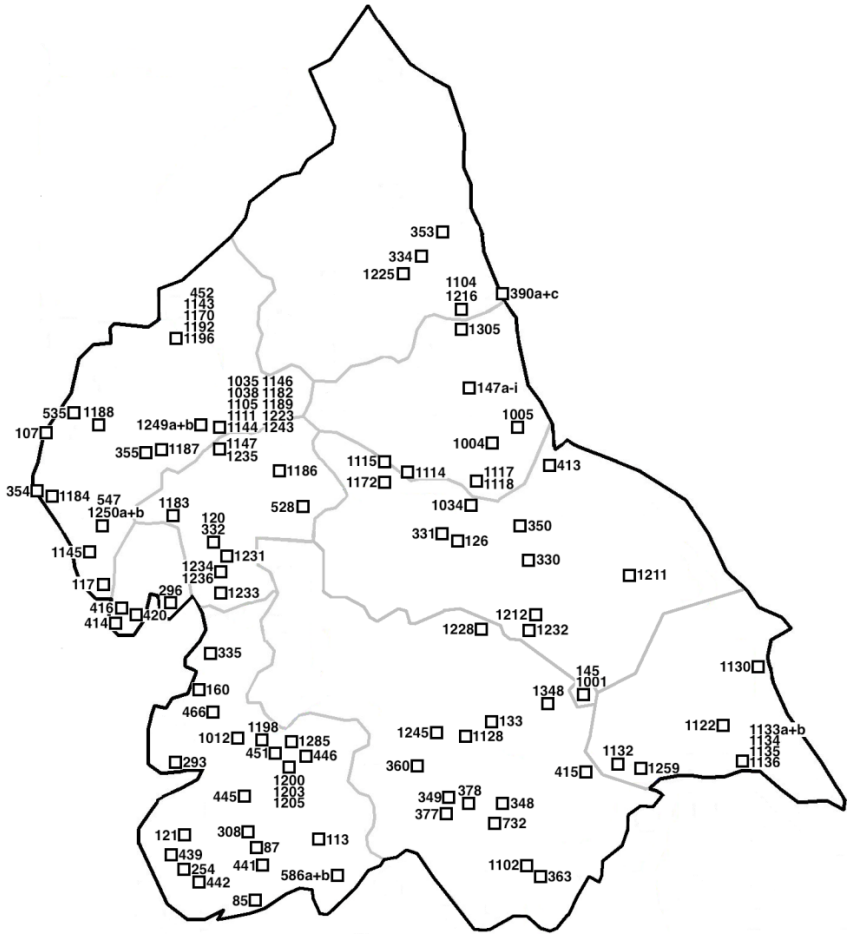


Fig. 1. Spatial distribution of the legal documents

On the basis of the explicitly dated documents, the material represents a time span of one hundred and forty-two years, with the earliest document form the year 1363 and the oldest from 1505. Yet, a substantial amount of documents used in the study dates back to the fifteenth century, with only a small number of documents from the second half of the fourteenth century or the first decade of the sixteenth century. Texts with only approximate dating were also included in the study. They are dated accordingly: four documents dated to the middle of the fifteenth century, two documents dated to the first half of the fifteenth century, one document dated to the second half of the fifteenth century, one document dated to the end fifteenth or the beginning of the sixteenth century, two documents dated imprecisely to the time between 1438–55 and 1472–83, one document with the unknown dating. The documents mentioned above were not incorporated in the diachronic analysis, which relied on precisely dated documents. Nonetheless, the design of the diachronic analysis used in the present study allowed documents, such as L0147g dated 1446–47, with a slightly imprecise dating to be included. Both explicit and approximate dates of document composition are listed in the Appendix in the right column.

4. Procedure

The process of transcription of the legal documents into a machine-readable format was thoroughly described in the manual accompanying the MEG-C (Stenroos & Mäkinen 2011). One can easily access the manual through the website of the *Middle English Grammar Project*. It is, however, worth mentioning that thanks to the authors of the corpus the annotation employed in their project enabled an in-depth analysis of late Middle English spelling conventions. As for the sampling method, longer pieces of writing, for example religious prose or verse, were included in the corpus in tranches of 3,000 words; legal documents used in the present study entered the corpus in their entirety. Hence, the material used in the study comprised of entire documents only, making it possible to reject any possible disparity with the original manuscripts.

In accordance with the theory outlined in the section devoted to the variable (th), data extracted from the part of the MEG-C were structured in relation to the word type further divided into grammatical and lexical words. Grammatical words like *the*, *that* or *they* were treated as those containing word-initial voiced dental fricative, whereas lexical words, for example *thing* or *think,* as containing

voiceless dental fricative initially. Furthermore, due to the limited occurrence in the corpus, lexical words were collected as an open category leaving out such obvious mistakes on the part of the scribe as, for instance, yAPPur[TENaNTz (L0147g), a merger of 'the' and 'appurtenance' and yEER~ (L0147g) 'þeer~ > year'. Conversely, grammatical words were clearly divided into eleven subcategories corresponding to particular grammatical forms: *this*, *these*, *those*, *there*, *they*, *them*, *their*, *theirs*, *then*, *that* and *the*. Complex conjunctive adverbs, for example, *therefore* or unusual YANEWITH, along with adverbs such as YIDERWARD (L0586a), were excluded from the scope of the research and were not included in any type of quantitative analysis employed in the present study. The grammatical word *through* was also excluded. In terms of adverbs, it is difficult to decide whether they should be treated as grammatical or lexical words. *Through*, however, appeared in numbers too small to allow the incorporation of the word in the analysis.

Data were extracted from the MEG-C using AntConc 3.2.4m, freeware software designed for the purpose of corpus linguistic analysis. Due to a large variety of spellings available for each word, regular expressions were used to simplify the process of data extraction and to yield more precise results. Although it was possible to create a single regular expression for the most of the grammatical words provided above, in the case of *there* and *their*, whose spellings overlap in some cases, each instance found in the corpus had to be analysed separately paying special attention to their concordances. Treating lexical words as an open category required a slightly different approach, that is using a regular expression extracting all words with word-initial <th>, <y> or <þ> and selecting those which matched the word type and the remaining requirements provided above. Upon the completion of data extraction, results were quantified as numerical values separately for lexical and grammatical word; the latter type was further divided into eleven subcategories. Numerical figures obtained in the course of the study were then converted into a standardised numerical value: frequency per one thousand words often used in research conducted in the area of corpus linguistics. Converted numbers made quantitative data significantly more comparable by avoiding the negative effect of sample size on the results of the study.

Data extracted from the MEG-C and structured according to the aforementioned requirements were then checked for statistical significance using a Chi-square statistical test. After combining two possible spellings: <þ/y> and <th> with two word types: grammatical and lexical, a statistical test was applied to both the overall results from six combined counties and each of the counties separately.

It was assumed that lexical words with a word-initial voiceless dental fricative should show a significantly higher proportion of <th>, conversely, grammatical words beginning with a voiced dental fricative should exhibit a tendency for <þ> or <y>. The results of a Chi square test were presented in a form of a table in the following section devoted to the presentation of the results. *P*-values acquired in the statistical analysis were used to accept or disprove the null hypothesis stating that there is no relationship between the word type and the variant of the variable (th) for p-value < 0.05. If a Chi square test confirmed the statistical significance of the relationship between the variables used in the present study, it would allow to reject the null hypothesis and to validate the hypothesis stating that there is a relationship between the word type and the variant of the variable (th). Along with the table presenting the results of the statistical test, a graph showing the proportion of variants of the variable (th) for both word types was included.

Following the statistical analysis and presentation of the overall results, a number of other methods were used in order to identify the pattern of distribution of the variable (th) in the selected collection of legal documents from the MEG-C. Since the distribution of the variable (th) in the open category of lexical words in the present study proved to be fully homogeneous, with the variant <th> occurring in every instance extracted from the corpus, the remaining methods were not applied to this word type. It might be assumed that the lack of variation in lexical words reflected the state of the graphemic system described by Benskin (1977) and commented by Stenroos (2004: 267), where a word-initial voiceless dental fricative is spelled <th> and the voiced one <þ> or <y>. The distribution of the variable (th) in grammatical words, however, appeared to be visibly more varied. The following methods were employed to account for this variability. Firstly, is accordance with Jensen (2012) claiming that

> Linguistic variation in Middle English texts is [thought to be] most commonly studied in terms of geography, and, [for this reason], regional patterns must be expected to account for much of the variation during the Late Middle English period. At the same time, variables other than geography must be assumed to have contributed to synchronic variation.

A spatial dispersion of the variants of the variable (th) was shown in the form of a map indicating a dominant spelling of word-initial fricative for each legal document bound to a particular locality, as they were shown in Figure 1. Having done

that, the chronological distribution of <þ/y> and <th> spellings was provided in a form of a bar plot showing changes in frequency of two variants form the year 1363 to 1505. To avoid unnecessary obscurity in the form of a bar plot, documents were grouped into intervals of one or two decades. The division into decades relied upon the number of documents available from each period to maintain the chronology, for instance, because the corpus did not contain documents from the period between 1460 and 1469, but a number of documents were dated between 1450-59, the interval used stretched over two decades, 1450-69 instead of just one followed by a lacuna. Thirdly, the frequency of two variants was set against the number of legal documents to check the distribution of frequencies of two variants across the documents. For instance, to validate the possibility of <th> being spread across a large amount of documents in relatively low frequency, whereas <þ> or <y> appearing in a significantly larger frequency in a comparable number of documents. The results of this part of the analysis were presented in a form of two histograms showing the distribution of both variants separately, which in the end were merged into a single overlapping histogram identifying differences between two patterns of distribution. Finally, a frequency of each variant was presented separately for each of the eleven subcategories of grammatical words in order to discern a possible pattern of distribution of the variable (th) across different words. The results of lexical distribution were shown in a form of eleven separate bar plots.

5. Results

Data extracted from the MEG-C and converted into a standardised value of a frequency per 1000 words rounded to units are shown in Table 1. The vertical header provides labels for two word types combined with the variants of the variable (th). Labels for the counties selected for the purpose of the study are shown in the horizontal header.

As was mentioned earlier, lexical words were fully homogenous in terms of realisation of the variable (th), and for this reason, one might assume that the state reflected in the data, with word-initial voiceless dental fricatives spelled as <th>, may correspond to the state of the system described by Benskin (1977) and specified by Stenroos (2004: 267). However, grammatical words were much more varied with respect to the occurrence of the variants of the variable (th). The entirety of the data, provided in the rightmost column, indicates that although the assumed variant <þ/y> was found dominant, <th> spelling in grammatical

words also appeared in numbers suggesting the interplay of other factors in the distribution of the variable (th) in this word type. Furthermore, the frequency of variants varied across the counties and ridings. For instance, divisions such as Lancashire, Cumberland, East Riding of Yorkshire, West Riding of Yorkshire and the City of York shown <þ/y> as the dominant spelling of the word-initial variant. On the contrary, Northumberland, Durham, Westmorland and Northern Riding of Yorkshire exhibited a tendency for the initial <th> spelling in grammatical words.

Table 1. Frequencies of variants of the variable (th) matched up with word types

	Nhb	Dhm	Lancs	Cumb	Wml	Ery	Nry	Wry	York	Total
Gram. <þ/y>	47	49	95	99	55	81	48	79	137	80
Gram. <th>	83	51	49	30	76	21	72	56	6	50
Lexical <þ/y>	0	0	0	0	0	0	0	0	0	0
Lexical <th>	3	8	9	8	8	3	12	7	8	8

Table 2 shows the results of a Chi-square test applied to the entirety of the data as well as particular counties in order to confirm statistical significance of the data extracted in the present study. Results were treated as statistically significant when p-value < 0.05. Values were rounded to the thousandth place.

Table 2. Results of Chi-square test applied to frequencies per 1000 words

	Nhb	Dhm	Lancs	Cumb	Wml	Ery	Nry	Wry	York	Total
P-value	**0.493**	0.02	0.001	0.001	0.047	0.011	0.015	0.008	0.001	**0.002**

P-value for the entirety of the data extracted from the corpus proved to be statistically significant. In terms of particular counties, however, Northumberland was the only one for which the results were statistically insignificant. The failure of statistical test may suggest that the tendency observed in this particular part of the data proved the lack of relationship between the word type and spelling, in particular, grammatical word type bound with <þ/y> spelling and lexical word type with <th> spelling. At the same time, by looking at the values used in the statistical analysis, one might claim that a significantly lower frequency of lexical

words per 1000 words in terms of Northumberland may have influenced the result of ta Chi-square test. Nevertheless, on the basis of the result of the test performed on the totality of the data, with *p*-value = 0.002, the null hypothesis stating that there is no relationship between the word type and the variant of the variable (th) may be rejected. At the same time, the hypothesis stating that the relationship between the word type and the variant of the variable (th) may be accepted. It is worth to bear in mind that the results of the statistical analysis conformed to previous studies related to the distribution of the variable (th) in the north of England (Jensen 2012; Stenroos 2004).

Following the order provided in the section devoted to the procedure used in the present study, Figure 2 shows the spatial distribution of the variable (th) in grammatical words. Localities with the dominant <þ/y> variant of the variable (th) are marked with a triangular shape and localities in which <th> variant appeared as dominant are marked with a square.

Looking at the map provided, one may be relatively certain that <þ/y> proved to be a dominant variant in grammatical words. Yet, it would seem unreasonable to treat spatial distribution as homogenous. Cumberland and East Riding of Yorkshire indicate a strong preference for <þ/y> in grammatical words manifested in an unvaried distribution of this variant across localities. Two remaining ridings, namely, West and North Riding of Yorkshire seem to go in line with the dominance of <þ/y>, however, with slight variance also present. L0363, L0348, L0415, L0133 and L1232 form a line stretching across the eastern part of West Riding of Yorkshire and ending in L1232 right after the border of North Riding of Yorkshire. Two pockets, the first one extending over West Westmorland and Northwest Lancashire, and the second forming a linear shape at the border of Durham with North Riding of Yorkshire, display a preference for <th> in grammatical words.

Northumberland and, even more, Lancashire provide an indiscernible pattern of spatial distribution. In the case of Northumberland, it may be due to insufficient amount of localities, whereas Lancashire may exhibit a good example of transition area between the northern system preferring <þ/y> initially in grammatical words and the encroaching standardised spelling conventions opting for <th>. Although slightly varied, <þ/y> seems to be a dominant variant of the variable (th) in terms of spatial distribution. One should not forget that the distribution shown on the map was based on the dominance of one variant over the other. In most of the localities both <þ/y> and <th> appeared, however, with varying frequency.

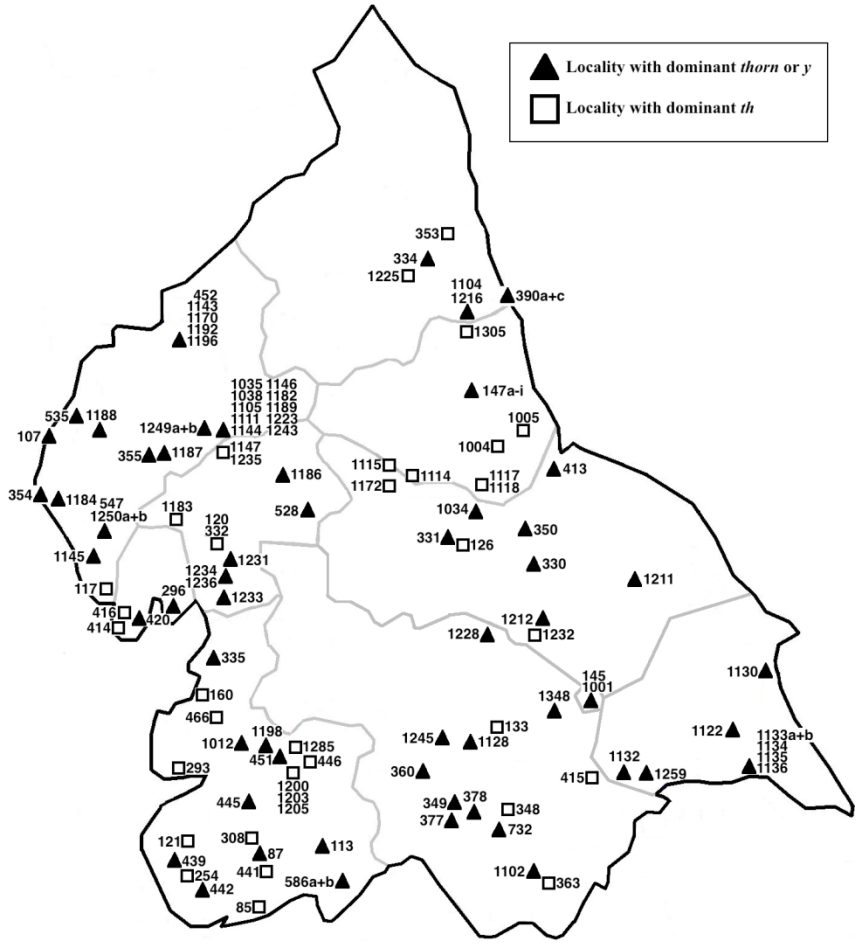


Fig. 2. Spatial distribution of the variants of the variable (th) in grammatical words

The chronological distribution of both variants of the variable (th) in grammatical words is shown in Figure 3. The distribution is based on documents dated precisely enough to match the intervals used in the analysis. The variants are indicated by two colours: <þ/y> by dark grey, <th> by light grey. Frequencies of both variants per 1000 words are presented for each interval. As shown in the figure, <þ/y> seems to be a more frequent variant throughout the entire period of one hundred and forty-two years from 1363 to 1505. Two deviating intervals, 1400–19 and 1440–49, showing a higher proportion of <th> may be considered insufficient to reject the dominance of the presumed variant in Late Middle English north-

ern legal documents. Yet, one might observe that there is a higher proportion of <th> in relation to <þ/y> in the period from 1430 to 1469. This may be treated as a tentative indication of the new spelling convention starting to be implemented in the documents. Still, one might argue that the retreat from this tendency seen in the following intervals disproves the tentative indication in favour of <þ/y> as a dominant variant throughout the period.

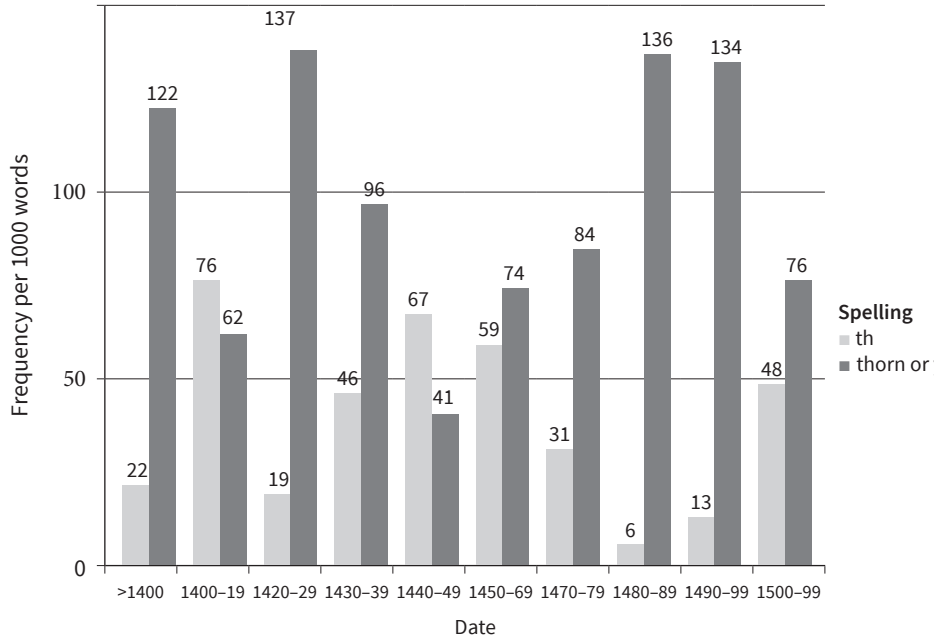


Fig. 3. Chronological distribution of the variants of the variable (th) in grammatical words

Chronological and spatial distributions seem to conform to the view that the variant <þ/y> was the dominant one in grammatical words. Having looked at two types of distribution, it may prove valuable to give some attention to the actual frequency of variants spread across the legal documents used in the study. Figure 4 shown on the following page comprises of two separate histograms displaying the number of documents according to the frequency of one of the variants in grammatical words. Vertical axis indicates the number of documents, while on the horizontal axis, growing frequencies per one thousand words are given. Histograms were set against each other in order to further check the predominance of <þ/y> in grammatical words.

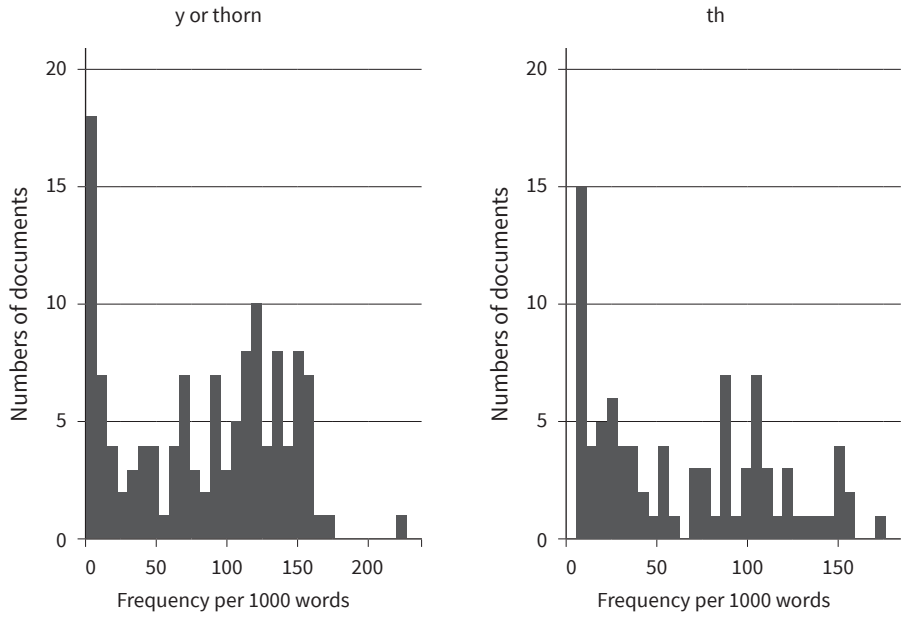


Fig. 4. Frequencies of the variants of the variable (th) in legal documents

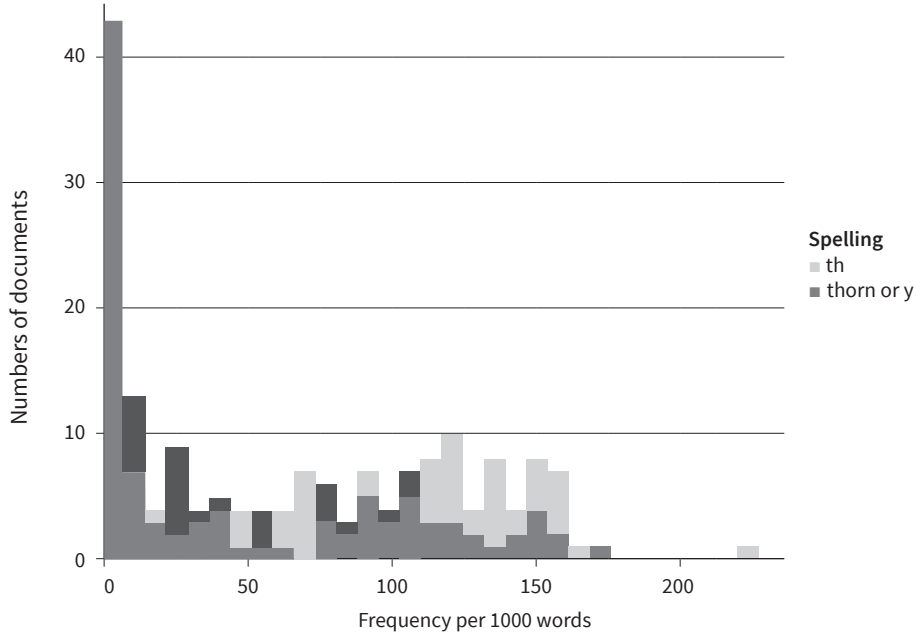


Fig. 5. Overlapping histogram of the variants of the variable (th) in legal documents

Although the variant <þ/y> of the variable (th) seems to be present in legal documents in high frequency, which is indicated in the left-hand side histogram, the <th> variant has a tendency to appear in low numbers over a large amount of documents. Almost sixty documents contain small amount of <th> or even no <th> whatsoever. In the case of <þ/y>, this number is lower by almost a half. A significant difference in the distribution of the variants across the legal documents may be considered as another sign of the preference for <þ/y> word-initially in grammatical words. The differences between the distributions of the two variants can be clearly demonstrated in a form of an overlapping histogram.

Despite the fact that the analysis of the data presented so far may seem to prove the dominance of the variable (th) realised as the variant <þ/y> in grammatical words and <th> as the only realisation of the variable in lexical words, quantitative analysis of separate grammatical words sheds new light on the distribution of the variants. Figures 6a and 6b showing the lexical distribution are presented on the following pages. Two figures comprise of eleven bar plots for eleven grammatical words showing frequencies of both variants of the variable (th) for each word separately. As shown in the legend, <th> is indicated with black and <þ/y> with light grey. For some grammatical words, realisations found in the legal documents were distributed relatively evenly between the two variants; this is the case of *this*, *these*, *them*, *their* and *theirs*. For *there* and *then* data show a higher proportion of <þ/y>, which seems to go in line with the results presented above. Because of a limited occurrence, it would be difficult to claim that *those* is the only grammatical word with <þ/y> only. For this reason, *those* was not considered a convincing piece of evidence for the preferred use of <þ/y> in grammatical words. Yet, *that*, *they* and *the* show an extraordinary preference for <þ/y> word-initially in grammatical words. The distribution of the three words might suggest that the dominance of <þ/y> is constrained lexically.

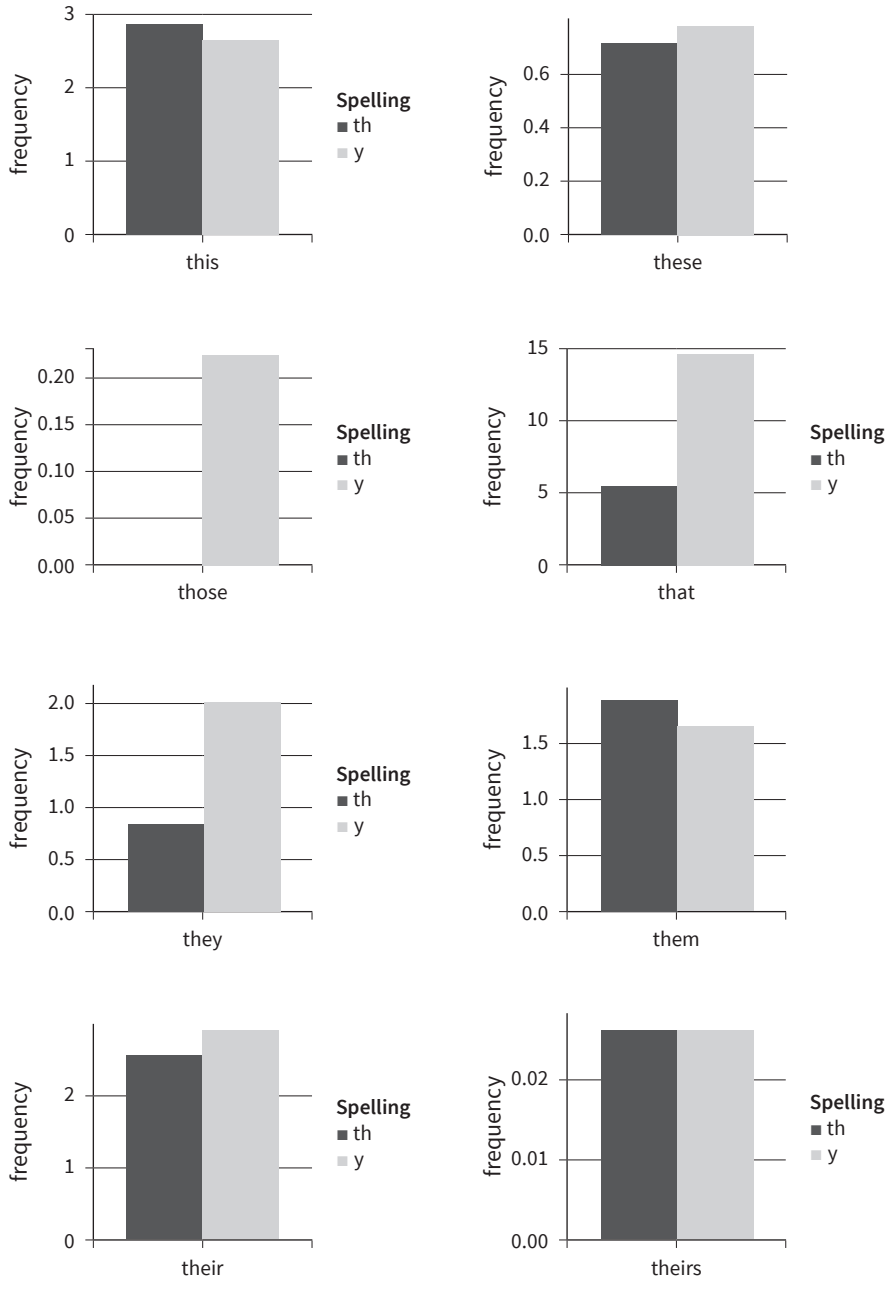


Fig. 6a. Lexical distribution of the variants of the variable (th) for *this*, *these*, *those, that*, *they*, *them*, *their*, *theirs*

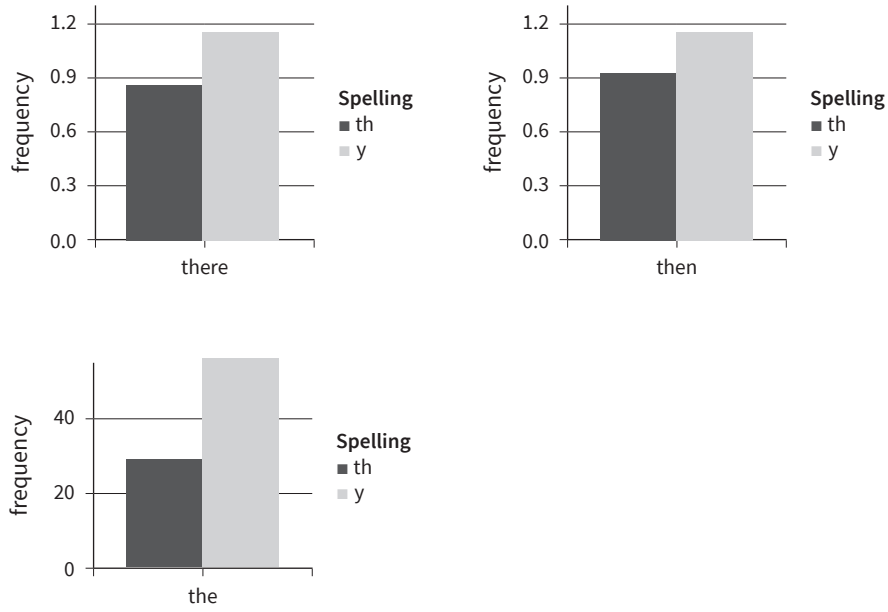


Fig. 6b. Lexical distribution of the variants of the variable (th) for *there*, *then*, *the*

6. Discussion

The Material analysed in the present study seems to indicate a strong preference for the use of <th> in lexical words and <þ> or <y> in grammatical words. This was confirmed by the statistical analysis employed in the study. However indicative the test may be, it also became evident that <th> appeared in a considerable amounts in grammatical words. Because one of the aims of the paper was to prove or refute the existence of the Northern System, every disparity with the systemic distinction between the voiced and voiceless dental fricative by means of different graphemes has to be addressed by analysing possible factors influencing the variance in the use of the aforementioned letters. Since the distribution of the variable (th) in lexical words was homogenous and in line with the distinction present in the Northern System, the analysis was focused on grammatical words selected for the purpose of the study.

While studying the pattern of spatial distribution of <th> and <þ/y> in grammatical words, it was found that <þ/y> proved to be the dominant variant. Only

a small number of localities favoured <th>. A number of them appeared in isolated linguistic pockets in Westmorland, Durham, Northumberland, Lancashire, North Riding of Yorkshire and West Riding of Yorkshire. The pattern behind these enclaves proved to be indiscernible. These localities may be worth exploring in terms of a combination of standardised features, for instance, present participle *-ing* in order to verify a possible existence of pockets showing preference for standardised written English in the fifteenth century northern England. Lancashire, however, stood as a good example of a transitional area between two dialects or two spelling conventions. It can be easily observed in a mixed pattern of localities favouring one or the other variant. Some areas were not covered in the present study. As it was stated by Jensen (2012), this may be due to two options: the reflection of demography or accidental survival of the legal documents. Furthermore, the upland character of the area known as the Yorkshire Dales may have had its impact on demography and, consequently, the survival of the legal documents.

Similarly to the spatial distribution, <þ/y> proved to be dominant throughout the entire period of one hundred and forty-two years from 1363 to 1505. Yet, the assumed dominance of one variant was not constant at all times. As was mentioned above, the period from 1430 to 1469 showed a higher proportion of <th> in comparison to the remaining intervals. It might be argued that the increase in the use of this variant may occur due to the slow incorporation of the standardised spelling conventions. Although followed by the withdrawal from this tendency, some may argue that it can be considered an indication of the standard that was to come. Conversely, it may be claimed that the retreat that followed the increase in the use of <th> may disprove the possible implementation of the standard English spelling conventions. Furthermore, it might be stated that the relatively stable dominance of <þ/y>, with <th> appearing in small quantity from the year 1363 to 1505, may reflect an in-between stage after the third Benskin's (1977) stage assuming the use of <th> word-medially and the fourth one employing <th> word-initially. Despite that, in order to arrive at the actual sequence of <th> implementation, the material used in the study should incorporate texts arranged chronologically for a longer period of time. Still, as far as one hundred and forty-two years covered in the present study are concerned, one can be relatively certain that the <þ/y> variant was dominant in grammatical words, but the state recorded in the material did not reflect the stage in which all voiced dental fricatives were spelt <þ/y>.

As far as the frequency of variants spread across the legal documents is concerned, the <þ/y> variant of the variable (th) once again proved to be dominant. A significant difference between the patterns of distribution of <þ/y> and <th>, with the variant <þ/y> appearing in larger frequencies in the legal documents, may result from the preference for <þ/y> in grammatical words over <th>. Yet, similarly to spatial and chronological distribution, one can observe that the presence of <th> in some instances may suggest that in the fifteenth century the systemic use of the two graphemes might have already been disrupted by the introduction of the standardised spelling conventions using <th> in all positions. The frequency of occurrence of the two variants may also serve as a further specification of the spatial distribution, which in the present study was focused only on the dominant variant for each locality. It clearly shows that for a large number of localities, along with the dominant variant, there were also some instances of the secondary variant. Although the situation in which <th> appears in small numbers distributed evenly across the legal documents may indicate the encroaching standard, the distribution preferring <þ/y> over <th> in grammatical words may be used as strong evidence for the existence of the Northern System.

It is possible that the distribution of the two variants may be conditioned by another factor. In fact, lexical distribution may play a vital role in the explanation of the patterns behind the occurrence of <þ/y> and <th>. As was shown in the previous section devoted to the presentation of the results, three grammatical words: *that*, *they* and *the* exhibited by far the greatest tendency for the use of <þ/y> word-initially. The remaining ones showed significantly more variance in this respect by using both variants in almost equal numbers. On the basis of the analysis, it might be argued that the three grammatical words, in the present material, can be treated as the most "conservative" lexical units retaining the assumed northern spelling convention, while the rest would exhibit more readiness for <th>. The fact that *the* and *that* proved to be most likely to retain <þ/y> may originate in them being the most often used determiners in Middle English. In the case of *that*, the tendency may be further strengthened by the fact that it was a commonly used conjunction. The predominance of the variant <þ/y> word-initially in *the* and *that* may have been caused by the fact that medieval scribes were much more punctilious in the use of the said variant in the most commonly used grammatical words. The tendency for the retention of the discussed variant may be further seen in forms *ye* and *y*^{t} being still in use in the Early Middle English period (Lass 1999: 36). *They*, however, proved to be much more difficult to explain. According

to Lass (ibid.: 120-121), the introduction of the Scandinavian third person plural paradigm, which replaced the Old English one, progressed in three consecutive stages: (1) *þei / her(e) / hem*; (2) *þei / her(e) ~ þeir / hem*; (3) *þei / þeir / hem ~ þem*. As the change progressed through the fifteenth century, one might see that the third person personal pronoun was the first one to adopt the Scandinavian paradigm. Hence, *they*, in comparison to *their(s)* and *them*, may be treated as the first pronoun with the word-initial voiced dental fricative. *They*, similarly to *the* and *that*, may retain the variant <þ/y> due to the fact that it appeared early in the fifteenth century, whereas the remaining *their(s)* and *them*, which appeared later in the century, were more likely to adopt the standardised spelling <th>.

Finally, looking at the analysis in its entirety, it seems that it provided strong evidence for the existence of the Northern System. The variants of the variable (th) showed a tendency to appear in word types with pre-assumed voiced or voiceless dental fricative. Although the material displayed a certain dose of variance in terms of grammatical words employing both <þ/y> and <th> word-initially to represent the voiced dental fricative, in lexical words, only <th> was used. It might be argued that the state recorded in the material showed the Northern System in its slow decline. Spatial and chronological distribution, along with the rate of occurrence of the two variants in the legal documents may be used to confirm a statement that the standardised variant <th> was already entering the system. The retention of the assumed variant in the most common and, at the same time, the earliest grammatical words used in the present study, may further validate this view.

References

Benskin, M. 1977. Local archives and Middle English dialects. *Journal of the Society of Archivists* 8: 500–514.

Benskin, M. 1982. The Letters <þ> and <y> in later Middle English, and some related matters. *Journal of the Society of Archivists* 1: 13–30.

Benskin, M., M. Laing, V. Karaiskos, and K. Williamson. 2013. An Electronic Version of A Linguistic Atlas of Late Mediaeval English (Version 1.1). Retrieved from http://www.lel.ed.ac.uk/ihd/elalme/elalme.html, 14.12.2013.

Fulk, R.D. 2012. *An Introduction to Middle English Grammar and Text*. Peterborough, ON, Canada: Broadview Press.

Hogg, M.R. 1992. Phonology and Morphology. In *The Cambridge history of English language*, vol. 1, ed. M.R. Hogg. Cambridge: Cambridge University Press, 67–164.

Jensen, V. 2012. The consonantal element (th) in some Late Middle English Yorkshire texts. In *Studies in Variation, Contacts and Change in English* 10, eds. J. Tyrkkö, M. Kilpiö, T. Nevalainen, and S. Rissanen. Retrieved from http://www.helsinki.fi/varieng/series/volumes/10/jensen/, 08.01.2014.

Lass, R. 1999. Phonology and Morphology. In *The Cambridge history of English language*, ed. N. Blake, Cambridge: Cambridge University Press, 23–154.

Quirk, R., and C.L. Wrenn. 1957. *An Old English Grammar* (2nd Ed.). London: Methuen and Co.

Scragg, D.G. 1974. *A History of English Spelling.* Manchester: Manchester University Press.

Stenroos, M. 2004. Regional dialects and spelling conventions in late Middle English: searches for (th) in the LALME data. In *Methods and Data in English Historical Dialectology*, eds. M. Dossena, and R. Lass, Bern: Peter Lang, 257–285.

Stenroos, M. 2011. MEG-C Catalogue of Sources, (Version 2011.1). Stavanger: University of Stavanger. Retrieved from http://www.uis.no/getfile.php/Forskning/Kultur/MEG/Catalogue_2011_Master_3.pdf, 02.11. 2013.

Stenroos, M., and M. Mäkinen. 2011. MEG-C Corpus Manual (Version 2011.1). Stavanger: University of Stavanger. Retrieved from http://www.uis.no/ getfile.php/Forskning/Kultur/MEG/Corpus_manual_%202011 _1.pdf, 05.11.2013.

Stenroos, M., M. Mäkinen, S. Horobin, and J. Smith. 2011. The Middle English Grammar Corpus (Version 2011.1, concordance version). Retrieved from http://www.uis.no/research/culture/the_middle_english_ grammar_project/>, 02.11.2013.

Upward, C., and G. Davidson. 2011. *The History of English Spelling.* Chichester: Wiley-Blackwell.

Appendix

The present table contains the information about the repository shelfmark for each transcription of the original scribal text used in the present study. All information given in the table was based on the *Catalogue of Sources* accompanying the MEG-C and refers only to the portions of MSS used in this corpus.

Legal Document	Source text used for the compilation of the MEG-C electronic corpus	Date
1	2	3
	NORTH	
	Cumberland	
L0107	Whitehaven, Cumbria Record Office: DCU/4/178. Lease	1435
L0117	Carlisle, Cumbria Record Office: D/Stan/21. Award	1459
L0354	Carlisle, Cumbria Record Office: D/Stan/24. Exchange	1489–90
L0355	Carlisle, Cumbria Record Office: D/Lons/L/Deeds/Lowther 116. Marriage settlement	1502
L0452	Carlisle, Cumbria Record Office: 1 Ca/Misc/Deeds/15th c. Commissioning agreement	1434
L0535	Carlisle, Cumbria Record Office: D/S/Eaglesfield Deeds/1440-41. Agreement	1441
L0547	Whitehaven, Cumbria Record Office: D/Stan/46. Grant	1503
L1035	Carlisle, Cumbria Record Office: D/Lons/L/Deeds/Askham 56. Gift	1450
L1038	Carlisle, Cumbria Record Office: Ca 5/1/20, box 1397-1794 / folder 1430-1747. Commitment to arbitration	1430
L1105	Gosforth, Northumberland Record Office: ZHW 1/85. Lease	1448
L1111	Gosforth, Northumberland Record Office: ZHW 1/76. Lease	1432
L1143	Carlisle, Cumbria Record Office: D/Lons/L/Deeds/D. 54. Lease	1429

1	2	3
L1144	Carlisle, Cumbria Record Office: D/Mus/Penrith/Medieval Deeds. Gift	15ab
L1145	Carlisle, Cumbria Record Office: D/Penn/28 no. 20 (Bretby Bundle). Lease	1439
L1146	Carlisle, Cumbria Record Office: D/Lons/L/Deeds/Lo. 111. Marriage settlement	1456
L1170	Carlisle, Cumbria Record Office: Ca 2/15. Ordinance	1445
L1182	Kendal, Cumbria Record Office: WD/Ry/92/107. Lease	1483
L1184	Kendal, Cumbria Record Office: WD/Ry/92/93. Award	1453
L1187	Kendal, Cumbria Record Office: WD/Ry/92/87. Enfeoffment	1438
L1188	Kendal, Cumbria Record Office: WD/Ry/92/77. Commitment to arbitration	1422
L1189	Kendal, Cumbria Record Office: WDEC/2. Enfeoffment	1436
L1192	Kendal, Cumbria Record Office: WD/Ry/92/79. Surety	1425
L1196	Carlisle, Cumbria Record Office: DMus/Edenhall 2/2/100. Memorandum	15/16
L1223	Gosforth, Northumberland Record Office: ZHW 1/97. Lease	1494
L1243	Durham, Prior's Kitchen, Dean and Chapter Muniments: Misc. Charter 51. Award	1433
L1249a	Carlisle, Cumbria Record Office: D/Lons/L/Deeds/Wg.15. Marriage settlement	1472
L1249b	Carlisle, Cumbria Record Office: D/Lons/L/Deeds/Askham 68. Condition of obligation	1472
L1250a	Whitehaven, Cumbria Record Office: DSTAN/1/15 . Declaration	15a2
L1250b	Whitehaven, Cumbria Record Office: DSTAN/1/16. Memorandum	1432
	Durham	
L0147a	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register I, ff. 101v line 19 to 102r line 14. Letter/Document	1439
L0147b	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register I. f. 122v lines 1-18 and f. 127v. Letter/Document	1440

Cont.

1	2	3
L0147c	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register I. ff. 142r line 18 to 149r line 5. Letter/Document	1441
L0147d	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register I. ff. 149v line 9 to 150r line 10. Letter/Document	1442
L0147e	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register I. ff. 152v line 8 to 154v. Letter/Document	1442
L0147f	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register I. f. 188v lines 1-22. Letter/Document	1444
L0147g	Durham, Prior's Kitchen, Dean and Chapter Muniments: Small Prior's Register II. ff. 9v line 15 to 23v line 8. Letter, Lease	1446–47
L0147h	Durham, Prior's Kitchen, Dean and Chapter Muniments: Prior's Register III. f. 41r lines 22-39. Letter of appointment	1414
L0147i	Durham, Prior's Kitchen, Dean and Chapter Muniments: Prior's Register III. f. 273r line 14 to 273v line 15 and f. 287v line 5 seq. Letter of appointment	1440–42
L1004	Durham, Durham County Record Office: D/Ch/D 92. Bond	1433
L1005	Durham, Durham County Record Office: D/Lo/F 322. Commissioning agreement	1414–15
L1114	Durham, Durham County Record Office: D/St/D1/2/13. Attestation	1452
L1117	Durham, Prior's Kitchen, Dean and Chapter Muniments: 3.10. Spec. 45.a and 45.c. Lease	1470
L1118	Durham, Prior's Kitchen, Dean and Chapter Muniments: 3.10. Spec. 45.b. Copy of the text in 3.10. Lease	1470
L1305	Durham, Prior's Kitchen, Dean and Chapter Muniments: 2.3. Spec. 63. Lease	1448
	Lancashire	
L0085	Preston, Lancashire County Record Office: DDIb (Ireland of Blackburne of Hale). Award	1431
L0087	London, British Library: Add. Charter 17692. Lease	1420
L0113	Preston, Lancashire County Record Office: DDIb (Ireland of Blackburne of Hale). Testimonies	1411
L0121	Preston, Lancashire County Record Office: DDSc (Scarisbrick of Scarisbrick Deeds) 439/162. Affidavit	1445

1	2	3
L0160	London, British Library: Add. 37769 (Chartulary of Cockersand Abbey), f. 18r. Deed	1363
L0254	London, Public Record Office: E 40/9307. Lease	1459
L0293	Durham, Prior's Kitchen, Dean and Chapter Muniments: Locellus IX.35 (recto). Complaints	<?>
L0296	London, Public Record Office: E 40/8559. Award	1450
L0308	London, British Library: Add. Charter 52290. Award	1436
L0335	London, Public Record Office: DL 25/L 691. Bond	1426
L0414	London, Public Record Office: DL 25/398 (Coucher Book of Furness Abbey). Award	1424
L0416	London, Public Record Office: E 40/10386. Award	1458
L0420	London, Public Record Office: DL 25/399. Agreement	1431
L0439	Preston, Lancashire County Record Office: DDBl (Blundell of Crosby) 55/20. Enfeoffment	1405
L0441	Preston, Lancashire County Record Office: DDSh (Crosse of Shaw Hill, Whittle-le-Woods) 1/132. Attestation	1419
L0442	London, Public Record Office: E 40/5631. Assignment	1422
L0445	Preston, Lancashire County Record Office: DDF (Farington of Worden, Leyland) 1932. Lease	1423
L0446	Preston, Lancashire County Record Office: DDPt (Petre and Walmesley of Dunkenhalgh) 24 (1432). Accord	1432
L0451	Preston, Lancashire County Record Office: DDPt (Petre and Walmesley of Dunkenhalgh) 24 (1432). Award	1434
L0586a	Oxford, Bodleian Library: Rawlinson B 460. ff. 91r.9-93v.11. Memorandum of evidences	1424–25
L0586b	Oxford, Bodleian Library: Rawlinson B 460 (The Black Book of Clayton). ff. 93v.12-96r.21. Award	1425
L1012	Leeds, Yorkshire Archaeological Society: DD 53/III/41 (Grantley MSS). Lease	1456
L1198	Preston, Lancashire County Record Office: DDPt (Petre and Walmesley of Dunkenhalgh) 24 (1454/5). Award	1455
L1200	Preston, Lancashire County Record Office: DDPt (Petre and Walmesley of Dunkenhalgh) 24 (1448). Bond	1448

Cont.

1	2	3
L1203	Preston, Lancashire County Record Office: DDPt (Petre and Walmesley of Dunkenhalgh) 22 (1453). Petition	1452
L1205	London, British Library: Add. Charter 62408. Agreement	1425
L1285	Preston, Lancashire County Record Office: DDPt (Petre and Walmesley of Dunkenhalgh) 24 (1430). Award	1430
	Northumberland	
L0334	Gosforth, Northumberland Record Office: ZSW 2/51. Agreement	1426
L0353	Oxford, Merton College: Merton Records 572. Memorandum	1438–1455
L0390a	Durham, Prior's Kitchen, Dean and Chapter Muniments: Locellus V.45 (dorse). Memorandum	1431
L0390c	Durham, Prior's Kitchen, Dean and Chapter Muniments: Locellus V.45 (dorse). Award	1430
L1104	Gosforth, Northumberland Record Office: ZSW 1/150. Accord	1414
L1216	Gosforth, Northumberland Record Office SANT-GUI-NCL-06-01-01. Ordinance	1459
L1225	Gosforth, Northumberland Record Office: ZSW 2/70. Enfeoffment	1505
	Westmorland	
L0120	Carlisle, Cumbria Record Office: D/Stan/26. Gift	1441
L0332	Manchester University, John Rylands Library: Rylands Charter 1945. Enfeoffment	1447
L0528	Carlisle, Cumbria Record Office: D/Mus/Medieval Deeds, box 'Cumberland and Westmorland - Carlisle', Nateby file. Agreement	1455
L1147	Carlisle, Cumbria Record Office: DLons/L5/1/3/82. Award	1478
L1183	Kendal, Cumbria Record Office: WD/Ry/92/101. Lease	1475
L1186	Kendal, Cumbria Record Office: WD/HH/63. Condition of obligation	1487
L1231	Kendal, Cumbria Record Office: Box A/71. Lease	1458
L1233	Sizergh Castle, Kendal: album, no. 20 of Henry VI. Commitment to arbitration	1430

1	2	3
L1234	Sizergh Castle, Kendal: album no 21 of Henry VI. Will	1430–31
L1235	Sizergh Castle, Kendal: Album no 31 of Henry VI. Use (Indenture)	1444
L1236	Sizergh Castle, Kendal: Album no 32 of Henry VI. Commitment to arbitration	1445–46
	The City of York	
L0145	York City Archives: York Memorandum Book A/Y 255. ff. 264v-267v. Memorandum	1428
L1001	York Minster Chapter Library: Dean and Chapter H.1(3), Chapter Acts 1352-1426, ff. 100v-101r. Ordinance	1371
L1348	York, Borthwick Institute: R.I.19. ff. 332v-333v. 25. Revocation, order and confession	15ab
	Yorkshire, East Riding	
L1122	Beverley, Corporation Records: Great Guild Book, f. 23r. Award	1431
L1130	Beverley, Humberside County Record Office: DDCC/19/I. ff. 1v-6v. Boundary survey	1473
L1132	Durham, Prior's Kitchen, Dean and Chapter Muniments: 2.2. Ebor. 19.a. Declaration of gifts	15a
L1133a	Kingston-upon-Hull Corporation Archives: Bench book 1. f. 12r-v. Jurament	15ab
L1133b	Kingston-upon-Hull Corporation Archives: Bench book 2. f. 243. Award	15a
L1134	Kingston-upon-Hull, Corporation Archives: Bench Book 2. f. 212. Award	1417
L1135	Kingston-upon-Hull, Corporation Archives: Bench Book 2. f. 251. Memorandum	1413
L1136	Kingston-upon-Hull, Corporation Archives: Bench Book 2. f. 164. Enactment	1434
L1259	Nottingham University Library: Galway MSS G 9262. Marriage settlement	1477
	Yorkshire, North Riding	
L0126	Northallerton, North Yorkshire County Record Office: ZRL 1/20. Commissioning agreement	1412

Cont.

1	2	3
L0330	Northallerton, North Yorkshire County Record Office: ZDU. Partition	1451
L0331	Northallerton, North Yorkshire County Record Office: ZAZ.z. Declaration	1446
L0350	Northallerton, North Yorkshire County Record Office: ZFL 29. Lease	1430
L0413	Leeds Central Library, Archives Department: RA/M9. Exchange	1447
L1034	Northallerton, North Yorkshire County Record Office: ZQH 1. f. 55r. Grant	1449
L1115	Durham, Durham County Record Office: D/St/D1/2/12. Declaration	1449
L1172	Durham, Durham County Record Office: D/St/D1/2/17. Memorandum	1475–80
L1211	Northallerton, North Yorkshire County Record Office: ZDS I 1/56. Power of attorney	1404
L1212	Northallerton, North Yorkshire County Record Office: ZDS I 2/1. Agreement for re-enfeoffment	1431
L1232	Sizergh Castle, Kendal: Thornton Briggs box, 'Old Deeds', bundle 'Henry VI'. Will	1441
	Yorkshire, West Riding	
L0133	London, British Library: Harley Charter 112.F.1. Will	1412
L0348	London, British Library: Add. Charter 16916. Surrender (Indenture)	1432
L0349	Leeds, Yorkshire Archaeological Society: DD 12/II/3/9/16 (1). Award	15ab
L0360	Leeds Central Library, Archives Department: TN/HX/A13. Affidavit	1479
L0363	Leeds, Yorkshire Archaeological Society: DD 53/III/262. Marriage settlement	1451
L0377a	Huddersfield Central Library: WBD/VIII/10. Lease	1436
L0377b	Huddersfield Central Library: WBM/2. Affidavit	1446
L0378	Huddersfield Central Library: WBD/IX/7. Exchange	1431

1	2	3
L0415	Hull University Library: DDLO 21/27, 21/28, 21/30, 21/32, 21/35, 21/40 (Selby Court Rolls). Court roll	1472–83
L0732	Leeds, Yorkshire Archaeological Society: DD 57/C/W.123. Award	1451
L1102	Doncaster, Bentley Library: DZ FL 1/1 and DZ FL 1/48. Lease (Two indentures)	1472, 1474
L1128	Beverley, Humberside County Record Office: DDCS 44/1. Bond	1415
L1228	Northallerton, North Yorkshire County Record Office: ZFL 59. Award	1440
L1245	Bradford Central Library: WPB 5/18. Lease	1497

Multiple negation in Chaucer's *The Canterbury Tales* as a marker of social status. A pilot study

http://dx.doi.org/10.18778/8088-065-8.03

Maciej Grabski

University of Lodz

Abstract

The aim of the following paper is to examine fragments of Geoffrey Chaucer's *The Canterbury Tales*, one of the most celebrated specimens of Middle English poetry, with regard to the presence or absence of multiple negation. Negative concord, as the structure in question is often referred to, was firmly ingrained in the language in the Old English period, and, having undergone some formal and syntactic modifications, carried into Middle English. The pattern of its decline in the latter parts of the 15th century is observed to correlate with the social status of the speaker, the change originating in the higher tiers of the society. Disfavoring negative concord possibly had sources in the administrative and legal language, the subtleties of which Chaucer, having held a number of official posts with the court and chancery, would have most likely been versed in. Consequently, the paper proposes to at least partly account for Chaucer's choices as regards negative concord from a sociolinguistic perspective and establish a possible connection between the structure's distributional pattern and the status of the *Canterbury Tales* fictional speakers, who come from very different walks of life. Other factors which may have informed or influenced the author's morphosyntactic choices will also be mentioned.

1. Multiple negation and the history of negation in English

Multiple negation is "clauses with two or more negatives which do not cancel each other out" (Iyeiri 1998: 121). Nevalainen and Tieken-Boon van Ostade (2006: 271) observe, after Trudgill, that multiple negation (or negative concord, which

terms they use interchangeably; both will be used with reference to the phenomenon henceforth) is presently among "the socially most marked features" and is completely absent from the modern standard English, which only permits "[s]ingle negation followed by non-assertive indefinites," so "sentences like *I don't want none*" are decidedly characteristic of non-standard varieties of English, and their structural counterparts not uncommon in other Indo-European standard languages. These include, among others, Polish, Portuguese, Persian, Russian, Spanish, or Ukrainian. Multiple negation is generally observed to be absent from West Germanic languages, such as German and, as remarked above, English. In Old English, however, multiple negation, though optional, was widespread and encoded no social information about the speaker. The obligatory negative particle of Old English was *ne*, which was a reflex of one of the reconstructed Proto-Indo-European negators, **ne*; this is also visible in Sanskrit, Latvian, Lithuanian, Old Church Slavonic, or Old High German (Forston 2004: 133). Aside from *ne*, which in Old English would immediately precede the verb it negated, "the addition of more negative adverbs to a sentence adds emphasis to its negativity" (Baker 2012), as demonstrated in (1):

(1) *Ne* derode Iobe *naht* **þæs deofles costnung.**

Not harmed Job *not* the devil's temptation

Job was not harmed by the devil's temptation

(Fischer and van der Wurff 2006: 157, emphasis mine, translation mine)

According to Fischer and van der Wurff (2006: 157), in time, the usage of the extra negative element to the right of the verb would gain ground until the operation became obligatory in the Middle English period, making the negative concord a rule. In ME, therefore, "sentential negation typically consisted of two parts, *ne* and *not*. Indefinites in negative clauses were also expressed by negative forms, a sentential negator co-occurring with *no*, *never*, *neither*, etc." (Nevalainen and Raumolin-Brunberg 2003: 71). With some variation of *not* (deriving from *naht*, *noht*, or *nawiht*) always present in the post-verbal position, the circumstances conducive to the eventual elimination of *ne*, and, consequently, the negative concord, were in place, and *not* took over as the sole negative particle, with variation

between its original, post-verbal, and new, pre-verbal position (Fischer and van der Wurff 2006: 157). These and the subsequent developments of English negation can be described as follows:

Old English	Middle English	Modern English
ne + verb (other negator)	(ne) + verb + not; not + verb	Aux + not + verb; (verb + not)

(Fischer and van der Wurff 2006: 112)

Nevalainen (2009: 580), basing her findings on the Corpus of Early English Correspondence, places the beginning of the decline of multiple negation "in the late fifteenth and sixteenth centuries". The disappearance of pre-verbal *ne* and the replacement of the indefinite negatives with the non-assertive ones were completed by the end of the fifteenth and seventeenth centuries respectively (Nevalainen and Raumolin-Brunberg 2003: 71). The real-time dynamics of the change are presented by Fig. 1.

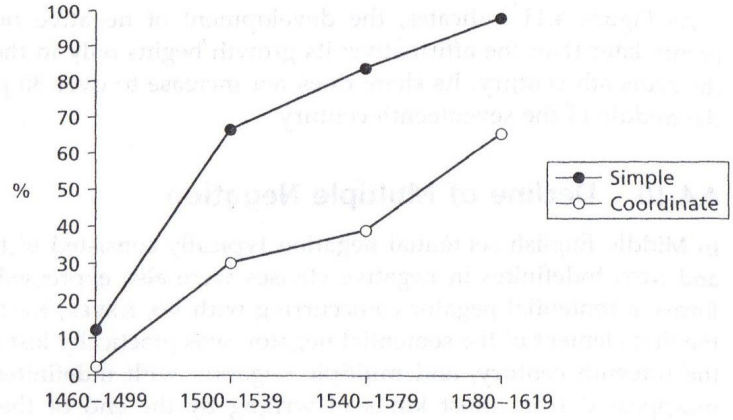


Fig. 1. Single vs. multiple negation with nonassertive indefinites. Percentages of single negation in simple and coordinate constructions. CEEC 1998 and Supplements; adapted from Nevalainen and Raumolin-Brunberg (2003: 72)

Nevalainen (2009: 580) then describes the deterioration of negative concord as "a selective process from above in terms of the speaker-writer's education and social status … promoted by male professional circles in the middle and upper so-

cial ranks", and Nevalainen and Raumolin-Brunberg (2003: 145) observe that "the lower strata lag far behind in the change from multiple negation to single negation". The advancing change can therefore be said to have been socially stratified, and though the leadership of the process may have been alternating over decades between the middle and high ranking males, its provenance is definitely in the non-low echelons (Nevalainen and Raumolin-Brunberg 2003: 149–50). Nevalainen and Raumolin-Brunberg (2003: 150) additionally remark, after Rissanen, that "it is most likely that the model for the non-use of multiple negation derived from administrative language," and that negative concord was scarce in "early legal English". In a similar vein, Mazzon (2004: 83), referring to her earlier study, notices that "scientific and legal prose [in Middle English] presented a much lower number of multiple negations as compared to religious or historical prose". What is more, the trend, "borrowed from the prestigious formal variety," was possibly catching on also among gentlemen and top chancery officers (Nevalainen and Raumolin-Brunberg 2003: 150). Disfavoring the negative concord might have also been additionally informed by "algebraic logic", whereby two negative elements cancel each other out, resulting in an affirmative statement ("The standardisation," 2006).

2. Chaucer and negative concord

A number of authors point to Chaucer's reputation for the "extensive use of double, and even triple and quadruple negatives", and this particular remark of Pereltsvaig's (2010) refers to *The Canterbury Tales*; Mazzon (2004: 83) is of the opinion that Chaucer "appears rather isolated among his contemporaries in his heavy use of multiple negation". However, a closer look at Chaucer's magnum opus actually reveals rather long passages where single negation visibly outstrips multiple negation, but the latter is indeed strongly represented elsewhere. One possible explanation of this constant alternation is simply through referring to a somewhat general linguistic disarray of the time bracket in question; Mazzon (2004: 83) observes "that this period (and late ME more so than EModE) is one in which individual variation emerged freely" and, for example, "Chaucer's *Boece* … showed many more cases of double negation than the earlier *Parker Chronicle* or *Seinte Marherete* … ." Thus, "speaking about the language of Chaucer or the language of Shakespeare is misleading, since there is a considerable amount of variation within the works of these writers".

However, it stands to reason to assume that Chaucer, who studied law and was professionally connected to the court, would have been aware of the change in the making[1]. Although Nevalainen and Raumolin-Brunberg's data clearly show that multiple negation still carried the day for the best part of the 15th century, theirs is a study of personal letters, which are widely considered the most "oral" of written genres, i.e. providing a reliable insight into how the spoken language of the past may have looked like (Nevalainen and Raumolin-Brunberg 2003: 29). Negative concord, in its turn, was observed to have most likely spread from above, both in terms of social stratification, as well as the level of perception, i.e. it was a process induced consciously, to have originally appeared in formal writing, and only later made it to more informal discourses. Therefore, it is not impossible that the personal writings from the Corpus of Early English Correspondence would only start recording the change after a time lag. In his considerations on the dynamics of the change in the periphrastic *do* structure, Warner (2009: 63) tries to account for the temporal discrepancies in his data precisely through reference to generic diversity, and remarks that "the onset of the change in evaluation [of periphrastic *do*] takes effect later in personal letters than in the more public types of writing". I take the liberty of invoking a similar line of reasoning in my subsequent attempts to establish a sociolinguistic pattern of the negative concord's distribution in Chaucer's *Canterbury Tales*, penned in the latter parts of the 14th century.

3. Aims

The aim of the following study is to analyze if the distributional pattern of multiple negative structures in Geoffrey Chaucer's *The Canterbury Tales* might mirror the dynamics of the structure's evolution in the Middle English period. According to a number of sources, the incidence of multiple negation tends to decrease with the speaker's/writer's higher education. Consequently, the study samples the fictional speech of selected characters which are chosen on the strength of their varying levels of education to confirm that in his verse, Chaucer might have been informed by the emergent trend.

1 It has been remarked that the poet was an acute observer of the English language, and this perspicacity of his would often show in his works: it was noticed, for example, with regard to "the third person –s ending [, which] was recognized in the South as a distinctively Northern form. Chaucer puts these forms into the mouths of his north-country students in *The Reeve's Tale* … , whereas the narrator … uses *hath* and *speaketh*" (Crystal 2004: 209).

4. Methodology

The textual material selected for analysis is *The Miller's Tale*, which contains 668 verses (ca. 5200 words), and *The Friar's Tale*, containing 364 verses (ca. 2900 words). In the analysis, all instances of and potential context for multiple negation are counted and later analyzed. The following are typical Middle English negative elements, and the co-occurrence of at least two of these within a clause is counted as an instance of multiple negation: negative particles *ne*, *not* (*nat*, etc.) and their contracted forms such as *nys* (*ne is*), *nolde* (*ne wolde*) etc., and negative indefinites *no*, *never*, *neither*, *nothing*, *none*, etc.

I will be working on the hypothesis that Chaucer was aware that higher registers (legal documents, etc.) had already started to discourage multiple negatives, and he may have further imbued the structure with social markedness so the incidence of negative concord be higher among the lower class, essentially uneducated characters than among the non-lower class, essentially educated protagonists. Thus, the speaker's education will be the most important variable considered in this study.

In addition, negative concord is often rhythm-sensitive, which may well obscure the suggested sociolinguistic account. The question is if rhythm should be considered as another (independent or controlled) variable which could influence the frequency of multiple negation in the analyzed material. Chaucer's poem consistently follows decasyllable meter, and alternating between single and multiple negation, and thus removing or adding an element to a verse, would have often had a bearing on the number of syllables. For example, in "*That noon of us* ***ne*** *speke* ***nat*** *a word*," scrapping negative concord would have truncated the verse to nine syllables, and in "*That to* ***no*** *wight thou* ***shalt*** *this conseil wreye*," attaching the negative particle to the modal would have resulted in the eleventh syllable. Meter, therefore, cannot be ruled out as a possible factor that informed Chaucer's choice. Such explanation, however, will not always hold: in „*This nicholas* ***no*** *lenger* ***wolde*** *tarie*," *wolde*, to establish negative concord, could have been easily replaced with its Middle English contracted negative counterpart, *nolde*, with no effect on the meter, while in „*There* ***nys no*** *man so wys that koude thence*," negative concord is in place, but had Chaucer chosen to go for single negation instead, it could have been seamlessly accommodated into the line by substituting *nys*, short for *ne ys* (*is not*), for the non-negated and also one-syllable-long *ys*. Metrical factors, then, would be a good candidate for what Milroy and Gordon (2003) refer to as the "don't count" cases, i.e. the ones that should be excluded right at the study's outset as somewhat skewed (181 and further).

However, for the time being, the assumption is that Chaucer, widely considered one of the greatest English poets, would have been able to come up with an alternative phrasing of a verse if he was indeed informed by the emerging social markedness of negative concord – and since *The Canterbury Tales* is built around the idea of a story-telling context, whose entrants come from very different sorts and try to assert their individualities, it is not untenable to postulate that Chaucer saw to it that the language of his characters would conform to their respective backgrounds. Under such working assumption, the analysis, intended merely as a pilot study, will tentatively disregard the obvious caveats briefly outlined above, and adopt a sociolinguistic perspective in an attempt to explain the apparently erratic behavior of negative concord. Therefore, the meter-sensitive contexts for negation are included and treated on a par with the remaining cases (although they receive occasional attention as a possible departure point for a future, more comprehensive study).

Toward the proposed sociolinguistic perspective, it is, to be sure, far from a straightforward task to decide on the model of class division for a rather dynamic period of societal changes that the late Middle Ages in England certainly was. However, the case of negative concord may be less problematic in that with regard to the choice of the structure, it is the level of one's education and administrative position that would somewhat superimpose on the social rank as such (Nevalainen and Raumolin-Brunberg 2003: 150). Consequently, the *Canterbury Tales* "speakers" were selected for this pilot study precisely on the strength of these criteria. From *The Miller's Tale*, they are the eponymous miller, the tale's narrator, and a carpenter named John as the representatives of the uneducated circles. There are two other characters in the story whose speech is considered: one is Nicholas, John's tenant, a young scholar, who convinces his host that the second Biblical deluge is imminent, and while the gullible tradesman is hiding, has sex with the carpenter's wife, Alison; the other is Absalom, a parish clerk, who is also attracted to Allison, but not as lucky in his advances as Nicholas. Both are considered essentially educated, with Absalom additionally holding an administrative post. The incidence of negative concord would be therefore expected higher for the first pair of characters.

The Friar's Tale is the story of a summoner on his way to collect a fictional debt from a woman. As he travels through the land, he is joined by a yeoman, who soon reveals that he is a demon, and, toward the story's end, takes the crooked summoner to hell. Friars were typically well-educated, and summoners

held posts in ecclesiastical courts – negative concord is therefore expected to be less prevalent in the speech of the story's narrator and its main villain. Women were generally denied any educational opportunities in the Middle Ages, so the bent officer's victim would likely prefer negative concord. As for the yeoman, his status seems problematic – on the one hand, he represents a group of literate, but not necessarily educated landowning farmers, but on the other, this is just his earthly incarnation, while in fact, he is a demon, acts as such, and makes no secret of his provenance right from the story's outset. His speech is considered, nevertheless.

In accordance with Labov's principle of accountability, wherever double negation is expected, its occurrences, as well as non-occurrences are counted, and vice versa. As regards Middle English negation, it is either multiple or non-multiple, so the condition of the existence of what Labov (qtd. in Milroy and Gordon, 2003: 180–181) refers to as "a closed set of variants" is met, which allows for adopting the occurrence vs. non-occurrence model, instead of reporting the variant's "frequency of occurrence in some globally defined section of speech".

5. Results

In the analysis of *The Miller's Tale,* four characters were considered: Miller (432 lines), John (41 lines), Absalom (58 lines), and Nicolas (110 lines). Allison and some minor characters also "speaking" in the story are not included in the study, showing too few contexts for negation for a quantitative analysis. Table 1 presents the global figures for multiple and single negation for the story's speakers as considered collectively within the educated/uneducated categories (percentages rounded off).

Table 1. The incidence of single and multiple negation in "The Miller's Tale" by the character's education

	single	multiple
educated	23 (82%)	5 (18%)
uneducated	17 (65%)	9 (35%)

Table 2 presents the individual breakdown of single and multiple negation in all the characters considered (percentages rounded off).

Table 2. The incidence of single and multiple negation in "The Miller's Tale" by individual characters

	single	multiple
Miller (uneducated)	11 (65%)	6 (35%)
John (uneducated)	6 (67%)	3 (33%)
Absalom (educated)	5 (71%)	2 (29%)
Nicholas (educated)	18 (86%)	3 (14%)

It transpires that single negation is generally preferred by both educated and uneducated speakers, but the scales tip in its favor more visibly among the former group. Absalom the clerk uses single negation in five out of seven possible contexts. On the other hand, John, the cuckold husband, presented throughout the tale as a crude simpleton, scores 67% for single negation, which puts him almost on a par with Absalom's 71% from a similar number of contexts (six out of nine). However, a more qualitative perspective would possibly lend some support to the working hypothesis. For example, in the passage where Nicholas urges John to keep quiet about the impending disaster, the carpenter is apparently happy to have been taken in young academic's confidence and rejects the notion that he could ever be so stupid as to even flirt with the idea of going public with such confidential intelligence he probably is honored to have been entrusted with; actually, it is where John's stupidity and naivety are at their most prominent, as he unknowingly dances to the cunning scholar's tune. It may have been intentional on Chaucer's part to put double negation in John's mouth here:

> *Quod tho this sely man, I* ***nam no*** *labbe*
> *Ne, though I seye, I* ***nam nat*** *lief to gabbe*
> [*Said then this simple man: I am no blab*
> *Nor, though I say it, am I fond of gab*]

Similarly, Nicholas, describing the upcoming deluge, informs John

> *That half so greet* ***was nevere*** *noes flood*

and uses single negation in the process. On a side note, all three contexts are meter-independent: *nam* could have been replaced with *am*, and *was* changed for *nas*, to discard negative concord in John's and insert it in Nicholas's respective parts,

and the fact that it was not may be regarded as somewhat supporting Chaucer's alignment to the sociolinguistic pattern of the structure's distribution. To be sure, this is a two-edged sword: some cases were discovered which had double negation where single was expected, and vice versa, and also occurred in the rhythm-independent contexts. A case in point could be the following from the miller:

For curteisie, he seyde, he ***wolde noon***

Wolde could have easily been substituted for *nolde* to establish negative concord in the speech of a character expected to use it. Furthermore, the overall incidence of single negatives for the miller is higher than that of multiple negation (eleven to six). However, it was suggested to me that the miller could be seen as a somewhat intermediary character, or even a social aspirer, and going into *The Canterbury Tales* one interpretative layer deeper, it turns out that *The Miller's Tale* comes right on the heels of *The Knight's Tale*; it is therefore possible that the miller, entering the contest immediately after a more sophisticated character, wanted himself to come across as such, or more unconsciously accommodated to the previous, higher-ranking speaker (Patrick Maiwald, personal communication). This interesting observation requires further research. Finally, Nicholas, an up-and-coming scholar, who can be essentially classified as the story's main hero and the ultimate opposite to John, a simple carpenter, employs multiple negation on three occasion only, the remaining eighteen being in tune with the advancing change.

In *The Friar's Tale*, four characters are analyzed as well, and they are the friar (134 lines), the summoner (88 lines), the woman (30 lines), and the yeoman (103 lines). Table 3 presents the figures for single and multiple negation for individual speakers in the story:

Table 3. The incidence of single and multiple negation in "The Friar's Tale" by individual characters

	single	multiple
Friar (educated)	9 (75%)	3 (25%)
Summoner (educated)	9 (82%)	2 (18%)
Woman (uneducated)	2 (33%)	4 (67%)
Yeoman (unclear)	12 (86%)	2 (14%)

The tale's two educated characters, its narrator and the summoner, prefer single negation in nine contexts out of twelve and eleven respectively, while the proportions are reversed in the speech of the woman, who employs it in two out of six possible contexts. Again, this might suggest that the distribution of the structure is not entirely random if a sociolinguistic explanation is put forward. It is also noteworthy that out of three instances of triple negation in the tale (the remaining multiple negatives being double), two are uttered by the woman, and thus account for half the multiple negatives she uses:

> ***Ne** was I **nevere** er now, wydwe **ne** wyf,*
>
> *[Never was I, till now, widow or wife]*
>
> ***Ne nevere** I **nas** but of my body trewe*
>
> *[Nor ever of my body was I untrue!]*

As regards the fourth character under scrutiny, the yeoman, he shows a clear preference for single negation, which outnumbers negative concord at twelve to two. Although his status, as hinted earlier, seems unclear, Chaucer portrays him as a figure held in somewhat high regard by both the summoner and the tale's narrator – under the assumption, therefore, that the distribution of the negative structures is to reflect one's societal status, the author would have sooner intended for the demon to sound in keeping with how he was perceived. The decidedly higher incidence of single negatives in his speech may thus not necessarily go against the working hypothesis. Should the demon be tentatively considered as a representative of the "educated", the distribution of negatives among the characters falling within their respective "educated" and "uneducated" circles would be as follows:

Table 4. The incidence of single and multiple negation in "The Friar's Tale" by the character's education

	single	multiple
educated	30 (81%)	7 (19%)
uneducated	2 (33%)	4 (67%)

5. Concluding remarks

Mazzon's (2004: 83) remark on Chaucer's "heavy use of multiple negation" seems barely to hold for *The Miller's Tale* and *The Friar's Tale*. Single negation dominates in the analyzed fragments, and as it alternates with negative concord, the pattern may not be entirely random, which randomness was elsewhere observed to be generally characteristic of later Middle English (Mazzon 2004: 83). Instead, the incidence of negative concord might be related to the status that Chaucer endowed each of his fictional characters with, and the pattern may be in keeping with other findings pertaining to the direction of the structure's development, but based on the later material gathered from real-life speakers (private correspondence). A number of caveats definitely merit further research, such as the verse-specific metrical constraints on negative structures, not considered in this pilot study in favor of a purely sociolinguistic interpretation. Fuller immersion into the world depicted by Chaucer to consider the relations between the characters he brought to life could result in a more comprehensive and convincing picture of the sociolinguistic patterning of the distribution of negatives.

References

Baker, P. S. 2012. *Introduction to Old English*. Oxford: Wiley-Blackwell.

Crystal, D. 2004. *The Stories of English*. London: Penguin.

Fischer, O., and W. van der Wurff. 2006. Syntax. In *A history of the English language,* eds. R. Hogg and D. Denison, 109–198. Cambridge: Cambridge UP.

Fortson, B.W. 2004. *Indo-European Language and Culture.* Hoboken, New Jersey: Blackwell Publishing

Iyeiri, Y. 1998. Multiple negation on Middle English verse. In *Negation in the history of English,* ed. I. Tieken-Boon van Ostade, 121–146. Berlin: Mouton de Gruyter.

Mazzon, G. 2004. *A history of English negation*. Harlow: Pearson.

Milroy, L., and M. Gordon, 2003. *Sociolinguistics. Method and interpretation*. Oxford: Blackwell.

Nevalainen, T. 2009. Historical sociolinguistics and language change. In *The Handbook of the History of English*, eds. A. van Kemenade and B. Los, 558–588. Oxford: Blackwell.

Nevalainen, T., and H. Raumolin-Brunberg. 2003. *Historical sociolinguistics: Language change in Tudor and Stuart England*. London: Longman.

Nevalainen, T., and I. Tieken-Boon van Ostade. 2006. *Standardisation*. In *A history of the English language*, eds. R. Hogg and D. Denison, 271–311. Cambridge: Cambridge UP.

Pereltsvaig, A. 2010. [Web log message]. Retrieved from http://languagesoftheworld.info/syntax/dont-do-no-double-negatives.html, 16.12.2014.

The standardisation of English. 2006. Retrieved from http://courses.nus.edu.sg/course/elltankw/history/Standardisation/C.htm, 07.03.2014.

Warner, A. 2009. Variation and the interpretation of change in periphrastic do. In *The Handbook of the History of English*, eds. A. van Kemenade and B. Los, 45–67. Oxford: Blackwell.

The effect of previous language experience and 'proper' L2 input on the aspiration of English voiceless stops by Polish adult immigrants to London

http://dx.doi.org/10.18778/8088-065-8.04

Aleksandra Matysiak

University of Lodz

Abstract

The study explores the effect of language experience (with emphasis on the level of L2[1] proficiency on arrival in the UK and the quality of L2 input) in 24 Polish adult immigrants to London who have been learning their L2 on a daily basis in natural surroundings.

Participants were divided into groups according to the abovementioned criteria. The phonetic parameter under investigation is VOT[2] in voiceless aspirated stops /p/, /t/ and /k/ in word-initial positions, analysed on the basis of a reading task and then measured in Praat. The qualitative data were collected by means of a structured interview.

The results suggest the importance of the level of English proficiency in L2 on arrival in the L2 country and its influence on the phonetic system developed by each learner individually. L2 pronunciation level can be developed mainly through the frequent use of the L2 in communication with native-speakers of the language.

1. Introduction

As English has become a language of international communication across the whole world, it is spoken by many non-native learners as their second language. The fact that Poland became a member of the European Union in 2004 created

1 Second language (in this study: English).

2 Voice Onset Time.

conditions for more direct contact with English in L2 speech communities (such as Great Britain, Wales, Scotland or Ireland) for thousands of Polish people who decided to settle down in the British Isles. The wave of mass immigration to the UK started shortly after the enlargement of the EU in May 2004. The majority of Polish people went there to seek employment and better opportunities in general. However, there are also many people who decided to emigrate in order to begin or finish their studies while others initially came as tourists – but in the end they decided to stay there a bit longer.

Exploring the effect of everyday life exposure to the L2 in natural surroundings may be of interest not only from the scientific point of view but it can also be important for teaching and learning English as a second language. Polish immigrants to the UK apply different strategies and represent various approaches towards the area they live in or the language itself, which can be expected to affect the process of second language acquisition. Flege et al. (2001) point out that a wide range of factors can contribute to the effectiveness of the degree of L2 acquisition. These are both external and internal factors. The former include the age of the L2 learner, the length of residence in the L2 speaking country or the learner's gender, while the latter comprise such aspects as, for instance, motivation, L2 learning aptitude, approach to the native-speakers of a given language, exposure to L2 or the amount of L1 and L2 used in everyday life situations. Among other factors, language input and proficiency level in L2 on arrival in the L2 country seem to be the most significant (Flege 1997, 2001, 2009).

The present paper is a follow-up study to the one conducted by Matysiak (2013) in which the influence of such factors as length of residence (LoR) and the amount of L1 and L2 used on a daily basis and their effect on the use of aspiration in Polish immigrants was analysed. Results of the aforementioned study support Flege's (1997, 1999, 2001, 2009) findings that LoR as such has no particular influence on the development of L2 pronunciation skills as some other factors that occur on the way have to be taken into account as well. One of those factors was language proficiency on arrival (Flege 2001; Waniek-Klimczak 2011). As regards L2 input, it was revealed that the frequent use of English on a daily basis by Polish immigrants is not enough to acquire native-like pronunciation as the quality of L2 input was not specified (Matysiak 2013). According to Flege (1999, 2001), L2 learners are supposed to develop their L2 pronunciation skills when they mainly interact with the native-speakers of a given L2. The present follow-up study hopes to shed more light on the issue

of immigrant English by determining whether such factors as the quality of L2 input and L2 level on arrival in the UK can possibly affect the degree of L2 proficiency.

2. Selected factors affecting the degree of SLA

The issue of SLA development in L2 learners has been investigated in a large number of experimental studies. Contrary to what one may think, it is not easy to determine which factors affect the overall degree of SLA as it has always been a broad and complex process. One of the possible explanations may result from the differences in design and methodology of particular studies and this "has led researchers to draw rather different conclusions about the influence that certain factors have on degree of L2 foreign accent" (Piske et al. 2001: 195). The authors mainly focus on such factors as length of residence in an L2 speaking environment (referred to as 'LoR'), the amount of L1 and L2 use in day-to-day communication with L2 speech community and the attitude towards the L2 itself and the L2 environment. When it comes to the language input and L2 proficiency on arrival in the L2 country, it seems that those factors have obtained limited attention from numerous researchers so far. It can be explained by the fact that – on the contrary to such factors as LoR or the age of arrival (AoA), which are relatively easy to measure – it is hard to assess L2 speakers' language proficiency at the moment of arrival in the UK, not to mention the amount of L2 use in interaction with the native-speakers of English. Flege supports that view. The author has already made some attempts to take a closer look at the two factors mentioned above.

2.1. L2 proficiency level on arrival in the L2 country

In his numerous studies on the subject, Flege (1992, 1997, 1999, 2001) focuses mainly on the age of arrival in the L2 country (AoA) and suggests that L2 speakers ought to be divided into two groups: early and late learners. According to the author (1999, 2001), those L2 speakers who started learning L2 relatively early (up to the age of 15) are more likely to acquire native-like pronunciation than those who had their first contact with the second language after that period. A study by Flege, Bohn and Jang (1997) conducted among experienced and inexperienced non-native subjects revealed that the former produced English vowel sounds more accurately than the latter. Hence, it can be concluded

that the earlier one starts L2 learning, the more effective the SLA process is in such a learner. Unfortunately, it is not easy to find studies devoted to the issue of L2 proficiency level on the arrival in the UK in Polish immigrants and its influence on the overall SLA process. However, on the basis of Flege's (1997, 1999, 2001, 2009) previous work it can be assumed that those immigrants who came to the UK with relatively high level of spoken and written English are less likely to have problems with every-day life communication with the L2 community. Consequently, they tend to be more open and use more English on a daily basis. On the contrary, those who came to the UK with the basic level of L2 (or even with no previous L2 experience at all) can have problems with day-to-day interaction with the L2 community as the so-called affective filter and language shock they experience simply hinders the process of second language acquisition. A more recent study by Waniek-Klimczak (2011) conducted among proficient English learners who decided to settle down in the UK confirms the assumption that such people are at an advantage. What is more, the overall attitude towards L2 and the use of different acculturation strategies seem to be dependent on the L2 level at the very start. However, it has to be mentioned that in the abovementioned study only highly proficient L2 learners were taken into consideration.

2.2. Proper vs. improper L2 input

Literature devoted to the issue of L2 input seems to confirm the hypothesis that in this case two factors can possibly contribute to the development of L2 proficiency (especially in the area of pronunciation). These are: the amount of L1 and L2 used on a daily basis and the quality of L2 input: native (referred to as a 'proper' input) vs. non-native (the so-called 'improper') input. Numerous studies conducted by Flege et al. (1997, 1999, 2001, 2009, 2011) seem to support the view that L2 speakers who receive substantial L2 input from native speakers of a given L2 are more likely to acquire native-like pronunciation than those who communicate mainly with other L1 speakers or non-native speakers' community in the L2 environment. In his studies, Flege (1997, 1999) divided immigrants into various groups on the basis of such factors as the age of arrival (early vs. late bilinguals) or the age at which the first contact with L2 took place (early vs. late learners). Those factors are related to each other and if we take those into account, previous language experience of a given L2 learner in the immigrant

society would tell us more about the ability to acquire their L2 in the so-called 'natural context', that is through day-to-day interactions with the members of the target community.

It seems that many authors have not been clear as regards the notion of 'L2 input'. The question about the importance of L2 input was very often posed by Flege (2009: 175) who understands this term as "all L2 vocal utterances the learner has heard and comprehended, including his own, regardless of whether these utterances have been produced correctly by L2 native speakers or incorrectly by other non-native speakers of L2." According to the author (ibid.), such a phenomenon is related to the spoken rather than the written language as "reading seems to have a negligible effect on L2 speech learning, apart from the occasional 'spelling' pronunciation of certain words that have been read but never heard."

Previous studies on L1 and L2 input conducted by Flege (2009) indicate that L1 input would be more adequate than the L2 one and it would always influence L2 pronunciation in adult immigrants: both early and late learners. The reason is that when children learn L1 phonemes, they develop long-term representations of each contrastive unit and perceptually implement them into the L1 speech. Although early and late learners may receive equally proper L2 input, they differ in the frequency of exposure to such input or the use of it. It is strictly connected with the so-called 'critical period' of L2 learning (that takes place no longer than up to 15 years old). Flege (1997, 2001) reported that the immigrants who are early learners (and early bilinguals at the same time) are more likely to achieve native-like pronunciation than early or late learners who became late bilinguals. According to the aforementioned studies, there are two types of L2 input: native (proper) and non-native (improper). Many experiments conducted so far have revealed that those immigrants who interact mainly with native speakers of L2 in the L2 environment are more likely to develop their L2 pronunciation level. However, sometimes it is not easy to distinguish between 'proper' and 'improper' L2 input as after arrival in a predominantly L2 speaking area, immigrants interact either with non-native speakers or native speakers from various dialect backgrounds and they hear different accent varieties of L2. Flege (2009: 177) claims that "the L1-inspired foreign accents of the compatriots tend to match the immigrants' own foreign accents and thus tend to reinforce them". That poses many questions, among which one seems to be particularly significant – how to assess the quality of L2 input effectively?

3. The use of aspiration by Polish adult immigrants to London – a study

The study reported here is divided into two parts and aims at exploring the effect of previous language experience and the quality of L2 input on the use of aspiration in English. It ought to be explained that 'previous language experience' is understood as L2 proficiency level at the moment of arrival in the L2 country. As it is practically impossible to measure it, data had to be obtained by means of a detailed questionnaire. In the first part of the study, possible L2 level on arrival in Polish immigrants to London was assessed on the basis of respondents' answers related to the questionnaire. As for the second part of it, the same questionnaire was analysed in order to find out more about the quantity and quality of L2 input. The study aimed at finding out whether Polish immigrants have a tendency to aspirate word-initial voiceless stops /p/,/t/ and /k/ in a given word (followed by a vowel sound) and then to investigate the possible relationship between their L2 proficiency on arrival in the UK (previous L2 experience) and the quality of L2 input (native vs. non-native) and VOT measurements for /p/,/t/ and /k/. Moreover, the study was conducted in order to investigate possible differences between participants divided into two groups on the basis of such criteria as L2 proficiency at the moment of arrival and the quality of L2 input. For the purpose of the study, T-tests for two-tailed and one-tailed independent samples were applied. The following hypotheses were formulated for the purpose of the study – first of all, previous L2 experience affects the overall level of aspiration in Polish immigrants to London. It was assumed that those immigrants who came to London with a certain level of spoken and written English would achieve better VOT results. Secondly, the quality of L2 input on a daily basis plays a significant role in successful SLA. According to the hypothesis, it was expected that those immigrants who receive more native-like input in day-to-day interaction aspirate certain sounds better than those who receive mainly non-native speech input. VOT was chosen as a phonetic parameter for two reasons. First of all, aspiration is considered to be one of the most salient features of English pronunciation. Secondly, it can be treated as an indicator of a positive attitude towards the L2 speech community (Waniek-Klimczak 2011).

3.1. Method

As in the previous study (Matysiak, 2014), recorded data were analysed. All participants were asked to read out 38 English words related to the picture of a busy street (Figure 1) and were recorded by the researcher. For the purpose of the study the following six words creating positive conditions for the use of longer VOT values (aspiration) in English were chosen for analysis: *cafe, car, pipes, police car, policeman* and *taxi*. For the analysis two types of variables were taken into account: two independent ones (which varied among speakers – previous L2 experience understood as L2 proficiency on arrival in the L2 country and the quality of L2 input understood as the distinction between native and non-native input) and one dependent variable (VOT measurements in English voiceless aspirated stops: /p/, /t/ and /k/).

Fig. 1. "The street" (adapted from "My First Thousand Words in English" by Usborne Publishing – Usborne Children's Books)

The speakers were also given a questionnaire composed of 21 questions (Figure 2) covering such issues as age at the moment of immigrating to London/the UK (LoR), previous language experience, motivation for L2 using and learning, attitude towards L2 speech community and the language itself, amount of L1 and L2 used in everyday life situations or their plans for the future (connected with possible settlement). Participants had to read out the questions (given in English) and answer those in their L2 in the form of a structured interview. They were first familiarized with the material and then, after a short time, they were asked to read out the words and then the questions, possibly at a natural speed. The recordings were made in each speaker's place of residence (or the place of meeting suggested by each participant) between August and September 2012 in London, UK.

QUESTIONNAIRE:

1. When and where were you born?
2. What's your mother (first) language?
3. What's your second language?
4. Are there any other languages you speak?
5. When did you come to London? How old were you at that time?
6. Why did you decide to come here? To find a job/to study/to improve your English?
7. Did you learn English before coming to the UK? If yes, how long was that and how did you learn the language (regular school classes, special courses etc.)
8. How do you learn English in the UK? Is it important for you to improve your language skills?
9. How would you assess your English before you came here and now?
10. Do you speak more Polish or English in everyday life situations?
11. How much Polish and English do you speak at home/at work/among friends/ when you have to communicate with British people (while doing the shopping etc.)?
12. Are there more Polish or English people in the community you live in?
13. Do you read any Polish newspapers/magazines or watch TV/radio programmes or films in Polish? How often do you do that?
14. Are you interested in what happens in Poland? Do you follow the news about the country of your origin?
15. How often do you go to Poland? Do you miss your country when you are in London?
16. How important is it for you to be recognized as a person of Polish origin?
17. Do you think the fact that you are Polish helps you in everyday life situations (like looking for a job etc.) or not? Are there any stereotypes of Polish people in the UK?
18. What do you think about English itself? Do you like the language, its melody etc.?
19. Do you like spending your free time with British people or do you prefer to have contact with your Polish friends? Do you take an active part in your community's social life?
20. What was the most difficult for you when you first came here? What kind of problems did you have as regards your new job, everyday life routine etc.?
21. Do you plan to settle down in London for good? Why?

Fig. 2. Questionnaire used for the purpose of the study

3.2. Participants

The first part of the study was conducted among 24 adult Polish immigrants to London: 14 male and 10 female speakers. All of the participants were born in Poland, and their L1 is Polish. English was declared as L2 by 22 speakers, while 2 speakers claimed that English is their third language. The age of participants ranged from 20 to 35. The speakers also declared varied LoR (ranging from minimum 6 months up to 15 years). The same situation could be observed as regards previous language experience: starting from those who declared practically no contact with English before coming to the UK, ending with those who rated their proficiency level up to B2 on the arrival. The amount of L1/L2 used on the daily basis varied as well, some immigrant L2 learners stated that they used more Polish and some stated they used more English in day-to-day interactions with the L2 speech community. However, the L2 input was different among the speakers who declared more L2 used in every-day situations as some participants claimed that they communicate mainly with native speakers of English while others admitted that although they use a lot of English on a daily basis, such interaction take place mainly between them and non-native speakers of L2.

As the main objective of the study was to explore the effect of previous L2 experience among Polish adult immigrants to London and the quality of L2 input, participants were divided into different groups according to a given variable taken into consideration when grouping the results. Flege (2001) has revealed that the use of L1 has no significant influence on acquiring native-like pronunciation.

Hence, it was reasonable to divide the speakers into two groups with respect to their L2 proficiency on the arrival: those immigrants whose level ranged from B1 to B2 (n=11, Table 1) and those with the L2 level ranging from A1 to A2 (n=13, Table 2).

As regards the amount of L1 and L2 use, it was reasonable to take only those who use English more often than Polish into account and then divide them into 2 groups according to the quality of L2 input they receive: the first group comprised speakers who declared that they communicate mainly with native speakers of English on a daily basis (n=9, Table 3), the second – those who claimed that they most often interact with non-natives (n=9, Table 4).

Table 1. VOT values for /p/, /t/ and /k/ in the group of Polish immigrants with B1 – B2 level of L2 on the arrival (n= 11) in investigated words

Speaker/word – Vot [ms]	Café	Car	Police car	Policeman	Pipes	Taxi
aR	71.02	133.71	59.53	68.13	41.15	42.4
KK-M	72.09	78.26	65.03	46.68	42.26	42.83
NL	94.19	108.12	79.29	79.09	82.71	104.27
KH	59.39	98.3	102.43	78.39	61.09	52.94
AA	73.72	116.04	75.32	104.21	79.03	42.06
RB	98.45	110.1	41.03	36.12	80.1	44.73
PH	63.78	96.43	66.75	62.21	80.52	40.21
KK	47.13	70.97	73.82	59.61	36.04	25.14
IK	66.84	89.31	51.04	71.57	50.22	32.62
DK	44.23	81.08	38.94	53.07	62.81	37.66
MB	64.06	111.69	79.16	46.01	99.55	26.02
MEAN VALUE	**68.6**	**99.5**	**66.6**	**64.1**	**65.0**	**44.6**

Table 2. VOT values and for /p/, /t/ and /k/ in the group of Polish immigrants with A1 – A2 level of L2 on arrival (n= 13) in investigated words

Speaker/word – Vot [ms]	Café	Car	Police car	Policeman	Pipes	Taxi
aL	55.01	102.08	83.03	53.43	42.09	34.06
MJ	59.97	95.91	39.27	47.24	35.48	31.47
MP	89.36	26.17	47.01	44.32	38.81	38.17
CS	76.88	58.49	68.42	51.16	53.27	69.18
JP	59.01	38.8	22.14	28.45	21.03	11.48
MO	48.43	70.96	36.01	54.45	33.02	42.05
SN	71.45	90.54	38.02	37.01	40.63	22.01
WK	57.82	69.85	50.15	39.58	33.01	34.03
BK	62.25	66.04	59.12	50.08	23.62	48.33
PW	38.13	43.58	35.02	33.32	42.03	38.33
MK	56.83	67.19	68.21	49.01	56.88	42.25
EM	32.05	84.07	24.57	26.08	30.14	29.03
DS	56.36	45.68	23.08	22.03	20.07	26.04
MEAN VALUE	**58.7**	**66.1**	**45.7**	**41.2**	**36.2**	**35.9**

Table 3. VOT values for /p/, /t/ and /k/ in the group of Polish immigrants (n=9) who declared they receive more native-like L2 input on the daily basis interaction with L2 speech community

Speaker/word – Vot [ms]	Café	Car	Police car	Policeman	Pipes	Taxi
kK-M	72.09	78.26	65.03	46.68	42.26	42.83
NL	94.19	108.12	79.29	79.09	82.71	104.27
KH	59.39	98.3	102.43	78.39	61.09	52.94
AA	73.72	116.04	75.32	104.21	79.03	42.06
RB	98.45	110.1	41.03	36.12	80.1	44.73
IK	66.84	89.31	51.04	71.57	50.22	32.62
SN	71.45	90.54	38.02	37.01	40.63	22.01
BK	62.25	66.04	59.12	50.08	23.62	48.33
MK	56.83	67.19	68.21	49.01	56.88	42.25
MEAN VALUE	**72.8**	**91.5**	**64.4**	**61.4**	**57.4**	**48.0**

Table 4. VOT values for /p/, /t/ and /k/ in the group of Polish (n=9) who declared they receive more non-native L2 input on the daily basis interaction with L2 speech community

Speaker/word – Vot [ms]	Café	Car	Police car	Policeman	Pipes	Taxi
aR	71.02	133.71	59.53	68.13	41.15	42.4
EM	32.05	84.07	24.57	26.08	30.14	29.03
MO	48.43	70.96	36.01	54.45	33.02	42.05
MJ	59.97	95.91	39.27	47.24	35.48	31.47
MP	89.36	26.17	47.01	44.32	38.81	38.17
AL	55.01	102.08	83.03	53.43	42.09	34.06
KK	47.13	70.97	73.82	59.61	36.04	25.14
WK	57.82	69.85	50.15	39.58	33.01	34.03
DK	44.23	81.08	38.94	53.07	62.81	37.66
MEAN VALUE	**56.1**	**81.6**	**50.3**	**49.5**	**39.2**	**34.9**

The VOT values were measured on the basis of spectrograms and waveforms generated with Praat (2001). As the division mentioned above does not distinguish between more and less proficient L2 learners, it was reasonable to create another group. This time the immigrants with higher level of English on arrival interacting mainly in English were divided into those who are exposed to more non-native L2 input (Table 5) and more native-like L2 input (Table 6).

Table 5. VOT values for /p/, /t/ and /k/ in the group of Polish immigrants (n=6) with higher level of L2 on the arrival and non-native L2 input on the daily basis interaction with L2 speech community

Speaker/word – Vot [ms]	Café	Car	Police car	Policeman	Pipes	Taxi
aR	71.02	133.71	59.53	68.13	41.15	42.4
KK-M	72.09	78.26	65.03	46.68	42.26	42.83
PH	63.78	96.43	66.75	62.21	80.52	40.21
KK	47.13	70.97	73.82	59.61	36.04	25.14
DK	44.23	81.08	38.94	53.07	62.81	37.66
MB	64.06	111.69	79.16	46.01	99.55	26.02
MEAN VALUE	**60.4**	**95.4**	**63.9**	**56.0**	**60.4**	**35.7**

Table 6. VOT values for /p/, /t/ and /k/ in the group of Polish immigrants (n=6) with higher level of L2 on the arrival and native-like L2 input on the daily basis interaction with L2 speech community

Speaker/word – Vot [ms]	Café	Car	Police car	Policeman	Pipes	Taxi
nL	94.19	108.12	79.29	79.09	82.71	104.27
KH	59.39	98.3	102.43	78.39	61.09	52.94
AA	73.72	116.04	75.32	104.21	79.03	42.06
RB	98.45	110.1	41.03	36.12	80.1	44.73
IK	66.84	89.31	51.04	71.57	50.22	32.62
MEAN VALUE	**78.5**	**104.4**	**69.8**	**73.9**	**70.6**	**55.3**

3.3. Results

The results are presented by means of VOT values given in milliseconds, mean values and the results of t-tests for independent samples for each hypothesis investigated in the study. The data were analysed for particular groups on all words and then the groups were compared with respect to L2 level on arrival in the L2 country and the quality of L2 input (native vs. non-native).

3.3.1. Previous L2 experience – does it really matter?

The analysis for group results for those immigrants whose L2 proficiency on the arrival was lower (Table 2) than more proficient ones (Table 1) indicates that there are statistically significant differences in 4 out of 6 instances of words in which aspiration was likely to occur. If we take a look at mean values of VOT measurements (Tables 1 and 2) for particular words it is noticeable that the more proficient group was definitely more successful in aspirating particular voiceless stops at the beginning of each word. It may result from the fact that such L2 learners are more aware of the existence of some salient features of English pronunciation and selected aspects of connected speech as they are more experienced in learning of L2.

Table 7. The significance level of the t-test applied to the independent samples in two groups (group 1; n= 11 and group 2; n= 13) with different L2 proficiency on arrival – ranging from A1 to A2 for the 1st group and from B1-B2 (the 2nd group). *P* value is given in the box

Café	Car	Police car	Policeman	Pipes	Taxi
0.1404831	0.0009219	0.013303	0.001386666	0.000275	0.240511

Interestingly enough, *car* turned out to be the word in which aspiration was the most audible. That would support the fact that /k/ is the strongest out of other fortis stops (aspiration in /k/ starts at above 50 milliseconds, while for /t/ and /p/ – nearly 30). However, if *car* was so strongly aspirated – what was the problem with the word *café* ($p>0.05$) in which VOT measurement are varied? One explanation that comes to mind is that this word turned out to be particularly problematic for Polish immigrants mainly because of its stress placement. Some of the participants got used to British-like stress pattern (the first syllable stressed), while others to the one typical for General American with the stress placement on the second syllable. It is commonly known that unstressed syllables are characterized by a much weaker aspiration level. Another important aspect could be the

frequency of 'car' – perhaps it occurs more often than 'café' in day-to-day communication with L2 speech community.

It also has to be mentioned that the voiceless stop /t/ turned out to be the one with the weakest level of aspiration – both in more and less proficient speakers, regardless of the L2 input quality.

All in all, the first hypothesis formulated at the beginning has to be confirmed as there is a certain relationship between the level of aspiration in English voiceless stops and the L2 proficiency on the arrival in the UK.

3.3.2. Is L2 input enough?

When asked about the quality of L2 input in communicating with L2 speech community, 9 participants declared that they interact mainly with native speakers of the second language (Table 3), 9 admitted that they communicate with non-native speakers of English on a daily basis (Table 4). However, the amount of previous language experience (L2 level on the arrival) was varied among those two groups. As far as the results are concerned, it is worth mentioning that statistically significant differences in VOT levels can be noticed in 2 words out of 6. These words are *café* and *pipes* – it indicates that inexperienced L2 speakers had problems with less frequent words (which corresponds to more recent studies on possible effects of word frequency conducted by Munro and Derwing 2008).

Table 8. The significance level of the t-test applied to the independent samples in two groups (group 1 – 'proper' L2 input and group 2 – 'improper' L2 input; in both groups n= 9) with varied L2 level on the arrival (*p* value given)

Café	Car	Police car	Policeman	Pipes	Taxi
0.025189	0.460124	0.2696227	0.261279108	0.047634	0.140399

On the basis of the results it can be assumed that the L2 input is not the crucial factor that would possibly affect aspiration in Polish immigrants. However, if we look at the mean value for VOT in Table 3 and 4, we can easily find out that the participants exposed to 'proper' L2 input (interaction mainly with native speakers of L2) achieved better VOT results than the group of those who communicate with the non-native speech community (receiving the so-called 'improper' input).

So as to confirm the hypothesis which states that the influence of L2 input ought to be correlated with L2 proficiency level on arrival, it was decided to investigate possible differences only within those speakers who assessed their L2 level as com-

municative (B1-B2) at the moment of arrival but the quality of L2 input was varied among them as some declared that they interact mainly with native speakers of L2 on a daily basis (n=5) and others – with non-native L2 speech community (n=6). There were 18 participants in total. For the purpose of investigation for possible differences in VOT within those speakers, a t-test was applied (one-tailed, independent samples). As it can be seen from the table below, a statistically significant difference can be noticed in the production of VOT in only one word, which is *café*. It may be related to some problems with the primary stress placement mentioned above. What was particularly striking here was the fact that both groups achieved similar VOT values in case of *police car* (with the initial /p/ sound as the aspirated one). Nonetheless, the weakest level of aspiration can be noticed in case of *taxi*.

Table 9. The significance level of the t-test applied to the independent samples in two groups (group 1; n= 9 and group 2; n= 9) with similar level of L2 proficiency on arrival – ranging from B1-B2 but different quality of L2 input (native vs. non-native input). *P* value is given in the box

Café	Car	Police car	Policeman	Pipes	Taxi
0.034075	0.227573	0.3113249	0.063583363	0.2229	0.068185

On the basis of the abovementioned VOT values for those speakers (Tables 5 and 6), it can be assumed that the quality of L2 input is not a factor that would influence the use of aspiration by Polish immigrants to London – both in those speakers who were more and less proficient on the arrival. Perhaps one of the reasons for such results can be the fact that all of the participants came to the UK as late-bilinguals. Flege (2009) claims that in general L2 input is less adequate in case of late-bilinguals than the L1 input they have received from the very early stage of their development. It can be noticed that even if one starts learning an L2, they receive the input of a foreign-accented teacher in their L1 country. However, after possible arrival in the L2 speaking environment, immigrants hear a variety of regional accents among native-speakers of L2, not to mention that in such metropolitan multicultural societies as London they are very often exposed to non-native speech, having no particular universal model of L2 pronunciation that they could stick to. That is why the second hypothesis formulated for the purpose of this follow-up study has to be rejected.

If we take a closer look at the results of mean values for aspiration, we can conclude that the word *car* had the highest level of aspiration, while *taxi* – the lowest in all groups. In most cases participants tend to aspirate English voiceless stops, yet there was no regular tendency in particular groups as regards VOT values for

particular stops. On the basis of the results related to the participants divided into two groups according to the L2 proficiency level on arrival in the L2 country, it can be said that only two words were not statistically significant. T-tests applied for the purpose of this study has shown that in most cases $p<0.05$, which means that there are statistically significant differences between the groups divided according to the L2 level on the arrival.

The situation is slightly different when it comes to VOT values for L2 input quality measured within more and less experienced L2 learners. It can be pointed out that less experienced participants have certain problems with the production of less frequently used words such as *café* or *pipe*s. Many inexperienced immigrants had a tendency to mispronounce those words as they were not familiar with them.

Consequently, the hypothesis related to L2 proficiency on arrival and formulated at the beginning has to be confirmed. Unfortunately, the second hypothesis that was formulated with reference to the quality of L2 input has to be rejected. There is a need for a follow-up study investigating some other factors that could possibly affect the degree of the L2 acquisition.

4. Discussion

Interpretation of the results and comparing those with participants' responses obtained by means of the questionnaire used in this study suggest that previous language experience (understood as L2 proficiency on arrival) is one of the factors that possibly influences L2 pronunciation level in Polish immigrants to the UK. Such findings correspond to the previous research by Flege (1992, 1997, 1999, 2001, 2009, 2011). As regards L2 input, it was pointed out that native-speakers are a much better source of the so-called 'proper input', at least as regards pronunciation (Flege et. al 2001). Consequently, if an L1 speaker lacks such input in day-to-day interaction with members of the L2 speech community, there is a high probability that the process of L2 acquisition is not very dynamic and we cannot talk about the development of language or pronunciation skills. However, VOT measurements and the overall analysis of the qualitative data (in the form of a questionnaire) obtained for the purpose of the present study reveal that the quality of L2 input is not necessarily the most decisive factor in L2 acquisition as SLA is an extremely complex process in which many different aspects ought to be systematically explored and analysed. There are many reasons for such findings. One of the most important seems to be the fact that although some respondents

are early learners of L2, they acquired a certain L2 level in their country of origin and they arrived in the UK as adults – that classifies them as late bilinguals. Even if some of them declare that they interact mainly with native-speakers of English on a daily basis, it is extremely difficult to assess the quality of such input as native-like speech is often characterized by the presence of different regional accent varieties. Many native-speakers have various dialect backgrounds and hence immigrants hear more than one universal model of L2 pronunciation.

At this point it should be said that the present study has its limitations. To begin with, the relatively small amount of participants who agreed to take part in the recordings (24 speakers in total) may not be sufficient to investigate some regular features or patterns of pronunciation typical of Polish immigrants to London as a larger group. A bigger sample would be of use in this case.

Secondly, the design of the study is far from perfect as, for instance, the words connected with the picture of a busy street were given on a single sheet of paper – hence, some speakers made practically no pauses between given words and that could affect the quality of VOT. Although the words chosen for the purpose of the study create contexts for aspiration, word-stress (*policeman* or *police car* would have weaker aspiration as the main stress falls on the second syllable in each of those words) or the tempo of reading (more careful reading creates better conditions for aspiration) could significantly distort aspiration level.

Thirdly, some of the investigated words were far more frequent than others, which means that the participants were familiar with such words as *car* or *taxi* as they could often hear them and consequently, it could affect their performances to some extent. As regards the number of words investigated, it can be seen that there are limited contexts in which aspiration could be observed (i.e. there was just one word with /t/ as the initial sound). Further studies are needed to establish the pattern of VOT in voiceless aspirated stops /p/, /t/ and /k/ in broader contexts (on the basis of participants' responses obtained from the questionnaire).

Another important aspect is that some answers recorded in the form of a structured interview were imprecise – it was particularly hard to determine the exact time specification or the quality of previous language experience, i.e. how long they have been learning English, when they started (early vs. late learners), where it was (Poland or an English-speaking country) and were those classes regular or not (intensity of such classes/ courses ought to be pointed out). Perhaps the language exposure turned out to be less substantial than the speakers declared in questionnaire. On the basis of such observations, the following question arises: how to measure or

– at least – assess the quality of L2 input objectively? Perhaps longitudinal comparative studies ought to be conducted to take a closer look at this factor, but still there is no effective method of investigating such aspects as previous L2 experience or deciding whether the amount of L2 input is proper or not. As a result, researchers are forced to rely on participants' responses which can be unreliable because of the lack of precision or the ability of sensible assessment. With such a high level of variability among the participants, it would be reasonable to investigate possible factors affecting SLA on the basis of individual differences in the speakers.

Finally, it might be helpful to conduct a kind of comparative analysis exploring the use of aspiration in voiceless stops in two languages: English (L2, non-native language) and Polish (L1, native language). Thanks to such studies the effect of L1 on L2 pronunciation could be explored in detail. I believe it might be a good point of reference for a possible follow-up study or further studies on this aspect in general as there is still a need for filling the gap in literature devoted to the issue of immigrant English.

5. Conclusions

Regardless of the limited number of samples (either the amount of participants or investigated words) and a high level of variability in VOT measurements, it can be said that the results of the aforementioned study offer a number of findings which are worth investigating in further studies on the aspect of SLA and immigrant English in general.

The first finding suggests that L2 proficiency level on arrival in the L2-speaking area can be treated as a determiner of success in the acquisition of L2 pronunciation and can contribute to the overall level of L2 proficiency. This would confirm the previous findings (Flege 2009, 2011) and highlight the importance of the previous L2 experience on arrival in the L2 country.

Nevertheless, the study revealed that there are practically no statistically significant differences between participants as regards the quality of L2 input. As the previous studies on this issue (i.e. Flege 2001, 2009) suggest that language input could possibly refine L2 pronunciation, it cannot be confirmed in case of the present study. The results reported above suggest that there is a need to examine L2 input and its quality thoroughly – this time in combination with other factors (such as length of residence, age of arrival or acculturation strategy) characteristic in case of individual speakers.

References

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341–345.

Flege, J.E. 2009. Give input a chance! In *Input Matters in SLA*, eds. T. Piske, and M. Young-Scholten, 175-190. Bristol: Multilingual Matters.

Flege, J.E., O-S. Bohn, and S. Jang. 1997. The effect of experience on non-native subjects' production and perception of English vowels. *Journal of Phonetics* 25: 437–470.

Flege, J. E., M.E. Frieda, and T. Nozawa. 1997. Amount of native language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 25: 169–186.

Flege, J. E., and K.L. Fletcher. 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America* 91: 370–389.

Flege, J. E. The role of input in second language (L2) speech learning. Paper presented during Accents 2012 Conference on native and non-native accents of English (December 2012, Łódź, Poland).

Flege, J.E., and I. MacKay. 2011. What accounts for "age" effects on overall degree of foreign accent? In *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010*, Vol. 2, eds. M. Wrembel, M. Kul and K. Dziubalska-Kołaczyk, 65–82. Bern: Peter Lang.

Flege, J.E., G. Yeni-Komshian, and H. Liu. 1999. Age constraints on second language acquisition. *Journal of Memory & Language* 41: 78–104.

Matysiak, A. 2013 VOT in Polish immigrants to London: The effect of language experience on the use of aspiration in English. *Gavagai Journal* 3: 36–49. http://filologia.uni.lodz.pl/gavagai/3rdgavagai.pdf. Accessed July 8 2014.

Munro, M.J., and T.M. Derwing, 2008. Segmental Acquisition in Adult ESL Learners: A Longitudinal Study of Vowel Production. *Language Learning* 58(3): 479–502.

Piske, T., I. McKay, and J.E. Flege. 2001. Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics* 29: 191–215.

Waniek-Klimczak, E. 2011. Aspiration and Style: A sociophonetic study of the VOT in Polish learners of English. In *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010,* eds. M. Wrembel, M. Kul, and K. Dziubalska-Kołaczyk, 303–317. Bern: Peter Lang.

Formulaic language in native and learner English – a corpus-based study of silent pauses

http://dx.doi.org/10.18778/8088-065-8.05

Wojciech Rajtar

University of Lodz

Abstract

The aim of this on-going study is to investigate whether intersegmental silent pauses with regard to formulaic sequences in native speaker English and Polish learner English occur in similar patterns, and whether the selection of the most common two- and three-word phrases in both types of English is alike. The analysis was conducted on the basis of the conversational sub-corpus of the British National Corpus and the spoken part of the PELCRA Learner English Corpus (Pęzik 2012), from which potential lexical bundles were extracted. Then, temporal analysis of audio samples from both corpora corresponding to the potential lexical bundles was performed in order to determine their prosodic features. As a result, the degree of relations between formulaicity of an utterance and its prosodic features, as well as the range of the most frequently used formulaic sequences was shown for both types of English analyzed.

1. Introduction and theoretical background

Fluency is one of the most important aims for learners of languages and language teachers. One of the factors that are believed to improve fluency is the use of formulaic language (van Lancker et al. 1981), sequences of which are "always produced fluently with an unbroken intonation contour and no hesitations for encoding" (Peters 1983: 8). Pawley and Syder (1983) also further elaborate on this assumption stating that formulaic language is used more effectively because

a memorized sequence, even if consisting of several words, is easier to process than a creatively generated utterance of the same length. That is one of the reasons why attention devoted to lexical bundles, "the most frequently occurring sequences in the register" (Biber et al. 2004), and their phonological features has increased recently; although the focus has been mainly directed at child language and EFL learners so far. In these areas, several prosodic features have been proven to be typical of formulaic sequences (e.g. high frequency phrases are more likely to be reduced (Bybee 2002, 2006); formulaic sequences less likely to contain pauses and hesitations (Raupach 1984; Dechert 1983). However, little analysis of such features with respect to adult native English or ESL learner English has been conducted as yet.

1.1. Lexical bundles

The first term that is of importance is what may be considered a unit of formulaic language. Lexical bundles, also called n-grams, were defined by Biber (Biber et al. 2004; Biber et al. 1999) as "the most frequently occurring lexical sequences in a register" and in the last decade they have received a lot of attention from scholars and researchers. Lexical bundles are not always equal with traditional language units like noun phrases or adjectival phrases, and may cross over several structures e.g. *In this study we, should be noted that* (Allen 2009).

One of the features of language that benefits from a speaker's knowledge and usage of formulaic language is naturalness, which is very important for language learners. What is more, as shown in a study by Neil Millar, misuse of formulaic language can lead to potential communication difficulties (Millar 2009). In the study, Millar measured speaker reading times of collocations taken from native speaker academic writing and Japanese learners' academic writing. The study showed that those learner collocations that were different from native speaker norms take more time to process in reading.

Lexical bundles are also a convenient research subject because of the method for their discovery. They are extracted purely on the basis of frequency in language through corpus data; in the case of this study – the two spoken language corpora. Because of this method, they can be considered solid, empirical data.

1.2. Native-like selection

The notion of lexical bundles is connected to a large extent to two linguistic capacities discussed in a 1983 essay by Pawley and Syder entitled *Two puzzles for linguistic theory: nativelike selection and nativelike fluency*. The first notion is introduced by the authors as:

> the ability of the native speaker routinely to convey his meaning by an expression that is not only grammatical but also nativelike; what is puzzling about this is how he selects a sentence that is natural and idiomatic from among the range of grammatically correct paraphrases, many of which are non-native-like or highly marked usages (191).

The second notion is introduced as:

> the native speaker's ability to produce fluent stretches of spontaneous connected discourse; there is a puzzle here in that human capacities for encoding novel speech in advance or while speaking appear to be severely limited, yet speakers commonly produce fluent multi-clause utterances which exceed these limits (191).

Pawley and Syder begin explaining the 'puzzle of nativelike selection' by choosing generative grammar as a reference point for their theory. This stance on grammar proposed by Noam Chomsky (1957) is based on a claim that acquiring a set of rules enables a speaker of a language to produce an infinite number of correct sentences and to distinguish them from ungrammatical ones. This approach focuses on a set syntactic rules called grammar that allow speakers to have a creative power of composing an infinitely large set of grammatical sentences. The authors then argue that if a speaker were to fully exploit the "creative potential of syntactic rules", their utterances would not be considered as showing native-like properties. In reality, most of the possible sentences in a language, even though grammatically correct, when looked at, would be "judged to be 'unidiomatic', 'odd' or 'foreignisms'", while only a small per cent of the infinite set would be accepted as native-like by other speakers of the given language.

To illustrate this assumption, Pawley and Syder (1983: 194) provide a piece of narrative, spoken by an elderly New Zealander remembering his family's circum-

stances at the time World War I broke out, and juxtapose it with a paraphrase of the same narrative that is fully grammatical but highly unnatural:

(1) I had four uncles

they all volunteered to go away

and ah that was one Christmas

th't I'll always remember …

(2) The brothers of my parents were four

Their offering to soldier in lands elsewhere in the army of our

country had occurred.

There is not a time when my remembering that Christmas will not take place...

This juxtaposition shows that a set of syntactic rules of generative grammar is not all that a language learner has to acquire in order for his speech to be seen as native-like. Apart from that, a language learner needs to "learn a means for knowing which of the well-formed sentences are nativelike – a way of distinguishing those usages that are normal or unmarked from those that are unnatural or highly marked" (Pawley and Syder 1983: 194).

Making this distinction is particularly difficult for language learners who have learned the grammar of the language from a grammar book without being exposed to the living language in the way it is used by native speakers. Such learners usually produce utterances that, even though grammatically correct, will be perceived as awkward and unidiomatic. The situation is different when one is learning a foreign language by immersion, where a learner acquires a language in a native-like manner. In the process of learning, such a person will be able to acquire the idiomatic use of language along with the command of grammar. The distinction between what is unmarked and marked, or natural and unnatural is what the authors call "the puzzle of nativelike selection."

1.3. Native-like fluency

The second puzzle from the essay by Pawley and Syder is called "nativelike fluency", which is the native speaker's ability to produce continuous, fluent utterances lasting for ten or even twenty seconds (Pawley and Syder 1983). Such a skill is very difficult to achieve for a language learner, and the process of acquiring it takes a very long time. This may even become a problem for native speakers when they, for example, try to "express (…) thoughts on an unfamiliar subject, or to deliver an unrehearsed monologue to a silent audience, as when tape-recording a letter or radio talk, or when called upon to speak in a public address or formal interview" (Pawley and Syder 198: 199–200). In such cases, brain activity increases, making the speaker 'look for words', which often results in phonological and prosodic changes in their speech, e.g. increased number of pauses, more hesitations, changes in tempo.

1.4. Idiom principle vs. open choice principle

John McHardy Sinclair (1991) proposes two principles which determine our strategies in the process of selecting the vocabulary we use in speech. Both principles have much common ground with the notions of native-like selection and native-like fluency. The first one is the idiom principle, which enables a speaker to choose from an array of prefabricated multi-word units. Sinclair (1991: 110) defines this notion so "The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments."

The open choice principle corresponds more to the traditional approach of generative grammar, where the speaker combines single lexical items using syntactic rules. Sinclair states that the two principles work alternatively in our speech, and so we produce alternating sequences of spontaneous creative utterances and prefabricated pieces.

1.5. Holistic storage

The two "linguistic puzzles" proposed by Pawley and Syder (1983) show that in order to achieve native-like performance in a foreign language, we need more than just the Chomskyan concept of generative grammar or just the open choice principle. In order to answer the question of what more is needed, they look at the concepts of what they call 'memorized sequences', which are sequences of words

that are retrieved from the memory as wholes, as opposed to being composed from individual words in the process of speech. This approach is further supported by phenomena like fossilized errors (Myles et al. 1999), multiword utterances in speech of children (Peters 1983), and fluent chunks (Dechert 1983; Raupach 1984) in the speech of language learners.

At first it might seem that a situation in which speech is just a process of recalling prefabricated data limits the creativity of a speaker that is theoretically enabled by the concept of generative grammar. But it is not necessarily so. As Pawley and Syder (1983: 208) put it:

> Possession of a large stock of memorized sentences and phrases simplifies his task in the following way. Coming ready-made, the memorized sequences need little encoding work. Freed from the task of composing such sequences word-by-word, so to speak, the speaker can channel his energies into other activities.

Bolinger (1976: 1) wrote that human language "does not expect us to build everything starting with lumber, nails, and blueprint, but provides us with an incredibly large number of prefabs, which have the magical property of persisting even when we knock some of them apart and put them together in unpredictable ways." This metaphor accurately demonstrates how formulaicity in a language can help its speakers achieve more efficiency than generative grammar allows for, and also how formulaic sequences, even though consisting of individual words, are closely connected in our mental lexicon.

As far as native speaker language is concerned, a number of factors seem to support the existence of such a holistic manner of storage: semantic non-compositionality (i.e. the meaning of a phrase is not the sum of meanings of the words in it), idiomatic expressions (e.g. *kick the bucket, black and blue*), cranberry collocations (e.g. *to and fro*), structures probably passed down from archaic English (e.g. *here comes…, believe you me*) (Moon 1998; Lin 2010). According to Lin (2010: 179) the theory of holistic storage

> [...] is also in line with the assumption that formulaic sequences should form single intonation units, have less internal dysfluencies such as hesitations and pauses, be uttered faster than rule-based language, and require specific accentual patterns or focus distinction.

Another phenomenon that speaks in favour of the theory of holistic storage are the previously mentioned multiword utterances, or formulaic sequences, in children speech and their phonological properties investigated by Peters (1977, 1983) and Plunkett (1990). A sequence of words is very probable to have been memorized by a child as a whole and stored as a single unit in the mental lexicon when the utterance is characterized by a distinctive intonation pattern (Peters 1977), an unbroken intonation contour (Peters 1983), and articulatory fluency (Plunkett 1990). These factors are criteria that have to be met for a string of words to manifest 'phonological coherence', a term which we will discuss later in this paper.

1.6. The frequency-based approach

Another approach used in academic research, which connects formulaic sequences to phonology and prosody is the frequency-based approach. It is concerned with psycholinguistic reasons that influence phonological features of formulaic language. As Bybee (2002) explains it, speech production is a series of "neuromotor routines" that gain efficiency as they are practiced, resulting in a gradually more fluent articulation with each repetition. That is why lexical bundles demonstrate fewer dysfluencies (like pauses or hesitations) and are expected to have increased speech rate and more reductions within their boundaries. In fact, this assumption finds support from Dechert's (1983) and Raupach's (1984) observations that second language learners can utter distinctively smooth and fluent stretches that seem to be formulaic amid their other dysfluent productions. As we can see, the theory of holistic storage and the frequency-based approach can work together well, as both look at phonology of formulaic sequences from different angles. That is why Lin (2010) suggests that the two points of view should be combined when investigating this kind of phenomena. The frequency-based approach also indirectly speaks in favour of using corpus data for investigating phenomena related to formulaic language and fluency.

1.7. Phonological coherence

The previously mentioned term of phonological coherence is a notion first used by Peters (1983) as a criterion used in identifying formulaic sequences in the speech of children. According to her, for an utterance to be phonologically coherent it has to fulfil two conditions (the dual criteria): there has to be an unbroken intonation

contour stretching between the boundaries of the sequence, and no hesitations can occur within the sequence. An utterance can be considered phonologically coherent only if the dual criteria are both met. As for formulaic language of adults, a similar assumption is made by Alison Wray (2004), who claims that within a formulaic sequence there should be fewer pauses and errors than in between them, and that lexical bundles are resistant to dysfluency and inaccuracy within them.

The theory of phonological coherence can be considered to be in agreement with the concept of holistic storage. As Wray suggests, such pieces of language (formulaic sequences) are recalled from memory as holistic units. The argument used by her to support this claim is that retrieval of prefabricated formulaic sequences is quicker than recalling a phrase word at a time. That is why retrieving sequences holistically should result in less hesitations, less pauses and a single intonation contour.

In fact, phonological coherence has been used as proof for psycholinguistic processes since as early as 1970s. During this period, studies of prosodic disambiguation were conducted in order to investigate the notions of deep structure and surface structure. A study by Lehiste (1973) aimed at examining the prosodic features that would facilitate disambiguation. In the study she uses sentences like *The old men and women stayed at home* or *The cop killed the robber with a gun* to illustrate how prosodic features can change the deep structure of an utterance. In the first sentence, the deep structure depends on whether a pause is present after the word *men* or not (the men as well as women were old vs. there were old men and some women). In the latter sentence, a pause after the word *robber* is expected if the cop is the one who has the gun. If there is no pause in this place, the robber is the one who has the gun. These examples illustrate how prosodic breaks within utterances delimit phraseological units, which can be considered evidence for a process called chunking.

The psycholinguistic theory behind this process says that chunks are units of language processing (Peters 1983). Lengthy utterances cannot be processed all at once because of the limitations in the 'computing power' of our brains, and that is why such stretches of speech have to be divided into smaller units, chunks, which are just small enough for our short-term memory to process at once. This further gives ground for the claim that the way in which the language is stored should be reflected in phonological and prosodic features of speech.

As Lin (2010) points out, there is a large interdependence between the theories of holistic storage and phonological coherence, and the two notions seem to prove

each other in research. In some research projects, phonological coherence takes the role of the basis for the claim that language is stored holistically (e.g. Wray 2004). In other cases the assumptions are given the other way round (e.g. Moon 1997). It is hard to choose which of the approaches has advantage over the other, but as Lin wrote,

> It is perhaps useful to point out that phonological coherence is more of a fact because it is measurable. Holistic storage, however, is more of a claim because its existence is inferred based on facts. On this foundation, we need to critically evaluate on a case to case basis if the evidence is strong enough for us to infer holistic storage from phonological evidence.

More evidence for mutual dependence of formulaicity and phonological features of speech is provided by a 1981 study by Van Lancker, Canter and Terbeek. In the study, native speakers were given a task of reading stretches of text containing specific phrases (e.g. *skating on thin ice*) which carried literal or idiomatic meaning. The researchers found that idiomatic expressions are spoken faster than phrases set in a literal context. What is more, a lower number of pauses occurred in the idiomatic phrases.

This research project, too, complies with the theories of phonological coherence and holistic storage/chunking. The phrases that are idiomatic are articulated faster and uninterrupted because they are retrieved holistically from the mental lexicon, whereas the literal strings are uttered more slowly because they have to be built compositionally, for they are not prefabricated.

In a study of a German learner of English by Dechert (1983), the author arrives at conclusions that the learner's spoken English was riddled with pauses, fillers and hesitations. But what is interesting, he also observed that there were also completely fluent stretches in the recordings of the subject that he called 'islands of reliability' (Dechert 1983: 184). The 'islands' were fragments of speech that the speaker was confident about, probably retrieved holistically, thus causing increases in fluency.

2. Methodology and analysis

The analysis was conducted on the basis of the conversational sub-corpus of the British National Corpus and the spoken part of the PELCRA Learner English Corpus (PLEC) (Pęzik 2012).

The PLEC corpus consists of a written part (2.8 million words), which contains samples of learner English, and a 200,000 word time-aligned spoken subcorpus of learner English, which is the part used for this study. The spoken subcorpus is a series of interviews with learners of English at different education and proficiency levels conducted by academic teachers and PhD students from the English Philology department, University of Łódź.

Data from the British National Corpus was extracted through BNCweb, an online interface for the BNC XML edition written at Lancaster university. The PLEC corpus was accessed through the website of the project (www.pelcra.pl/plec), using the search engine for time-aligned spoken data.

Firstly, the most frequent n-grams were extracted from the PLEC corpus using the formulaic expression browser available at the website of the project:

Table 1. Formulaic expression browser extraction

#	N-gram	Frequency (A)	Independence (B)	B/A	Joint prob	Dispersion	FScore
1	uh_huh	572	59	0.10314685314685315	7.70975	2	52.877613205595935
3	i_don_t_know	310	42	0.13548387096774195	11.5108	29	52.38747509019103
5	for_instance	130	87	0.6692307692307692	8.32041	1	46.05645672539412
6	a_lot_of	178	46	0.25842696629213485	9.38713	30	44.4266802838018
7	for_example	207	51	0.2463768115942029	7.82196	30	43.1284277000129
9	in_poland	125	63	0.504	7.3005	28	39.76749931575577
11	thank_you_very_much	65	30	0.46153846153846156	15.602	9	39.007299814330146
13	i_think_it_s	104	43	0.41346153846153844	9.18749	52	38.74235697577031
15	as_well	77	72	0.935064935064935	7.91003	21	38.41968216096479

Next, 7 most frequent two- and three-word n-grams were selected from the list, with exclusion of *uh huh* for not being a lexical item and also *for instance*, because the phrase was used by just one person (dispersion=1). The same process was applied for the data extracted from BNCweb. Chosen n-grams appear in table 2:

Table 2. N-grams extracted from BNC and PLEC

#	BNC	PLEC
1	I don't know	I don't know
2	isn't it	a lot of
3	that's right	for example
4	or something	in the future
5	you know	in Poland
6	I mean	I think it's
7	come on	as well

The next step of the process was the analysis of audio samples containing the potential lexical bundles seen in the table above. Ten randomly selected (random selection is possible at the BNCweb, in PLEC each third example was chosen) samples of each of the potential lexical bundles were extracted from both corpora, resulting in 140 samples that were subjected to further analysis. The samples analyzed were recordings of language learners only; samples of speech of the interviewers were disregarded.

The data was imported into PRAAT software (Boersma and Weeknink 2014), where the n-grams were isolated from the sound samples. Then, the number of silent pauses was calculated for audio sample and checked against the maximum possible number of intersegmental pauses in the phrase; one possible pause in each two-word phrase, and two possible pauses in each three-word phrase.

3. Results

In the case of the learner English lexical bundles sound samples extracted from the PLEC corpus, there were 110 potential spots where intersegmental silent pauses within the lexical bundles could possibly occur. For the recordings extracted from the BNC spoken subcorpus, there were 80 possible spots for the occurrence of pauses. In PLEC recordings, pauses occurred in 28 of potential spots (24% of possible occurrences). In BNC recordings, pauses occurred in 11 of potential spots (14% of possible occurrences).

The p-value for this data was computed through a Fisher test using R software for statistical analysis, giving a result of p=0.06806, showing no statistical significance.

The pauses within the lexical bundles were fairly evenly distributed throughout all the samples analyzed, with the exception of the phrase *I don't know* in both corpora. In the BNC samples most of the pauses occurred in the phrase *I don't know*, whereas in the PLEC samples, the phrase contained no silent pauses whatsoever.

The duration of silent pauses was also measured. In the native English samples, pause duration mean value was 66ms, while in the learner English lexical bundles it was 37ms. The p-value for the Fisher test regarding the duration of pauses between native and learner English is 0,02409, showing statistical significance.

The ranks of occurrence of the n-grams from BNC in PLEC is as follows:

I don't know – 1

isn't it – X

that's right – 7173

or something – 63

you know – 62

I mean – 17

come on – X

(1 – most common n-gram in the corpus, X – n-gram not present in the corpus)

4. Conclusions and discussion

Several conclusions can be drawn from the results of this study. Firstly, it is apparent that the differences in distribution of silent pauses in both types of English analyzed are not significant, thus showing that formulaic sequences are likely to behave similarly in both cases, thus complying with what Wray (2004) wrote as far as pause occurrence is concerned, i.e. pauses are unlikely to occur, and if they do, they are fairly short.

What is interesting, the corpora have only one of the potential lexical bundles in common. This might result from methods used for teaching English in Poland as well as a limited set of interview topics used during the process of gathering data for the spoken part of the PLEC corpus. Such phrases as *for example* can be perceived to have become memorized holistically because of the learning method where the teacher elicits such phrases, even though they do not seem to be that frequent in native speaker English. Also, the format of the interviews, i.e. questions about the learners' plans for the future and about their relations with other cultures is most likely the reason for the occurrence of the phrases *in the future* and *in Poland*.

What is more, most of the n-grams present in native English are rare, or in some cases even nonexistent in Polish learner English (with the exception of *I don't know*). The way English is taught in Poland is most likely the factor to blame for this situation. Textbooks and teaching methods used in Polish schools do not take into account formulaicity and frequency in native language with regard to vocabulary selection. Instead, learners acquire different phrases, which become formulaic for Polish learner English. Such phrases like *for example* or *I think it's* definitely do improve the fluency of the learners' utterances, but at the same time they do not bring them any closer to sounding native-like. What could be done to improve this situation is implementation of the most common native English formulaic sequences into methodology programmes for teaching English – first at the level of input (listening), and then at the level of production as well.

As far as the only n-gram common for samples from both corpora is concerned, there are some interesting observations to be made. Interestingly, the phrase *I don't know* contains the most pauses of the native English bundles analyzed, while no pauses occur in this phrase in learner English. The reason for this polar difference seems to be the way in which the two groups of speakers under scrutiny use the phrase. When we look at the context in which the given phrase was used, native speakers tend to put it in an entirely literal context, exercising the full literal meaning of the phrase (e.g. *I don't know anything about music*), while Polish learners seem to use a completely different strategy; they are likely to insert the phrase into the syntactic clause, devoid of literal meaning (e.g. *somehow you have to, I don't know, maintain it*). This strategy of a "clause-breaking" use of the phrase *I don't know* seems to be used in order for the speakers to "buy themselves time" that they can use for processing the rest of the utterance. The same use of the lexical bundles is visible among native speakers as well, but they

choose different phrases in such contexts; e.g. the phrases *you know* or *or something* seem to easily adopt this "clause breaking" function. A possible reason for the presence of this phenomenon in the speech of Polish learners of English is that the Polish counterpart of the phrase – *nie wiem* – seems to very frequently adopt this function in the Polish language. Because of this, Polish learners subconsciously "calque" the function of the memorized sequence present in their L1 onto its English counterpart.

This is a preliminary study, which is a part of a larger on-going project. The study has several limitations which will be considered during further research in this field. The sample for the study will be extended, and selected so that the data for analysis is as objective as possible. Other phonological features of formulaic sequences, such as filled pauses and intonation contour are to be considered in the full study as well.

References

Allen, D. 2009. Lexical Bundles in Learner Writing: An Analysis of Formulaic Language in the ALESS Learner Corpus. *Komaba Journal of English Education* 1: 105–127.

Biber, D., S. Conrad, and V. Cortes. 2004. If you look at…: Lexical Bundles in University Teaching and Textbooks. *Applied Linguistics* 25 (3): 371–405.

Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow, Essex: Pearson Education Ltd.

Boersma, P. and D. Weenink. 2014. Praat: doing phonetics by computer [Computer program]. Version 5.4.01, retrieved 9 November 2014 from http://www.praat.org/.

Bolinger, D. 1976. Meaning and memory. *Forum Linguisticum* 1: 1–14.

Bybee, J. 2002. Phonological evidence for exemplary storage of multiword sequences. *Studies in Second Language Acquisition* 24: 215–21.

Bybee, J. 2006. From usage to grammar: the mind's response to repetition. *Language* 82(4): 711–33.

Chomsky, N. 1957. *Syntactic Structures*. Pairs: Mouton.

Dechert, H.W. 1983. How a story is done in a second language. In *Strategies in interlanguage communication,* eds. C. Faerch, and G. Kasper, 175–95. London: Longman.

Lehiste, I. 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa* 7: 107–122.

Lin, P.M.S. 2010. The phonology of formulaic sequences: A review. In *Perspectives on formulaic language: Acquisition and communication*, ed. D. Wood, 174–193. London: Continuum.

Millar, N. 2009. Assessing the processing demands of learner collocation errors. Poster presented at Corpus Linguistics 2009, Liverpool, UK.

Moon, R., 1998. *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press.

Myles, F., R. Mitchell, and J. Hooper. 1999. Interrogative Chunks in French L2: A Basis for Creative Construction? *Studies in Second Language Acquisition* 21: 49–80.

Pawley, A., Syder, F. 1983. "Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency." *Language and Communication 191: 225.*

Peters, A.M., 1977. Language learning strategies: does the whole equal the sum of the parts? *Language* 53(3), 560–573

Peters, A. 1983. *The Units of Language Acquisition*. Cambridge: Cambridge University Press.

Pęzik, P. 2012. Towards the PELCRA Learner English Corpus. In *Corpus Data across Languages and Disciplines, Lodz Studies in Language. Vol. 28,* ed. P. Pęzik. Frankfurt am Main: Peter Lang. Forthcoming.

Plunkett, K., and S. Stromqvist. 1990. The Acquisition of Scandinavian Languages. *Gothenburg Papers in Theoretical Linguistics* 59.

Raupach, M. 1984. Formulae in second language speech production. In *Second language productions*, eds. H.W. Dechert, D. Möhle, and M. Raupach, 114–37.Tubingen: Gunter Narr.

Sinclair, J. 1991. *Corpus, concordance, collocation: Describing English language*. Oxford: Oxford University Press.

The British National Corpus, version 3 (BNC XML Edition). 2007. Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Retreived from http://www.natcorp.ox.ac.uk/, 17.11.2014.

Van Lancker, D., J. Canter, and D. Terbeek. 1981. Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech and Hearing Research,* 24: 330–335.

Wray, A. 2004. 'Here's one I prepared earlier': formulaic language learning on television. In *The acquisition and use of formulaic sequences*, ed. N.Schmitt, 249–268. Amsterdam: John Benjamins.

Mandeville's Travels and the study of Middle English word geography: a corpus-based analysis of selected verbs

http://dx.doi.org/10.18778/8088-065-8.06

Paulina Rybińska

University of Lodz

Abstract

This paper presents an example of a study that deals with lexical choices in the Middle English period. It aims to first investigate whether the choice of the verbs in the two regionally distinct versions of the Late Middle English book *Mandeville's Travels* is text-dependent or region-dependent, which would then show to what extent the results of the comparison may be observed in other Middle English texts. In addition, it checks whether the choice of the verbs is influenced by their etymology. This study in progress is hoped to partially contribute both to the field of Middle English word geography and to the examination of the aforementioned text in general.

1. Historical dialectology

As stated by Laing and Lass (2006: 418), "[h]istorical dialectology is simply historical linguistics with a spatial emphasis; in the same sense, historical linguistics is simply linguistics with a temporal emphasis." Particularly, this area of linguistics pays attention to linguistic changes which occurred in certain time in the history and across different regions.

The main difference concerning present-day and historical dialectology is that the data in the latter case can only be collected on the basis of the existing written sources (Fisiak 1982). First of all, many texts or manuscripts did not survive till present. Secondly, the preserved texts reflect the written, not the spoken mode. Moreover, we do not have many information, from the so-

ciolinguistic point of view, about the people who wrote them. Despite of this, the tools used for historical dialectology make this discipline "an important laboratory for the achievement of increasingly reliable findings" (Dossena and Lass 2009: 7).

Fisiak (1982) proposes to divide the functions of historical dialectology into static (synchronic) and dynamic. Static functions aim at "establishing isoglosses (consequently producing atlases) for certain periods in the history of a language" (ibid.: 118). On the other hand, dynamic functions are directed towards giving information about changes regarding, for instance, dialect boundaries.

One of the major principles when studying languages' development is to put the emphasis on context – either regional or social (Dossena and Lass 2004). Even though these factors assume a divergent point of view on the study of language, they do not exclude each other (Meurman-Solin 2000a, 2000b). Actually, one should be aware of the fact that the influence of both of them on linguistic data is of crucial importance (Dossena and Lass 2004).

2. Middle English dialect situation

The linguistic situation in England after the Norman Conquest was complex. Despite of the fact that the prestigious functions of language were taken over by French and Latin, English was still spoken, mainly by the lower classes of society. The fact that there was no "standard" variety is the main reason for Middle English dialectal diversity (Crystal 2004). Scribes were trying to illustrate the way people spoke, so it can be assumed that the great variation concerning spelling is a reflection of people's pronunciation (McIntyre 2009). According to Crystal (2004: 212), "within each manuscript there will be distinctive linguistic features – spellings, words, and grammatical constructions – that can be assumed to be diagnostic of their locality."

In the 14th century, John Trevisa, a Cornish writer, distinguished between "Southeron, Northeron, and Myddel speche" of Middle English (Burrow and Turville-Petre 1996: 6). In fact, it is the least complicated division that can be made. Historical dialectologists suggest the division into five conventional dialect areas. Conventional, because of the great dialectal diversity that existed during those times.

The map representing the Middle English dialect areas is shown below:

Map 1. Middle English dialects of England (http://kids.britannica.com/comptons/art-143574/Middle-English-dialects-of-England)

In the preface to *Eneydos* translation, William Caxton included a comment which may be treated as an early attempt to characterize the variability in the English language during standardisation. "Loo, what sholde a man in thyse dayes now wryte, *egges* or *eyren*? Certaynly it is harde to playse euery man by cause of dyuersite & chaunge of langage." According to Caxton's comment, the most noticeable contrast was seen between northern and southern regions – not only distinct from each other in space, but also subjected to diverse linguistic influences. The word of Old Norse origin, *egg*, was predominant in the north, whereas the native form *ei* was much more popular in southern dialects.

The issue of the north-south divide has already been recognized in the writings of William of Malmesbury (1125). Moreover, throughout the whole Middle English period writers commented on people's attitudes towards these two dialects. In 1387, Ranulph Higden wrote in his *Polychronicon*:

> All the speech of the Northumbrians, and especially at York, is so harsh, piercing, and grating, and formless, that we Southern men can hardly understand such speech. I believe that this is because it is near to outlandish men and foreigners, who speak in a foreign language, and also because the kings of England always live far away from that region (from the translation of 1387 Ranulph Higden's *Polychronicon*; Crystal 2004: 216/217).

Not only southerners had problems with comprehension of the northern variety. The anonymous author of *Cursor mundi*, a chronicle written in the north, mentions that he had to translate a fragment of text written in southern English into his own variety, so that his readers could understand it: "In sotherin englis was it draun. And turnd it haue I till our aun language o northrin lede, Þat can nan oiþer englis rede" (Crystal 2004: 207).

The differences between the abovementioned dialects concerned not only spelling and grammatical structures, but also lexicon. Interestingly enough, research in dialect vocabulary of Middle English has been perceived by many scholars as a neglected field of historical linguistics (Kaiser 1937; McIntosh 1973, 1978; Benskin and Laing 1981; Hudson 1983; Hoad 1994; Fisiak 2000). As Fisiak (2000) states, "[h]istorical word-geography of English, particularly Old and Middle English, is still a very much underdeveloped area of research although its importance has been recognized for a long time."

According to Black (2000: 455), "[w]ord geography may fairly uncontroversially be defined as the mapping out of words across space: the part of dialectology that deals with lexis." One of the first representative investigations concerning lexical choices in the Middle English period was conducted by Rolf Kaiser in 1937. He examined how some words in the northern version of *Cursor mundi* were replaced by a southern scribe. Kaiser assumed that the replacement was due to the fact that certain lexical items were not present or not widely used in the southern dialect (Hoad 1994). Another approach to word geographical studies is represented by the scholars compiling the *Linguistic Atlas of Late Mediaeval English*. The assemblage of as many linguistic data as possible is the basis for further observations regarding the occurrence of regionally distinct lexicon (Hoad 1994).

One of the possible reasons why Middle English word-geography studies are regarded as a neglected field of historical linguistics is related to the problem of establishing strict criteria when conducting such research. Due to Middle English dialectal diversity, we discover a plethora of various forms which are sometimes ambiguous in terms of classification; in other words, the assignment of these

forms to particular linguistic categories, such as phonology (in Middle English represented through spelling), morphology, or lexis may be disputable. As Benskin and Laing (1981: 94) state, "[i]t may of course be arguable whether a given item represents translation lexical rather than orthographic, orthographic rather than morphological, or morphological rather than lexical." For example, some linguists decide to treat the difference between forms such as 'kirk' and 'chirch' as orthographic, but others perceive this as a lexical difference. To be sure, in such cases it is of greater importance to look for different word-forms, and exclude those forms which do not contribute to the studies concerning lexical choices. To sum up, the field of historical word-geography studies is certainly problematic when it comes to methodological issues, mainly when one attempts to define what may constitute a lexical difference as opposed to spelling or morphological variation.

3. Methodology

The study aims to first investigate whether the choice of the verbs in the two regionally distinct versions of *Mandeville's Travels* is text-dependent or region-dependent, and then to check if this phenomenon is determined by words' etymology or is independent of it.

For the purpose of the study, the following thesis statement was formulated: The choice of the verbs in the two regionally distinct versions of *Mandeville's Travels* is region-dependent and influenced by words' etymology. Furthermore, the following research questions were asked:

1) *Does the choice of the verbs depend on region?*

In other words, it will be investigated whether the choice of the verbs in the two regionally distinct versions of one text, namely *Mandeville's Travels*, is text-dependent or region-dependent. Then, it will be checked if the results from *MT* can also be noticed in other chosen Middle English texts from different regions.

2) *Is this phenomenon determined by words' etymology or is independent of it?*

Because of the influence of Old Norse, it is expected that borrowings from this language will be predominant in the chosen northern texts. On the other hand, French borrowings are assumed to be more widespread in texts from the south. According to Wardale (1958: 40), "it may be said that Old Norse came in first in the north-east and north, French in the south and south-east." As Crystal (2004: 148) states, "[t]he

loans took their time to move north: in early Middle English there were far more French loans in southern texts, and an even spread does not emerge until the later period." Presumably, native Old English words might appear both in the north and in the south. To be more precise, the following situations are expected:

1) words of Old Norse would appear predominantly in the north
2) words of Old Norse origin wouldn't appear exclusively in the south
3) words of native origin would appear both in the north and in the south
4) words of French origin would appear predominantly in the south
5) words of French origin wouldn't appear exclusively in the north

First stage of the analysis involved the comparison of the first three chapters of two geographically distinct versions of *Mandeville's Travels*: northern – London, British Library, Egerton 1982, and southern (written in the East Midland dialect) – London, British Library, Cotton Titus C.16. According to *Linguistic Atlas of Late Mediaeval English* (*LALME*), the northern MS. is localized in North Riding of Yorkshire; the southern one – in Hertfordshire.

Map 2. *Egerton* 1982 – Yorkshire (NRY) (http://en.wikipedia.org)

Map 3. *Cotton Titus* C. 16 – Hertfordshire (http://en.wikipedia.org)

Mandeville's Travels or *The Book of John Mandeville* was written in French c. 1356. The authorship is unknown. It comprises various stories of people and places during the journey from Europe to Jerusalem and Asia. The great popularity of the book resulted in many translations. After some time, a copy of the

French version was carried into England. The oldest English translation is known as the Defective version, which constitutes the source text for two regionally different copies: *Cotton* and *Egerton*. Probably, the northern version was transcribed from the southern one. Both copies are dated c. 1420 (Seymour 2002).

During the second stage of the analysis four representative texts both from the northern and East Midland regions (approximately the same time period) were chosen for further examination. When choosing the texts, geographical division made by Anna Hebda in one of her articles (2010) was mostly used as a basis.

Table 1. Selected northern and East Midland texts (based on Hebda 2010)

North	East Midland
The wars of Alexander (Ashmole 44)	*Guy of Warwick* (Auchinleck)
The pricke of conscience (Glb E. ix & Hrl)	*Confessio amantis* (Frf 3)
Works by Rolle	*Merlin* (Cmb Ff.3.11)
Mandeville's *Travels* (Egerton 1982)	Mandeville's *Travels* (Cotton Titus C. XVI)

As can be seen, both versions of the whole *Mandeville's Travels* book are also present in this classification.

First of all, the two versions of *Mandeville's Travels* had to be copied into *Excel* and arranged into two columns. Thanks to this, it was possible to observe variation between these two texts. So far, the first three chapters of each version were analyzed (c. 7200 words). Since the study is devoted to lexical choices, obvious spelling and grammatical/morphological variation was ignored. As 'lexical differences' I understand words/phrases that were used as equivalents in exactly the same contexts in both versions. For example, the difference between *thurgh* and þorgh will be treated as spelling rather than lexical one, whereas between *wenden* and *gon* – as lexical. An exemplary comparison is shown below:

Table 2. Exemplary comparison

North	South
he	he
will,	wole
wende	***go***
thurgh	þorgh
Almayne	Almayne

For the purpose of this article, twelve verb pairs have been chosen for further analysis. Originally, the main criterion was to choose the most frequently appearing verbs pairs, but because of the fact that only two verb pairs occurred more than one time (*callen-clepen* – nine times, *opposen-examynen* – three times), I decided to select them randomly: from the first chapter – *wenden-gon*, *callen-clepen*, *taken-receiven*, *gon-entren*; from the second chapter – *waten-knowen*, *stirren-meven*, *opposen-examinen*, *forsaken-denien*; from the third chapter – *grauen-beryen*, *trowen-hopen*, *lousen-assoilen*, *okeren-vsuren*[1].

The next stage of the research involved qualitative analysis. The aim of the qualitative analysis was to determine whether a given word may be treated as northern or southern by checking its etymology in the *Middle English Dictionary* or the *Online Etymology Dictionary*.

What is more, due to the availability of the electronic version of the *Linguistic Atlas of Late Mediaeval English*, the distribution of spelling variants of some verbs was possible to be found and then shown using dot maps. Unfortunately, the *Atlas* does not provide maps for the majority of verbs that were chosen for the analysis. Thus, quantitative analysis had to be conducted. The aim was to compare the distribution of the words in question in the chosen representative texts from the north and the East Midland region, and to check whether the words labeled as 'northern' appear also in southern texts, and if those labeled as 'southern' appear in the north. If such instances occur, possible explanations will be provided. This part of the analysis was conducted using the *Corpus of Middle English Prose and Verse*. It is a collection of digitized copies of about 150 works in Middle English, assembled from a number of sources including University of Michigan faculty, the Oxford Text Archive, and the Humanities Text Initiative. To be sure, only the selected texts (Table 1) were searched through using the *Corpus*.

1 The chosen verb pairs are listed as infinitives and according to their order of appearance in the texts.

4. Results and analysis

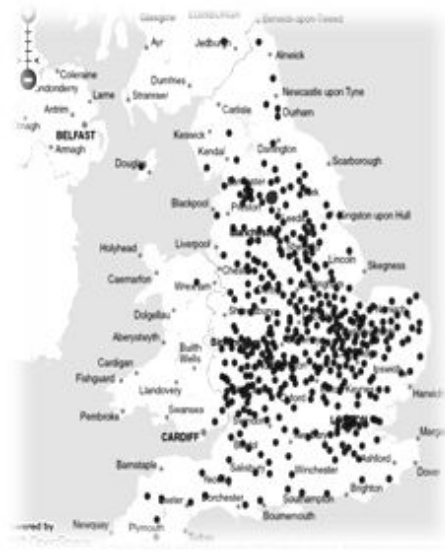
4.1. *Wenden/Gon*

Wenden (v.) [OE wendan, wændan, (Nhb.) woendan]. According to *Etymonline*, "wend (v.) *to proceed on*, Old English wendan *to turn, go.*

Gon (v.) [OE gān; sg. 2 gǣst; sg. 3 gǣþ; pl. & impv. pl. gāþ; inflected inf. tō gānne; p.ppl. -gān. In OE, the past forms are usually supplied by ēode & gangan; in ME, they are supplied by yēde, gangen, wenden, q.v.]

Wenden is expected to be predominant in the north rather than in the south, while *gon* – to be more popular in the south.

The distribution of both forms was checked in *eLALME*. The results are shown on the maps below:



Map 4. *Wenden* – distribution (generated from http://www.lel.ed.ac.uk/ihd/elalme/elalme.html.)

Map 5. *Gon* – distribution (generated from http://www.lel.ed.ac.uk/ihd/elalme/elalme.html.)

As can be seen, *wenden* is distributed across the whole country. *Gon* is predominant in the south. However, it is worth mentioning that northern variants with <a> are not present in the *Atlas* (forms such as *ga*, *gas*, etc.). Nevertheless, when searching through the *Corpus* the variants with <a> were included as well.

The results of the quantitative analysis actually differ from previous expectations as far as the northern texts are concerned.

Table 3. *Wenden/gon* – frequency (N)

Northern texts	Wenden		Gon	
Mandeville's Travels (N)	50	26%	139	74%
The wars of Alexander	41	68%	19	32%
The pricke of conscience	71	58%	51	42%
Works by Rolle	35	51%	33	49%
Total	**197**	**51%**	**242**	**49%**

Table 4. *Wenden/gon* – frequency (EM)

East Midland texts	Wenden		Gon	
Mandeville's Travels (S)	4	1%	291	99%
Guy of Warwick	186	34%	357	66%
Confessio amantis	88	15%	493	85%
Merlin	112	17%	539	83%
Total	**390**	**17%**	**1680**	**83%**

Both in the northern and East Midland texts we can notice a preference for *gon* rather than *wenden* forms. It is not surprising, because it was *gon* which eventually became a standard form. What is particularly visible is that *gon* in the first table outnumbers *wenden* forms only in one text, namely *Mandeville's Travels*. Presumably, the northern manuscript in this case demonstrates great affinity with the East Midland manuscript it was transcribed from.

4.2. *Callen/clepen*

Callen (v.) is a borrowing from Old Norse. Thus, it is expected to be found in northern rather than in East Midland texts.

It means particularly *to call (sth. by a certain name)*, *name (sb. sth.)*, *call (sb. good, etc.)*; *ppl. called, named*; *(b) ~ bi (to) name*, *to call (sb.) by (a name)*. The other meaning is *to summon (sb.)*, *call (sb. to a place)*, *call (sb. to do sth.)*.

Clēpen (v.) [A *cliopian*, *cleopian*, from West Saxon *clipian*, *clypian*.]. This word is expected to be characteristic of the south rather than of the north.

Again, two meanings of this word: *to apply (a name, epithet, title, expression, or designation to sb. or sth.)*, *name (sb. so and-so)*, and *to ask*, *request*, or order *(sb.) to appear (in a place or in someone's presence)*; *summon, call, send for*; *invite*.

Possible spelling variants of these two verbs can be found in *eLALME*:

Map 6. *Callen* – distribution (generated from http://www.lel.ed.ac.uk/ihd/elalme/elalme.html)

Map 7. *Clepen* – distribution (http://www.lel.ed.ac.uk/ihd/elalme/elalme.html)

As can be seen on the maps, the distribution of the variants suggests that *call* was more common in the north with a tendency to spread downwards. *Clepen* was certainly more popular across southern regions. Red dots denote provenance of the northern and East Midland versions of *Mandeville's Travels*.

The results of the quantitative analysis are presented in the tables below:

Table 5. *Callen*/*clepen* – frequency (N)

Northern texts	Callen		Clepen	
The wars of Alexander	46	100%	–	–
The pricke of conscience	56	100%	–	–
Works by Rolle	51	96%	2	4%
Mandeville's Travels (N)	393	100%	–	–
Total	**546**	**99%**	**2**	**1%**

As far as the northern texts are concerned, we have an overall number of only two instances of the use of the southern variant *clepen*. To compare, *callen* is used 546 times. The instances of *clepen* can be found in *Works by Rolle* in the examples shown below:

IS CLEPED

(..) þe first degre **is cleped** Insuperable, þe tother Inseparable, þe thridde Singuler.

(…) and þerfore it **is cleped** inseparable, for it may not be departed fro thoght of Ihesu Crist.

(Ms. Rawlinson A 389 [folio 81])

ES CALLED

Þe fyrst degre **es called** Insuperable. Þe secund Inseparabl. Þe third Singuler.

(Ms. Rawlinson C 285 [folio 40])

Surprisingly, we can see inconsistency, because the author used two various forms in exactly the same contexts.

Let us move on to the analysis of the East Midland texts:

Table 6. *Callen*/*clepen* – frequency (EM)

East Midland texts	Callen		Clepen	
Guy of Warwick	8	9%	77	91%
Confessio amantis	46	23%	153	77%
Merlin	29	12%	215	88%
Mandeville's Travels (S)	90	20%	353	80%
Total	**173**	**16%**	**798**	**84%**

As can be seen in the table, the southern variant *clepen* is much more frequent, but 173 contexts with the northern variant appeared. Most of these cases are not surprising, since it was the northern variant that eventually became a standard form. But let us look at some specific examples from *Guy of Warwick*:

CALL

And commaunded his dukes and barons **all**

To bee redy in armes at euery **call**.

CALLE

Þou art me leuest of oþer **alle,**

For þi 'leman' ichil the **calle;**

CALLED

ON WITSONDAYE **called** Pentecoste (title)

And Guye seide, 'my fader **is called** Sywarde

In *Guy of Warwick*, both *call* and *calle* appeared because of rhyming. In the examples presented above, *call* had to rhyme with *all*. Apart from this, *called* was used two times in this text: first –in the title, second – in the statement *my fader is called Sywarde*. As the present study is not a diachronic one, it cannot be proved that the phrase *is called* was becoming more and more popular in the south and that it was gradually replacing the older form *is cleped*. Nevertheless, quantitative analyses show that there existed variation in the southern texts as regards these two verbs.

In *Confessio amantis* all the instances of the northern variant are a matter of rhyming (with such words as *alle*, *befalle*, *falle*, *withalle*, *halle*, and *befalleth*, *falleth*). An interesting example is the one presented below. Both southern and northern forms are used here:

Noght upon on, bot upon alle

It is that men now **clepe and calle**

In *Merlin*, 29 instances may indicate a change in progress (fixed phrase *is called*). Of course, the southern form *clepen* is still predominant – 215 instances.

Finally, the southern version of *Mandeville's Travels*:

> CALLE
>
> And it was wont to be **clept** Collos & so **calle** it the Turkes 3it
>
> (…) but 3if þat the Emperour **calle** ony man to him þat him list to speke with ałł.
>
> **CALLED (70)** – either change in progress ('is called') or editor's notes.

As can be seen, *calle* was used twice: in the first example – probably to avoid repetition with *clept*; in the second example – we have different meaning of *call*, namely *to summon sb.*

70 instances of *called* indicate either change in progress – high frequency of the phrase *is called*, or editor's notes, written in present-day English, which should be excluded in the future so as not to influence the results.

To sum up, qualitative analysis proves that *callen* can be treated as a predominantly northern word, while *clepen* – as a southern one. Furthermore, it is not surprising that the northern form *callen* can also be found in the East Midland texts – this phenomenon may point to variation in this region. However, many of the uses of *callen* in the chosen East Midland texts are a matter of rhyming.

4.3. *Taken/Receiven*

Tāken (v.) [LOE **tacan**, p.sg. **tōc**, pl. **tōcon**, from ON (cp. OI **taka**, pr.sg. **tek**, **tekr**, p.sg. **tōk**, pl. **tōku**, ppl. **tekinn)**.]

Receiven (v.) [OF **recevoir**, **recever**, **receivre**, **recivre**, **rechever**]

The meaning is, among others, *to take (a material object) into one's hand or possession*, *accept possession of (sth.).*

The map in *eLALME* was available only for *taken*. As can be seen, not many forms were found. They are visible predominantly in the north.

Map 8. *Taken* – distribution (generated from http://www.lel.ed.ac.uk/ihd/elalme/elalme.html)

The results of the quantitative analyses are as follows:

Table 7. *Taken*/*receiven* – frequency (N)

Northern texts	Taken		Receiven	
Mandeville's Travels (N)	121	100%	–	–
The wars of Alexander	40	100%	–	–
The pricke of conscience	71	97%	2	3%
Works by Rolle	94	84%	18	16%
Total	**326**	**95%**	**20**	**5%**

Table 8. *Taken*/*receiven* – frequency (EM)

East Midland texts	Taken		Receiven	
Mandeville's Travels (S)	131	99%	1	1%
Guy of Warwick	148	99%	1	1%
Confessio amantis	395	89%	47	11%
Merlin	410	96%	19	4%
Total	**1084**	**96%**	**68**	**4%**

Taken is significantly more frequent in the north, but also in the East Midland region. This may testify to the growing popularity of this Old Norse borrowing across the whole country and its later standardization.

4.4. *Gon/entren*

Gon (v.) [OE gān; sg. 2 gǣst; sg. 3 gǣþ; pl. & impv. pl. gāþ; inflected inf. tō gānne; p.ppl. -gān. In OE, the past forms are usually supplied by ēode & gangan; in ME, they are supplied by yēde, gangen, wenden, q.v.]

Entren (v.) [OF **entrer**]

It means *to enter into a confined space or a situation*.

Table 9. *Gon/entren* – frequency (N)

Northern texts	Gon		Entren	
The wars of Alexander	19	44%	24	56%
The pricke of conscience	51	93%	4	7%
Works by Rolle	33	100%	–	–
Mandeville's Travels (N)	139	85%	24	15%
Total	**242**	**80,5%**	**52**	**19,5%**

Table 10. *Gon/entren* – frequency (EM)

East Midland texts	Gon		Entren	
Guy of Warwick	357	99%	1	1%
Confessio amantis	493	100%	–	–
Merlin	539	79%	141	21%
Mandeville's Travels (S)	291	80%	73	20%
Total	**1680**	**89,5%**	**215**	**10,5%**

According to the qualitative analyses, *entren*, despite its Old French origin, is not predominant in the southern texts. Moreover, even the basic text is not consistent as regards the appearance of the abovementioned verbs. It would be interesting, then, to look at specific examples in the future.

4.5. *Waten/Knowen*

The information about ***waten*** was impossible to find in *MED*. It is a very old word that we know from such Old English sentence as 'Ic nat' ('I don't know').

Knouen (v.) [OE **cnāwan**, **on-**, **ge-**, **tō-**; p. **-cnēow**; ppl. **-cnāwan**. In ME, ***ou*** occurs in present forms & p.ppl. of the M & S dialects, perh. also (rarely) in the N dialect; ***au*** (from ***ou***) occurs in the present forms of M, N, & K, in the p.ppl. of M, N, & S. P. forms are often used with a present subjunctive sense.]

The results of the quantitative analyses are shown in the tables below:

Table 11. *Waten/knowen* – frequency (N)

East Midland texts	Waten	Knowen	
The wars of Alexander	6	10%	6
The pricke of conscience	27	20%	27
Works by Rolle	47	46%	47
Mandeville's Travels (N)	30	37,5%	30
Total	**110**	**28%**	**110**

Table 12. *Waten/knowen* – frequency (EM)

East Midland texts	Waten	Knowen	
Guy of Warwick	–	54	100%
Confessio amantis	–	216	100%
Merlin	–	457	100%
Mandeville's Travels (S)	–	83	100%
Total	**0**	**810**	**100%**

Surprisingly, no instances of *waten* were found in the East Midland texts. It means that this word might have simply become obsolete. Some traces of this form can be observed in the north, but there is a preference for *knowen* anyway.

4.6. *Stiren/meven*

Stiren (v.) [OE styrian, stirian, stirgan].

It means *to change the location of (sth.)*, *move*, *shift*; *dislodge (sth.)*.

Meven (v.) [OF movoir, meuvre, muevre, moevre & AF moveir, muve(i)r; also cp. L movre.]. It means *to move (sb. or sth.)*, *shift*; *remove (sth.)*, *dislodge*; *move (sth.) about*.

It is expected that *stiren* would be predominant in the north, and *meven* – in the south.

The distribution of the words in question is presented below:

Table 13. *Stiren/meven* – frequency (N)

Northern texts	Stiren		Meven	
The wars of Alexander	7	29%	17	71%
The pricke of conscience	5	38%	8	62%
Works by Rolle	32	94%	2	6%
Mandeville's Travels (N)	12	75%	4	25%
Total	**56**	**59%**	**31**	**41%**

As the table shows, *stiren* is predominant in the north, but *meven* is also frequent. In the first two texts it even exceeds the number of *stiren* forms. An interesting example of the use of the two variants can be found in *Mandeville's Travels (N)*:

> for men may see þare þe erthe of þe toumbe many a tyme stirre and moue, as þer ware a qwikke thing vnder.

The table below shows the distribution of *stiren/meven* in the East Midland texts:

Table 14. *Stiren/meven* – frequency (EM)

East Midland texts	Stiren		Meven	
Guy of Warwick	–	–	–	–
Confessio amantis	–	–	1	100%
Merlin	1	2%	61	98%
Mandeville's Travels (S)	6	55%	5	45%
Total	**7**	**99%**	**67**	**91%**

What can be seen is that *meven* is generally more frequent (67 instances). Both *stiren* and *meven* are not present in *Guy of Warwick*. *Meven* visibly outnumbers *stiren* in *Merlin*. One instance of it is present in *Confessio amantis*. Surprisingly, *Mandeville's Travels (S)* book is not consistent as far as the choice between the two forms is concerned.

4.7. *Opposen/Examinen*

Opposen (v.) [OF **oposer**] According to *MED*, it means *to question or interrogate (sb.); examine (heart, conscience, confession); ask (sb.) a question; -- also without obj.; (b) to accuse (sb.) of (sth.), charge; (c) to torment (sb.); (d) to examine or audit (sth.).*

Examinen, **-ene(n** (v.) Also **exam(p)nen** [OF **examiner**, L **exāmināre**]. It means *(a) to investigate, examine (something); to scrutinize, consider critically, appraise.*

Taking into account the words' origin, it is expected that both forms will appear more frequently in the south, with no visible discrepancies between their distribution.

Table 15. *Opposen/examinen* – frequency (N)

Northern texts	Opposen		Examinen	
The wars of Alexander	–		–	
The pricke of conscience	–		–	
Works by Rolle	–		–	
Mandeville's Travels (N)	3	75%	1	25%
Total	**3**	**75%**	**1**	**25%**

Table 16. *Opposen/examinen* – frequency (EM)

East Midland texts	Opposen		Examine	
Guy of Warwick	–	–	–	–
Confessio amantis	22	96%	1	4%
Merlin	1	25%	3	75%
Mandeville's Travels (S)	9	53%	8	47%
Total	**32**	**58%**	**12**	**42%**

As the tables show, the words in question were present only in *Mandeville's Travels (N)*. *Opposen* appeared 3 times, while *examinen* – only once. However, in the East Midland texts both forms are more popular, with *opposen* being predominant. It is especially visible in *Confessio amantis*. The two versions of *Mandeville's Travels* are not consistent.

4.8. *Forsaken/Denien*

Forsāken (v.) P. forsōk, -sūk (N); ppl. forsāken, -sāked. [OE forsacan, forsōc; cp. also OE sacan *contend*, *disagree*, *accuse*, etc.].

Dēnīen (v.) Also denaien, denoien, disnoien. [OF denïier, deneiier, denoiier, desnoier (from L dēnegāre)].

It is expected that *forsaken* will be characteristic of the north; *denien*, being a French borrowing – characteristic of the East Midland region.

The following tables show the frequency of the words in question in the selected texts from the northern and East Midland regions.

Table 17. *Forsaken/denien* – frequency (N)

Northern texts	Forsaken		Denien	
The wars of Alexander	6	100%	–	–
The pricke of conscience	15	94%	1	6%
Works by Rolle	32	100%	–	–
Mandeville's Travels (N)	8	80%	2	20%
Total	**61**	**93,5%**	**3**	**6,5%**

Table 18. *Forsaken/denien* – frequency (EM)

East Midland texts	Forsaken		Denien	
Guy of Warwick	4	100%	–	–
Confessio amantis	44	100%	–	–
Merlin	24	83%	5	17%
Mandeville's Travels (S)	6	86%	1	14%
Total	**78**	**92%**	**6**	**8%**

As can be seen, *forsaken* is considerably more frequent in the north than *denien* (61 vs 3 instances). On the other hand, the high frequency of *forsaken* forms in the south may testify to its being popular in the East Midland region from the very beginning or its growing popularity.

4.9. *Grauen/beryen*

Grāven (v.) – [OE **grafan; grōf, grōfon; grafen.**]

It means *(a) To bury (a corpse), place (sb.) in a grave; fig. to swallow up (a damned soul); (b) to put (sth.) under the ground, cover with earth, bury.*

According to the *Online Etymology Dictionary*, OE *buryen* (v.) means *to raise a mound, hide, bury*, akin to beorgan "to shelter," from Proto-Germanic *burzjan- "protection, shelter".

These two lexical units of native origin are expected to be found in both regions.

Table 19. *Grauen/beryen* – frequency (N)

Northern texts	Grauen		Beryen	
The wars of Alexander	–	–	–	–
The pricke of conscience	–	–	–	–
Works by Rolle	4	80%	1	20%
Mandeville's Travels (N)	18	100%	–	–
Total	**22**	**90%**	**1**	**10%**

Table 20. *Grauen/beryen* – frequency (EM)

East Midland texts	Grauen		Beryen	
Guy of Warwick	3	33%	6	67%
Confessio amantis	–	–	–	–
Merlin	–	–	–	–
Mandeville's Travels (S)	–	–	22	100%
Total	**3**	**16,5%**	**28**	**83,5%**

As can be seen, generally, there is a preference for *grauen* forms in the north and for *beryen* forms in the south. The discrepancy between the *Mandeville* versions is caused probably due to the lacunae in the *Egerton* manuscript. Interestingly, as shown in the example below, on one occasion the phrase *puttez him in þe erthe* was used additionally:

> (…) and, when he es deed, þai bere him in to þe felde and ***puttez him in þe erthe*** (Egerton MS.)

> (…) and whan he drawetħ towardes the detħ euery man fleetħ out of the hous tiłł he be ded & after þat þei buryen him in the feldes (Cotton MS.)

4.10. *Trowen/hopen*

Trouen (v.) – [OE **trūwian**, **trūwigan**, impv. **trūa** & **trēowan**, **trȳwan** & **trēowian**, **trēowigan**, **trȳwian**].

According to *MED*, *trouen* means, among others, *(a) To have trust, be trustful; rely (on sb. or sth.), place one's confidence (in sb.), trust (in God).*

Hopen (v.) – [OE **hopian**]

It means *(b) to have trust, have confidence; assume (sth.) confidently, presume; trust (that sth. is the case); trust (to have sth.).*

Because of the words' native origin, they are expected to be distributed similarly in the north and in the south.

Table 21. *Trowen/hopen* – frequency (N)

Northern texts	Trowen		Hopen	
The wars of Alexander	21	55%	17	45%
The pricke of conscience	28	90%	3	10%
Works by Rolle	24	52%	22	48%
Mandeville's Travels (N)	60	98%	1	2%
Total	**133**	**74%**	**43**	**26%**

Table 22. *Trowen/hopen* – frequency (EM)

East Midland texts	Trowen		Hopen	
Guy of Warwick	18	72%	7	28%
Confessio amantis	26	70%	11	30%
Merlin	60	94%	4	6%
Mandeville's Travels (S)	35	97%	1	3%
Total	**139**	**83%**	**23**	**17%**

According to the tables presented above, *trowen* forms are more popular both in the northern and southern texts. Interestingly, in Gower's *Confessio amantis* both forms appeared together:

I speke it forth and noght ne leve:

And thogh it be beside hire leve,

I **hope and trowe** natheles

That I do noght ayein the pes.

4.11. *Lousen/assoilen*

Lōsen (v.) – this verb comes from the adjective **lōs, which is of Old Norse origin.**

According to *MED*, it means *(a) To free (sb. from physical constraint, prison, hell, etc.); untie (an ass, a dog), (b) to free (sb. from someone's control, from an obligation, from sin or distress, etc.); release (sb. from the cloister); absolve (sb.) from sin.*

Assoilen (v.) Also ***asoili(e***, ***as(s)oli***. [OF **assoiler**, **-ir, assolir**, **-ier**].

(a) To absolve (sb.) of sin by divine or sacerdotal authority; grant (sb.) remission of sins or penance.

The first verb is expected to be found in the northern texts, the second – in the southern ones.

Table 23. *Lousen/assoilen* – frequency (N)

Northern texts	Lousen		Assoilen	
The wars of Alexander	–	–	–	–
The pricke of conscience	5	62,5%	3	37,5%
Works by Rolle	2	100%	–	–
Mandeville's Travels (N)	2	100%	–	–
Total	**9**	**87,5%**	**3**	**12,5**

Table 24. *Lousen/assoilen* – frequency (EM)

East Midland texts	Lousen		Assoilen	
Guy of Warwick	–	–	1	100%
Confessio amantis	–	–	3	100%
Merlin	–	–	2	100%
Mandeville's Travels (S)	–	–	2	100%
Total	**–**	**–**	**8**	**100%**

As shown in the tables, the word of Old Norse origin, *lousen*, did not appear in the southern texts at all. As far as the northern texts are concerned, *assoilen* was found only in *The pricke of conscience*, but the instances of *lousen* still prevail.

4.12. Okeren/vsuren

Okeren (v.) – from **ōker** (n.1) of Old Norse origin; *(a) To lend (money, goods) at interest; (b) to make a loan (to sb.) at interest.*

Vsuren (v.) – from Old French ***usurer***; *To make a loan at interest, practice usury; lend money at interest (to sb.).*

Taking into account the words' etymology, *okeren* might be predominant in the north, whereas *vsuren* – in the south.

Table 25. *Okeren/vsuren* – frequency (N)

Northern texts	Okeren		Vsuren	
The wars of Alexander	–	–	–	–
The pricke of conscience	–	–	–	–
Works by Rolle	–	–	–	–
Mandeville's Travels (N)	1	100%	–	–
Total	**1**	**100%**	**–**	**–**

Table 26. *Okeren/vsuren* – frequency (EM)

East Midland texts	Okeren		Vsuren	
Guy of Watwick	–	–	–	–
Confessio amantis	–	–	–	–
Merlin	–	–	–	–
Mandeville's Travels (S)	–	–	1	100%
Total	**–**	**–**	**1**	**100%**

Surprisingly, these two verbs appeared only in the two versions of *Mandeville's Travels*. Therefore, it is difficult to make any statements regarding their general use in the given regions.

5. Conclusion

The aim of the presented study was to first investigate whether the choice of the verbs in the two versions of one text – northern and southern – is text-dependent or region-dependent, which would then show to what extent the results of the comparison may also be noticed in other Middle English texts. In addition, the aim was to check if words' etymology may be influential as regards the occurrence of the verbs in texts from different regions. In order to conduct the analysis, two geographically distinct versions of *Mandeville's Travels* were examined. After this, four representative texts from the north and the south were searched through using electronic corpora. Apart from the abovementioned objectives, the goal of

this study in progress was to improve the methodology which will be used in the forthcoming MA thesis.

As this small study shows, quantitative analyses indicate that some words of Old Norse origin (such as *call*) were more popular in the south than vernacular forms (such as *clepen*). In addition, contrary to what was expected after qualitative analyses, some words of Old French origin were not prevalent in the southern texts. It is not surprising, as the words might have spread from dialect to dialect during those times. Undoubtedly, this problem requires more thorough diachronic analysis. On the other hand, it is worth mentioning that none of the verbs of Old Norse origin appeared exclusively in the south, and those of French origin – exclusively in the north. Moreover, according to expectations, most of the vernacular forms were distributed both in the north and the south. Interestingly, some words started to disappear from the English language altogether (e.g. *waten*). It may be important, then, to check when exactly this process took place. Some cases were problematic because of the limited data (e.g. *okeren-vsuren*). It is extremely difficult to make any reliable judgments based on the very small overall number of instances.

Interestingly enough, even the two regionally distinct versions of one text proved to be inconsistent as far as the choice of the verbs is concerned. The inconsistencies may be caused by several reasons. For example, the book might have been transcribed in fragments by different scribes. As a result, some words might have been changed because of individual, not regional preferences. A given word could have been either unfamiliar to the copyist or he might have avoided some equivalents deliberately. Hence, a vital question arises: to what extent are regional differences truly influential when idiosyncratic/stylistic aspect has to be taken into account? On the other hand, if the frequency of the words is relatively high, it is hard to exclude the possibility of some regional tendencies regarding word use.

To sum up, the study in the area of lexical preferences in the Middle English texts is certainly worth further exploration. In the future study, special attention will be paid to the variability resulting from such factors as the origin of the analysed manuscript. Moreover, as the presented study examines only the selected verbs, more examples of lexical differences will be investigated in the future, including different parts of speech. According to preliminary observations, nouns demonstrate more visible regional tendencies than verbs. In addition, greater emphasis will be put on semantic differences between words. Despite the fact that this might be extremely time-consuming, all the northern and East Midland texts from the *Corpus* will be searched through in the future in order to make the whole study more reliable.

References

Benskin, M., and M. Laing. 1981. Translations and Mischsprachen in Middle English Manuscripts. In *So meny people longages and tonges: philological essays in Scots and mediaeval English presented to Angus McIntosh*, eds. M. Benskin, and M.L. Samuels, 55–106. Edinburgh: EUP.

Black, M. 2000. Putting words in their place: an approach to Middle English word geography. In *Generative Theory and Corpus Studies: A Dialogue from 10 ICEHL*, eds. R. Bermúdez-Otero et al., 455–480. Berlin: De Gruyter.

Burrow, J.A., and T. Turville-Petre. 1996. *A Book of Middle English*. Oxford: Wiley-Blackwell.

Crystal, D. 2004. *The stories of English.* London: Penguin.

Dossena M., and R. Lass. 2004. Introduction. In *Methods and data in English Historical Dialectology*, eds. M. Dossena and R. Lass, 7–20. Bern: Peter Lang.

Dossena M., and R. Lass. 2009. Introduction. In *Studies in English and European Historical Dialectology*, eds. M. Dossena and R. Lass, 7–14. Bern: Peter Lang.

Fisiak, J. 1982. Isophones or Isographs? A Problem in Historical Dialectology. In *Language form and linguistic variation,* ed. J.A. Anderson, 117–128. Amsterdam: John Benjamins Publishing Company.

Fisiak, J. 2000. Middle English *beck* in the Midlands: the place-name evidence. In *Words: structure, meaning, function,* eds. Ch. Dalton-Puffer, N. Ritt, and D. Kastovsky, 87–94. Berlin: Mouton de Gruyter.

Hebda, A. 2010. Onde and envy: a diachronic cognitive approach. In *Studies in Old and Middle English*, ed. J. Fisiak, 107–126. Łódź–Warszawa: WSIELL.

Hoad, T. 1994. Word Geography: Previous Approaches and Achievements. In *Speaking in Our Tongues,* eds. M. Laing, and K. Williamson, 197–204. Cambridge: D.S. Brewer.

Hudson, A. 1983. Observations on a northerner's vocabulary. In *Five Hundred Years of Words and Sounds: A Festschrift for Eric Dobson,* eds. E.G. Stanley, and D. Gray, 74–83. Cambridge: D.S. Brewer.

Kaiser, R. 1937. *Zur Geographie des mittelenglischen Wortschatzes*. Leipzig: Palestra 205.

Laing, M., and R. Lass. 2006. Early Middle English Dialectology: Problems and Prospects. In *The handbook of the history of English*, eds. A. van Kemenade, and B. Los, 417–451. Oxford: Blackwell Publishing Ltd.

McIntosh, A. 1978. Middle English word-geography: its potential role in the study of the long-term impact of the Scandinavian settlements upon English. In *The Vikings: Proceedings of the Symposium of the Faculty of Arts of Uppsala University, June 6–9, 1997*, eds. T. Anderson and K.I. Sandred, 124–30. Uppsala.

McIntyre, D. 2008. *History of English: A Resource Book for Students*. London: Routledge.

Meurman-Solin, A. 2000a. Change from above or from below? Mapping the loci of linguistic change in the history of Scottish English. In *The Development of Standard English, 1300–1800: theories, descriptions, conflicts,* ed. L. Wright, 155–70. Cambridge: Cambridge University Press.

Meurman-Solin, A. 2000b. On the conditioning of geographical and social distance in language variation and change in Renaissance Scots. In *The History of English in a Social Context. A Contribution to Historical Sociolinguistics*, eds. D. Kastovsky, and A. Mettinger, 227–55. Berlin: Mouton de Gruyter.

Seymour, M.C. 2002. *The Defective Version of Mandeville's Travels*. Oxford: Oxford.

Wardale, E.E. 1956. *An introduction to Middle English*. Oxford: Taylor & Francis.

Internet sources

CME – Corpus of Middle English Prose and Verse. http://quod.lib.umich.edu/c/cme/. Accessed June 2014.

eLALME – Linguistic Atlas of Late Mediaeval English online. http://www.lel.ed.ac.uk/ihd/elalme/elalme.html. Accessed May 2014.

Middle English Dictionary online. http://quod.lib.umich.edu/m/med/. Accessed June 2014.

Online Etymology Dictionary. http://www.etymonline.com/. Accessed May 2014.

Map 1 – http://kids.britannica.com/comptons/art-143574/Middle-English-dialects-of-England. Accessed June 2014.

Maps 2 and 3 – http://en.wikipedia.org. Accessed June 2014.

Second-person pronouns and their relation with nominal forms of address in Late Middle English and Early Modern English personal letters

http://dx.doi.org/10.18778/8088-065-8.07

Emilia Szczytko

University of Lodz

Abstract

Generally, little attention has been given to the role of selected linguistic and extralinguistic factors in the use of forms of address (Walker 2007). Therefore, the major theoretical concern behind this research is to examine quantitatively and qualitatively, based on selected letters from the CEECS corpus (1998), the influence of social stratification and family relations on the usage of pronominal forms of address. Apart from that, it also analyses the interrelation between second-person pronouns and nominal forms of address in Late Middle English and Early Modern English.

1. Introduction

Terms of address can be divided into two categories: pronominal and nominal. There have been a substantial number of scholars who devoted their studies to this phenomenon (Mulholland 1967; Barber 1981; Brown and Gilman 1960; Brown and Levinson 1987; Mazzon 2009; Kopytko 1993; U. Busse 2002, B; Busse 2006). However, most of the existent studies were performed from a pragmatic perspective using various modified versions of Brown and Levinson's politeness theory (Busse 2006). Hence, the present study seeks to explore the connection between the usage of second-person pronouns, social stratification and family relations in Late Middle English and Early Modern English. Apart from checking the effect

of social rank and type of relations, it aims at analysing the type of correlation between pronominal and nominal terms of address. Since many studies have been based on Shakespeare's dramatic works (Mulholland 1967; Barber 1981; Brown and Gilman 1960; Brown and Levinson 1987; Kopytko 1993; U. Busse 2002; B. Busse 2006), this study uses the collection of letters from the CEECS (1998) corpus as a material subjected to analysis.

This article is divided into five sections. The sections devoted to the description of theoretical background, results and analysis are numbered from two to five. The second section is a description of rank classification in Late Middle English and Early Modern English. Then the article progresses to a section devoted to epistolary conventions in the above-mentioned period. Another section describes research methodology and research questions. Finally, the last section serves to provide answers to the research questions and reevaluate the relevance of social stratification and family relations in pronoun selection. Apart from the discussion of results, the section reveals weaknesses of the study, and suggests some ideas for further research.

2. Rank classification in Late Middle English and Early Modern English

Social class is one of the key notions in sociolinguistics, since it has its roots in functionalist sociology (Saville-Troike 2003). The term may be approached from various perspectives. When describing the concept, Spolsky (1998) concentrates on economic aspects and notes that it is a set of divisions, which is determined based on such factors as income, occupation and education. Singh (2009) states that what is important in specifying the nature of one's social status is not only the economic situation of an individual but also the prestige of birth and the mode of living. Kerswill (2010) also emphasizes the fact that traditionally, the notion of social class is presented as a set of divisions in socioeconomic hierarchy. However, he provides his own definition as well, and describes it as one of the internal differentiations and constraints on one's usage of language, which enable categorisation of people into broad groups in a society. As far as the present study is concerned, the definition provided by Kerswill (2010) is more applicable, since this study is sociolinguistic in nature and its primary aim is to check the influence of social rank on language use with respect to terms of address.

In Late Middle English and Early Modern English, the society was highly stratified. Social status depended mostly on one's position on the social ladder (Laslett 1983). Furthermore, the sources from the sixteenth century reveal that the society was divided into four layers. The structure of society in the above-mentioned period is presented in the table on the following page:

Table 1. Detailed rank classification (Walker 2007:25)

	Code	Description	Official title	Occupation
Non-commoners	**A**	royalty, nobility and the high clergy	Queen, Duke, Archbishop, Baron, Bishop	
	B	knights and baronets	Sir	
	C C1	gentry	Esquire	lawyer, doctor, army officer, clergyman,
	C2	those in the professions, wealthy traders, wholesale merchants	Doctor, Colonel	teacher, financier,
Commoners	**D**	well-to-do farmers, and retailers, urban masters, and certain urban craftsmen		yeoman, shopkeeper, innkeeper, cutler
	E	poorer farmers and (especially) rural craftsmen		husbandman, weaver, blacksmith, shoemaker, alehouse keeper
	F	poor wage-earners, or those bound to a master		labourer, servant, apprentice
	G	unemployed, criminals		pauper, vagrant, whore, thief

The table above contains a detailed description of all the levels, the division into gentry and non-gentry, and official and occupational titles (Walker 2007). In the coding system of divisions, each capital letter represents a different layer of the

society. The top ranks of the society, namely people from groups A, B and C1 did not do any kind of manual work, and their income came from land ownership (Laslett 1983). Group A stands for royalty and high clergy, group B for knights and baronets, while C1 for the gentry. As far as group C2 is concerned, Walker (2007) states that it is difficult to place a group of professions in the social hierarchy, since the group does not fit into the division based on the ownership of land. He notes that in this group, various kinds of service or commerce are sources of wealth; therefore, he describes it as *pretended gentry*. When considering the differences between social groups, the greatest divide lines can be observed between non-commoners and commoners represented in groups from D to G. Laslett (1983) notes that downwards the social ladder one's status was defined only based on occupation and its position in the hierarchy. The groups of non-gentry illustrate a classification of the lower echelons of society who relied on manual labour solely. The system will be used in the study to classify the authors and addressees of the chosen letters form the corpus.

3. Epistolary conventions in Late Middle English and Early Modern English

In Late Medieval and Early Modern England, letter writing was considered one of the methods used to teach people classical rhetoric. It was claimed that in personal correspondence there were some traces of Renaissance humanism, which had influence on the epistolary conventions (Nevalainen and Raumolin-Brunberg 1995). Therefore, when writing letters one had to follow a set of rules related to form and content.

There were many manuals containing guidelines for writing personal letters. One of the most popular letter-writing manuals written by Fullwood (1558) is a detailed description of all the rules with regard to technical requirements and ways of addressing individuals. Since the focus of the present study is on terms of address, technical aspects of private correspondence, such as visual representation of one's social status by means of layout, will not be discussed. Fullwood (1568) suggests that when addressing members of higher or lower class one has to remember to emphasise social status of the addressee. He notes that when writing to social superiors one has to do it with honour, humility and reverence, and he or she should not address them with their first name. In addition, he points out that using first name instead of names denoting social rank accompanied by ap-

propriate modifiers would be disrespectful. When considering letters directed at social equals, he stresses the fact that one should express familiar reverence and politeness, and use one's name of rank and such words as *worshipful* or *honourable*. As far as pronominal terms of address are concerned, Fullwood (1568) adds that non-commoners, in other words, members of gentry and nobility, should always employ *you*. In contrast, he points out that in address to social inferiors one should show his or her authority and use *thou*.

Nevalainen and Raumolin-Brunberg (1995) agree with Fullwood (1568) and state that in the fifteenth and sixteenth century, the society favoured very complex forms of address. They enumerate various modifiers of nominal terms of address, which were considered clear indicators of addressee's position within the social hierarchy. The table on the following page contains a set of the most frequently encountered modifiers denoting social class together with their explanation:

Table 2. Typical modifiers on nominal forms of address in LME and EME (on the basis of Nevalainen and Raumloin-Brunberg 1995: 550)

Modifier	Meaning
generous	high-born
gentle, kind	well-born
honest	holding a honourable position
honourable	of distinguished social rank
noble	illustrious by rank, title, or birth
reverent	worthy of deep respect on account of rank, age or character
worshipful	distinguished in respect of character or rank
worthy	holding a prominent place in the community

The table above reveals that in early correspondence, there could be a tendency to emphasise addressee's social rank and to follow epistolary conventions (Hall 1908). In the present study, it will be checked if the chosen individuals followed all the rules and used terms of address in order to indicate social class differences.

Apart from non-kinship terms denoting social class membership, Braun (1988) and Nevalainen and Raumolin-Brunberg (1995) discuss the typical model of household in England. Nevalainen and Raumolin-Brunberg (1995) state that nuclear family consisting of two generations was the prevailing type. Moreover, they add that people tended to indicate a type of kinship in address terms, even if it was a very distant relation and even in addressing members of non-nuclear family. They argue that when the speaker and the addressee were connected by kinship ties, no-naming was a common phenomenon, and people usually addressed each other with kinship terms accompanied by modifiers and intensifiers such as *right* or *most*. Braun (1988) also comments on no-naming and she points out that using first name was a common practice only among the ranks below nobility. In contrast to the claims made by Nevalainen and Raumolin-Brunberg (1995), she states that in social relations, social rank always overrode kinship and in the case of status differentials, there should be no indication of family relations between the speakers. Walker (2007) agrees with her point and notes that when there is any status differential, one should always mark it in forms of address due to the importance of social stratification and strong tendency to signalise differences by means of language in the fifteenth and sixteenth century.

4. Methodology – The influence of social rank and family relations on pronoun selection

The primary aim of the present study is to investigate qualitatively and quantitatively the influence of social rank and family relations on the usage of pronominal forms of address in Late Middle English and Early Modern English. Apart from the analysis of pronoun selection, it additionally checks the type of correlation between pronominal and nominal forms of address. In order to check the impact of chosen non-linguistic factors on the pronoun usage, the following research questions have been constructed:

— Is addressee's social rank reflected by the usage of pronominal forms of address?
— Are family relations reflected by the usage of pronouns?
— What is the type of correlation between pronominal and nominal forms of address in the personal letters chosen from CEECS?

The study is based on the collection of forty letters from the fifteenth and sixteenth centuries, retrieved from the first part of *the Corpus of Early English Correspondence Sampler* (CEECS). The CEECS consists of two parts and the total number of tokens is 450,000. It is one of the elements of the *Corpus of Early English Correspondence,* which was compiled by Sociolinguistics and Language History Project Team at the Department of English at the University of Helsinki. The team consisted of such scholars as Helena Raumolin-Brunberg, Terttu Nevalainen, Minna Nevala, Arja Nurmi, Jukka Keranen or Minna Palander-Colin (Nurmi 1998).

As far as reliability of the corpus is concerned, Nurmi (1998) states that the CEECS proves to be a useful tool in all types of linguistic research apart from the studies of orthography, since spelling was not standardised then. In addition, she notes that despite the size, the social representativeness of the corpus is as wide as possible. Nevalainen (1996) and Raumolin-Brunberg (1995) also comment on the representativeness of the corpus. In contrast to Nurmi (1998), they do not consider the exact word count. They focus on the low level of literacy in Late Middle and Early Modern English. Nevalainen (1996) points out that due to the abovementioned problem, it was not possible to cover entire social hierarchy in the corpus, because most of the letters were written by members of the higher levels of society, who according to the figures presented by Laslett (1983) represented only around 5% of the whole society. However, she further notes that contrary to the problem of illiteracy and limitations set by it, the corpus contains appropriate kind of data for sociolinguistic investigation. Apart from the issue of reliability of the data from the corpus, Palander-Colin et al. (2009) also add that letters as a text type bear close resemblance to speech, since they are a kind of communication between identified individuals. Therefore, the collection of letters chosen from the CEECS seems to be a good choice when assessing the reliability of the materials subjected to analysis.

In the present study, twenty-eight identified individuals, who are the authors of the chosen letters from the CEECS corpus, are basic units of analysis. In order to obtain relatively high representativeness, the individuals had to represent higher and lower layers of the society. Apart from aiming at a relatively high representativeness, another factor was also taken into consideration, namely gender of the addressers and addressees. Since the influence of gender is not of particular interest in the context of the study, only letters written by men and addressed to men were chosen. The choice was also motivated by the fact that most of the letters

from the corpus were written by men and addressed to men, and women, who were mainly members of royalty and nobility, wrote only one-fifth of the letters. The vast majority of chosen letters subjected to analysis comes from the collection written by members of Stonor family and its servants. Apart from that, the data also contains the correspondence pertaining to the highest echelons of society, namely letters written by clergy and royalty. Due to relatively high illiteracy in the fifteenth and sixteenth century, the correspondence between members of the lowest levels of society is not available (Nevalainen and Raumolin-Brunberg 1995).

The letters were grouped according to the description of rank classification based on the studies done by Walker (2007) and Nevalainen and Raumolin-Brunberg (1995), which was presented in the first section of the present study. The table enclosed in the appendix shows the list of chosen letters from the corpus together with the date of composition, short descriptions of the authors and addressees and their social background, the number of letters written by each author, relations between the author of the given letter and the addressee, and expected results.

The study consisted of several stages. The first step was the choice of letters. Another one concerned the analysis of letters with the help of Wordsmith tools 5.0. The first element of the analysis performed with the help of the program involved the use of *concordance search*. This type of search was used in order to find all the possible contexts of usage of pronominal forms of address to check qualitatively the type of correlation between pronominal and nominal terms of address, and to investigate the influence of social stratification on the usage of second-person pronouns by applying the framework presented in the second section. Second element of the study was related to the quantitative part. It relied on generating the *frequency lists* by the program in order to obtain the exact number of second-person pronouns in all the case forms and to find the most frequently occurring nominal forms of address and the accompanying modifiers and intensifiers.

As far as the expected results are concerned, the usage of second-person pronouns by the chosen individuals is supposed to reflect the influence of social stratification and family relations. Furthermore, there should be a strong correlation between the employed pronominal and nominal forms of address. The letters should also fulfil all the requirements related to epistolary conventions in the fifteenth and sixteenth century presented by Fullwood (1568).

5. Results of the study

The present section is concentrated on the results obtained from the quantitative and qualitative study. It contains interpretations of numerical and qualitative data in the theoretical framework presented in the previous sections. Apart from findings and their interpretation, the section provides an overall summary of the obtained results and describes weaknesses of the study, and some suggestions for further research.

Before presenting a detailed analysis of pronominal forms of address with respect to social and linguistic factors, the overall distribution of second-person pronouns in the selected letters from the corpus will be presented in the table below:

Table 4. The number of occurrences of *you* in all the spelling variants

Form	Number of occurrences
YOU	86
YOUE	8
YOUR	213
YOURE	16
YOURRE	1
YOURS	1
YOV	4
YOVEN	1
YOW	83
YOWE	19
YOWER	6
YOWR	6
YOWRE	9
YOWRS	2
YOWUR	6

The table above shows that the total number of occurrences of *you* in all case forms and spelling variants was 464. As regards the investigated pronouns, quantitative analysis of the results performed with the help of Wordsmith tools 5.0 revealed, as it was expected, that there were no instances of the usage of *thou*. When considering letter-writing manuals, the usage of pronominal forms of address might have been determined by the epistolary conventions, which suggested that social rank should be presented and emphasised in all types of correspondence, even in family letters. Apart from that it might be also connected with the content of the letters, which was very formal, since in the majority of cases, business matters were the main issue discussed by the authors of letters. Another table contains a list of nominal terms of address and their modifiers together with the number of occurrences in letters written by each author:

Table 5. List of nominal forms of address and their modifiers

Author	Addressee	Results
1	**2**	**3**
GROUP A		
King Henry VII (royalty)	**a)** Sir Gilbert Talbot (group B-knight, Earl of Shrewsbury) **b**) Sir William Say(group C2-below nobility-Sir/Sheriff of Hertfordshire) **c)** Cardinal Wolsey (group A-high clergy, royal minister and Archbishop of York)	**a)** ***Trusty and well-beloved; well-beloved Knight and Sir*** **b)** ***Trusty and well-beloved; well-beloved knight*** **c)** ***Lord Cardinal; Lord; Good Cardinal***
Richard Duke of York (nobility)	the Citizens of Shrewsbury (lower class-below nobility and below C2-G)	***Right worshipful friends; worshipful friends***
Dr Cuthbert Tunstall (Bishop of Durham)	King Henry VIII (royalty)	your Grace (30)
John Abbot of Norton (Abbot of Norton)	William Stonor (group B-Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right worshipful and fullgood master (2)***

1	2	3
GROUP B		
Sir Thomas Boleyn (Sir/Earl of Wiltshire)	King Henry VIII (royalty)	***your Grace (15); your Highness (9)***
Lord Dacre (baron)	Cardinal Wolsey (group A-high clergy; royal minister and Archbishop of York)	***your Grace (5)***
Humphrey Forster (Sir, Sheriff of Gloucestershire)	Thomas Stonor (group B- Sir; knight)	***Right worshipful and good, kind brother (2); good Brother (4)***
Thomas Stonor (Sir; knight)	William Stonor(group B-Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***William Stonor; you+no naming***
Thomas Hampden (knight)	William Stonor (group B-Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right worshipful cousin (2)***
GROUP C2		
Hugh Unton (lawyer)	Thomas Stonor (group B- Sir; knight)	***Right worshipful master (2); sir***
William Goldwyn(physician)	John Byrell (C2-apothecary)	***Sir (3); master (2)***
Richard Page (lawyer)	William Stonor (group B-Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Good mastership; master Sir William; sir (3); right singular good master (2)***
Edmund Stonor (merchant)	William Stonor (group B-Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right reverent and worshipful brother (3); good brother (4)***
Thomas Mathew (bailiff)	Thomas Stonor (group B- Sir; knight)	***Right worshipful master***
Richard Pace (diplomat/ administrator; the Cardinal's secretary)	Cardinal Wolsey (code A- high clergy; royal minister and Archbishop of York)	***your Grace (4)***

Table 5. cont.

1	2	3
Thomas Betson (merchant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right worshipful sir; right worshipful and singular good master; your good mastership (3); sir (4)***
William Burbank (the Cardinal's secretary)	King Henry VIII (royalty)	***Your most noble Grace; your Grace***
GROUP F		
John Frende (family servant)	Thomas Stonor (group B- Sir; knight)	***Right worshipful master (2)***
John Yeme (family servant)	Thomas Stonor (group B- Sir; knight)	***Right reverent master (2)***
Thomas Mull (family servant)	Thomas Stonor (group B- Sir; knight)	***Master Stonor; sir; right worshipful master***
Walter Elmes (family servant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Master (2)***
Goddard Oxbryge (family servant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right worshipful and reverent sir (3); good master; sir(3); right worshipful and reverent master***
Henry Makney (family servant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Good master (4); sir (2)***
Thomas Henham (family servant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right reverent and worshipful master (2); your mastership (2); right honourable (1); sir (5)***
Henry Dogett (family servant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right reverent worshipful master (2)***

1	2	3
Richard Germyn (family servant)	William Stonor (group B- Sheriff of Oxfordshire, Berkshire &Devonshire, High Steward of Oxford University)	***Right reverent master (2); sir (2)***
NO INFORMATION ABOUT SOCIAL CLASS MEMBERSHIP		
Thomas Hampton	Thomas Stonor (group B- Sir; knight)	***Right worshipful cousin (2); sir (5)***

In the above table, the exact number of usages by the authors is given in cases when there was more than one example noted. The letters written by one of the members of the top level of society, namely by King Henry VII, contained the forms denoting social rank of the addressees. Other representatives of group A, namely Dr Cuthbert Tunstall and John Abbot of Norton, also employed terms of address indicating social class membership. However, one of the authors from group A, Richard Duke of York, did not emphasise social rank when addressing citizens of Shrewsbury. The reason why he used the form *friends* might have been related to the fact that he wanted to politely encourage people to enjoy the election of the new king and to fulfil some orders.

As far as the representatives of group B are concerned, not all of them used terms denoting social rank of the addressees. Two of them, namely Lord Dacre and Sir Thomas Boleyn emphasized social class membership. However, the rest, apart from William Stonor who addressed his son with first name and family name, tended to indicate the type of family relations. One of the representatives of group C, namely Edmund Stonor, also used kinship terms when addressing his brother. The authors of letters from group C who were not related by any family bonds always employed forms denoting social class membership. As far as members of the lower echelons of the society are concerned, that is family servants, they always expressed social rank of their masters by means of terms of address. All the representatives of group F usually used such forms as *master* and *sir*. In addition, they used such modifiers as *right, worshipful* or *reverent*, which were clear indicators of social position of the addressee.

When analysing the correlation between pronominal and nominal forms of address, the results prove that it is not possible to determine the exact strength of relation between the two, since there were no occurrences of *thou* noted. In the vast majority of cases, namely in the case of letters written by the representa-

tives of the top levels of society, *pretended gentry*, and servants, who were not related by any kind of family bonds, the usage of pronominal and nominal terms of address reflected social rank of the addresser and addressee. Surprisingly, the qualitative part of the study also shows that members of the Stonor family, who were related by different types of kinship, contrary to what was claimed by Brown (1988), always emphasised the type of family relation, not one's social position in the hierarchy.

The results obtained from quantitative and qualitative parts of the present study prove that social stratification seemed to have influence on the usage of pronominal forms of address in the selected letters. The data also confirm the claims made by Nevalainen and Raumolin-Brunberg (1995), since the authors of letters who were related by the ties of kinship tended to indicate type of family relations rather than social rank. The lack of *thou* in the selected letters could be the effect of epistolary conventions on the choice of the right form.

As far as the weaknesses of the study are concerned, the amount of the data analysed for the purpose of pilot study is too small to generalise about the whole society, therefore further studies are needed. Furthermore, in order to compare the instances of the usage of *you* and *thou,* bottom-to-top approach is required, and firstly the CEECS corpus should be checked for the exact number of occurrences of both pronominal forms of address. Apart from that, in order to draw some more general conclusions about the society in Late Medieval and Early Modern England, letters written by social equals from the lower echelons of the society should also be subjected to the analysis. In addition, letters written by women should also be investigated.

Conclusions

To sum up, the corpus-based investigation described above examined the use of terms of address in the selected letters from Late Middle English and the beginning of Early Modern English. The aim of the present pilot study was fulfilled. The results obtained from the quantitative and qualitative parts of the study corroborate the claim that the extralinguistic factor under scrutiny, namely social stratification had influence on the usage of terms of address in the vast majority of cases. When considering further investigation, the influence of a greater number of factors has to be checked to explain fully the mechanisms governing the use of terms of address in private letters in the above-mentioned period.

References

Primary sources

Chosen letters from the *Corpus of Early English Correspondence Sampler*. text version. 1998. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, Jukka Keränen, Minna Nevala, Arja Nurmi and Minna Palander-Collin. Helsinki: University of Helsinki.

Secondary sources

Andersen, G. 2010. How to use corpus linguistics in sociolinguistics. In *The Routledge Handbook of Corpus Linguistics*, eds. A. O'Keeffe, and M. McCarthy. 547–562. New York: Routledge.

Barber, Ch. 1981. You and Thou in Shakespeare's Richard III. In *Leeds Studies in English*, New Series XII, 273–289. Reprinted in *A Reader in the Language of Shakespearean Drama* comp. by Vivian Salmon & Edwina Burness 1987. 163–177. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Braun, F. 1988. *Terms of address: problems of patterns and usage in various languages and cultures*. Berlin: Mouton de Gruyter.

Brown, P., and S. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Brown, R., and A. Gilman. 1960. The Pronouns of Power and Solidarity. In *Style in Language,* ed. T.A. Sebeok, 253–276. Cambridge: MIT Press.

Busse, B. 2006. *Vocative Constructions in the Language of Shakespeare*. Amsterdam: John Benjamins Publishing Company.

Busse, U. 2002. *Linguistic Variation in the Shakespeare Corpus: Morpho-syntactic Variability of Second Person Pronouns*. Amsterdam: John Benjamins Publishing Company.

Coulmas, F. 1997. Introduction. In *The handbook of sociolinguistics,* ed. F. Coulmas, 5–11. Cornwall: Blackwell Publishing.

Fullwood, W. 1568. *The Enimie of Idlenesse: Teaching a Perfect Platforme how to Indite Epistles and Letters of All Sortes: As Well by Answere As Otherwise: No Lesse Profitable then Pleasaunt*. London: Henry Midleton.

Kerswill, P. 2010. Introduction. In *The SAGE Handbook of Sociolinguistics*, eds. B. Wodak, and P. Kerswill, 1–9. London: Sage publications Ltd.

Laslett, P. 1983. *The World We Have Lost: Further Explored (3rd ed.)*. London: Routledge.

Mazzon, G. 2009. *Interactive Dialogue Sequences in Middle English Drama.* Amsterdam: John Benjamins Publishing Company.

Mulholland, J. 1967. Thou and You in Shakespeare. A Study in Second Person Pronoun. *English Studies* 48: 34–43.

Nevalainen, T., and H. Raumolin-Brunberg. 1995. Constraints on politeness: the pragmatics of address formulae in Early English correspondence. In *Historical Pragmatics. Pragmatic Developments in the History of English*, ed. A. Jucker, 541–601. Amsterdam: John Benjamins Publishing Company.

Nevalainen, T., and H. Raumolin-Brunberg. 1996. *Sociolinguistics and Language History: Studies Based on the Corpus of Early English Correspondence*. Amsterdam: Rodopi .

Nevalainen, T., and S.K. Tanskanen, 2007. *Letter Writing.* Amsterdam: John Benjamins Publishing Company.

Nurmi, A. 1998. Manual for the Corpus of Early English Correspondence Sampler (CEECS). http://www.hit.uib.no/icame/ceecs/index.htm. Last accessed: May 10th,2014.

Saville-Troike, M. 2003. The *Ethnography of Communication: An Introduction (3rd ed.)*. Cornwall: Blackwell Publishing.

Singh, V.P. 2009. *Caste, class and democracy: changes in a stratification system*. New Jersey: Transaction Publishers.

Spolsky, B. 1998. *Sociolinguistics (1st ed.*). Oxford: Oxford University Press.

Palander-Collin, M., M. Nevala, and A. Nurmi. 2009. *The Language of Daily Life in England (1400–1800). Amsterdam: John Benjamins Publishing Company.*

Walker, T. 2007. *Thou and You in Early Modern English Dialogues.* Amsterdam: John Benjamins Publishing Company.

Wardhaugh, R. 2006. *An Introduction to Sociolinguistics (5th ed.).* Cornwall**:** Blackwell Publishing.

Appendix

Table 3. Units of analysis

Date	Author and his social background	Addressee and his social background	Number of letters	Type of relations	Expected results
1	2	3	4	5	6
1452	Richard Duke of York(code A-nobility- the higher class)	the Citizens of Shrewsbury (lower class-below nobility and below C2-G)	1	Non-kinship relation	The use of *you* + nominal forms denoting social group membership
1495	King Henry VII (code A-royalty)	Sir Gilbert Talbot (code B-knight, Earl of Shrewsbury)	1	Non-kinship relation	The use of *you* + *sir*
1490	King Henry VII (code A-royalty)	Sir William Say (code C2-below nobility-Sir/Sheriff of Hertfordshire)	1	Non-kinship relation	The use of *you* + *sir*
1514	William Burbank (Code C2- the Cardinal's secretary)	King Henry VIII (code A-royalty)	1	Non-kinship relation	The use of *your+Grace*
1516	King Henry VII (code A-royalty)	Cardinal Wolsey (code A- high clergy; royal minister and Archbishop of York)	2	Non-kinship relation	The use of *you+Cardinal*
1517	Dr Cuthbert Tunstall (code A- Bishop of Durham)	King Henry VIII (code A-royalty)	1	Non-kinship relation	The use of *your+Grace/ Highness*
1519	Sir Thomas Boleyn (code B-Sir/Earl of Wiltshire)	King Henry VIII (code A-royalty)	1	Non-kinship relation	The use of *your+Grace/ Highness*
1519	Richard Pace (code C2-diplomat/ administrator; the Cardinal's secretary)	Cardinal Wolsey (code A- high clergy; royal minister and Archbishop of York)	1	Non-kinship relation	The use of *you+Cardinal*

Table 3. cont.

1	2	3	4	5	6
1517	Lord Dacre (code B-baron)	Cardinal Wolsey (code A- high clergy; royal minister and Archbishop of York)	1	Non-kinship relation	The use of *you*+ *Cardinal*
1462	John Frende (code F- family servant)	Thomas Stonor (code B- Sir; knight)	3	Non-kinship relation	The use of *you+master*
1462	Thomas Hampton (no information given)	Thomas Stonor (code B- Sir; knight)	1	Kin-cousins	The use of *you+cousin*
1462	Hugh Unton (code C1-lawyer)	Thomas Stonor (code B- Sir; knight)	1	Non-kinship relation	The use of *you+master*
1466	John Yeme(code F- family servant)	Thomas Stonor (code B- Sir; knight)	1	Non-kinship relation	The use of *you+master*
1466	Humphrey Forster (code B-Sir, Sheriff of Gloucestershire)	Thomas Stonor (code B- Sir; knight)	2	Brothers-in-law	The use of *you+Sir*
1469	Thomas Stonor (code B-Sir; knight)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	2	Father-son	The use of *you*+ kinship terms, terms of endearment
1472	Thomas Mull (code F-family servant)	Thomas Stonor (code B-Sir; knight)	1	Non-kinship relation	The use of *you+master*
1469	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	Thomas Stonor (code B-Sir; knight)	1	Son-father	The use of *you*+ father, family name, terms of endearment

1473	Thomas Mathew (code C2-bailiff)	Thomas Stonor (code B-Sir; knight)	1	Non-kinship relation	The use of *you+master*
1475	Edmund Stonor(code C2-merchant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	2	brothers	The use of *you*+ brother, family name, terms of endearment
1481	Walter Elmes(code F-family servant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	1	Non-kinship relation	The use of *you*+master
1476	Thomas Betson (code C2-merchant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	2	Son-in-law to Father-in-law	The use of *you*+master
1476	Goddard Oxbryge (code F-family servant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	2	Non-kinship relation	The use of *you*+master
1477	Thomas Hampden (code B-knight)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	1	Kin- cousins	The use of *you*+ cousin
1478	Henry Makney(code F-family servant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	2	Non-kinship relation	The use of *you*+master
1477	John Abbot of Norton (Code A- Abbot of Norton)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	1	Non-kinship relation	The use of *you+sir*
1480	Richard Page (code C2-lawyer)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	2	Non-kinship relation	The use of *you+master*

Table 3. cont.

1	2	3	4	5	6
1479	Thomas Henham (code F-family servant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	1	Non-kinship relation	The use of *you+master*
1479	Henry Dogett (code F-family servant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	1	Non-kinship relation	The use of *you+master*
1480	William Goldwyn (code C2-physician)	John Byrell (Code C2- aphotecary)	1	Non-kinship relation	The use of *you+sir*
1480	Richard Germyn (code F-family servant)	William Stonor (code B- Sheriff of Oxfordshire, Berkshire, & Devonshire, High Steward of Oxford University)	1	Non-kinship relation	The use of *you+master*

Variability in L2 English pronunciation examined through the prism of phonetic imitation

http://dx.doi.org/10.18778/8088-065-8.08

Magdalena Zając

University of Lodz

Abstract

The paper is concerned with the imitation of vowel duration and quality upon exposure to native and non-native English speech. The participants were Polish learners of English recruited at the University of Lodz. The study aimed to determine whether the extent of phonetic imitation may be influenced by the model talker being a native or a non-native speaker of English and whether different imitation strategies may explain some of the variability in L2 speech. The results of the study suggest that phonetic imitation may account for some of the variability in L2 pronunciation and that the native/non-native status of the model talker may have a bearing on the direction of convergence. It was also found that the magnitude of imitation may depend on the degree to which a given L2 feature functions in the learners' interlanguage.

1. Variability in L2 pronunciation

The fact that there exist considerable distinctions between the L2 pronunciation of speakers with different native languages is well documented in SLA literature and appears to be closely linked to differences in perception. According to Native Language Magnet Model (Kuhl 2000), the representations of native sounds in our brains act like 'perceptual magnets' for L2 phones that bear resemblance to the L1 sounds. Perceptual Assimilation Model (Best 1995) and Speech Learning Model (Flege 1993) propose that the perception of an L2 phone involves comparing the sound with all sounds in the L1 system of the speaker. Indeed, the fact that the perception of a given sound depends on the listener's native language was confirmed in a number of studies, e.g. Brown (2000), Fox et al. (1995), Iverson et

al. (2001), and Rochet (1995). The differences in perception lead to a wide range of variability in the production of a given sound by speakers with different mother tongues. For instance, Davidian and Flege (1984) discovered that the native Polish subjects in their study devoiced word-final plosives in their English productions, whereas native Spanish and native Chinese participants deleted word-final stops. Livbjerg and Mees (1988) state that Danish learners may be disposed to replace English /ʌ/ and /ɒ/ with their native vowel /ɔ/, while Gonet, Szpyra-Kozłowska and Święciński (2010) show that English /æ/ is often substituted by /a/ or /e/ by native speakers of Polish.

It has also been argued that various cognitive and affective factors contribute to increased accent variability among different L2 speakers. Piske et al. (2001) argue that the degree of foreign accent in one's speech is largely determined by one's ability to mimic unfamiliar speech sounds. Suter (1976), Purcell and Suter (1980) and Elliot (1995) found that the amount of concern for L2 pronunciation accuracy had a considerable bearing on learners' L2 productions, indicating that motivation may play a major role in successful acquisition of L2 phonology. Numerous studies show that the age at which one starts learning a second language has a significant impact on the degree of foreign accent in one's speech and suggest that attaining native-like pronunciation is considerably more difficult for adult learners than for children (e.g. Flege 1988; Moyer 1999; Oyama 1976; Suter 1976; Tahta et al. 1981; Thompson 1991). Another factor which was found to affect non-native pronunciation is L2 input. For instance, Purcell and Suter (1980) provide some evidence that increased contact with native speakers may reduce the degree of foreign accent in learners' speech.

Other findings indicate that variability in L2 pronunciation is also strongly related to a number of social factors. For instance, non-native pronunciation appears to depend on the speaker's gender to some extent. Tahta et al. (1981) and Thompson (1991) discovered that women's L2 pronunciation was rated higher than men's, while Hartford (1978) found that female Mexican-American adolescents used more prestige forms in their English pronunciation than did male Mexican-American adolescents. Thompson (1976) concentrated on Chicano English and found that participants with lower socio-economic status used Spanish-influenced pronunciation features to a greater extent than subjects with higher socio-economic status. Gatbonton (1975) investigated the pronunciation of French-Canadian learners of English and reported that successful acquisition of English dental fricatives was conditioned by the strength

of affiliation with the English community. Similar observations were also made by Zuengler (1982).

An interesting aspect of L2 pronunciation is that it varies not only between individual speakers but also within one speaker. For instance, a given learner's production of a particular sound may differ depending on the phonetic environment of the sound in question. Anderson (1987) observed that native speakers of Mandarin Chinese omitted /r/ more frequently in post-vocalic positions and deleted word-final /t d/ more frequently in consonant clusters. The results of a study by Bayley (1996) showed that Chinese learners tended to omit English /t/ and /d/ more after liquids than nasals or obstruents. Benson (1988) found that consonant deletion in the productions of Vietnamese learners of English was connected with the preceding vocalic context. Tarone (1982) hypothesised that the production of a given second-language pronunciation feature is also affected by the amount of attention that a learner pays to speech form. She argued that attention to speech form increases when learners are asked to perform elicitation tasks such as reading of word lists and decreases in less formal tasks such as free speech. Tarone's claims were corroborated in a study by Dickerson and Dickerson (1977), who examined Japanese learners' realisations of /r/ in free speech, dialogue reading and word-list reading and found that /r/ was supplied only 50% of the time in the first task and almost 100% correctly in the last task. Similar results were obtained in a study on Thai learners' production of English /r/ (Beebe 1980). Interestingly, it was found that the number of target-like realisations of the investigated sounds depended not only on the amount of attention paid to speech but also on phonetic environment. Dowd (1984) examined L2 pronunciation of Mexican women and detected that the informants' production of certain features was affected by the type of question they were asked and that some of the investigated features shifted in opposite directions. When asked an emotional question, the participants tended to produce final consonant clusters less accurately but, at the same time, increased correct realisations of /r/. The findings of Gonet et al. (2010) suggest that some within-speaker variability may also be brought about by the existence of phonetic false friends in the learners' L1 and L2. It was found that Polish learners of English substituted /æ/ with either /e/ or /a/ and that in the majority of cases, the substitution pattern accorded with the vowels present in the corresponding Polish loanwords from English.

Overall, it appears that variability in L2 pronunciation occurs both between different and within individual speakers. A speaker's L2 phonetic performance

may be affected by certain social factors such as age, gender, personality traits or attitudes, language-related features such as the structure of the speaker's L1 sound system and language universals and, finally, cognitive factors such as the amount of attention paid to speech or language aptitude.

2. Phonetic imitation

The process of changing or adjusting one's speech upon exposure to the speech of others first attracted researchers' attention in the 1970s. Howard Giles and colleagues referred to the phenomenon as convergence or accommodation and developed a framework called Communication Accommodation Theory (CAT) to account for the accent and language shifts that individuals make when interacting with other people. The advocates of CAT were primarily concerned with speech behaviour in conversational interactions and the social-psychological factors that may affect language and accent shifts in socially rich settings (Bourhis and Giles 1977; Coupland 1984; Giles 1973; Giles et al. 1973). Similar studies were also carried out by Gregory and Hoyt (1982), Gregory and Webster (1996), Bilous and Krauss (1988), Natale (1975a, 1975b) and Welkowitz and colleagues (Welkowitz and Feldstein 1969; Welkowitz and Feldstein 1970; Welkowitz et al. 1972). Phonetic convergence in conversational interactions was investigated more recently by Pardo and colleagues (Pardo 2006; Pardo 2010; Pardo et al. 2012; Pardo et al. 2013), Kim et al. (2011), Llamas et al. (2009) and Lewandowski and colleagues (Lewandowski 2012; Schweitzer and Lewandowski 2012).

In the late 1990s the process of adjusting one's speech to the speech of others began to be referred to as phonetic imitation. As opposed to accommodation, phonetic imitation is examined in socially minimal, laboratory settings, where the participants are usually required to repeat pre-recorded single words. The focus in a vast number of phonetic imitation studies is on the mechanisms underlying speech production and perception and the phenomenon itself is often treated as an automatic reflex of the human brain rather than a socially or psychologically motivated process (Brouwer et al. 2010; Delvaux and Soquet 2007; Goldinger 1998; Goldinger and Azuma 2004; Honorof et al. 2011; Kim 2011; Mitterer and Ernestus 2008; Nielsen 2011; Shockley et al. 2004). What seems to be of interest in this particular strand of research is that the results of some of the studies have shown phonetic imitation to be sensitive to language structure. For instance, Mitterer and Ernestus (2008) examined convergence in the pronunciation of native speakers of Dutch and found that it

was only the phonologically relevant pronunciation features that were imitated by the participants. Nielsen (2011) reports that native speakers of American English imitated extended VOT values in word-initial voiceless stops but did not imitate reduced VOT values in the same phonetic context.

A number of studies merge the social-psychological aspects of accommodation in conversational interactions with the laboratory-based methodology used in phonetic imitation research. One of such studies was carried out by Namy et al. (2002), who explored the effect of gender on the magnitude of phonetic imitation and found that the participants converged to male talkers more than to female talkers and that female participants were more likely to converge than male participants. Babel (2009) investigated whether racial biases and perceived attractiveness influence the magnitude of convergence in the pronunciation of American English speakers. The results revealed that participants with a pro-black bias were more likely to imitate a black speaker and that the more attractive a given talker was considered, the more the female subjects tended to converge. It was also found that some of the investigated vowels were imitated to a greater extent than others. Similar results were obtained in Babel's subsequent study (Babel 2010), in which she focused on the imitation of Australian English vowels by speakers of New Zealand English. She observed that subjects who were disposed favourably towards Australia converged more than participants with a New Zealand-bias. Babel et al. (2012) confirmed the finding that voices that are considered attractive may induce more imitation and that different vowels may not be imitated to the same extent. Yu et al. (2013) examined the imitation of extended VOT values by speakers of American English and found that personal characteristics and cognitive abilities such as openness and high attention focus contributed to greater imitation effects. Taken together, the studies on phonetic imitation imply that the phenomenon of adjusting one's speech to the speech of others is conditioned by both linguistic and social-psychological factors.

Although the majority of accommodation and imitation studies are concerned with speech adjustments made by native speakers of a given language, several researchers set out to examine speech convergence in L2 speech. Earlier such studies were conducted within Communication Accommodation Theory and examined accent shifts in conversational interactions between native and non-native speakers. The participants in Beebe's (1981) study were Chinese-Thai bilingual children (brought up in Thailand by Chinese parents) who were interviewed in Thai by two female interlocutors, one Thai and one Chinese. The phonetic variables under investigation were 6 Thai vowels. The results of the

study revealed that the subjects converged towards the Chinese interlocutor by making some of the investigated vowels more Chinese-like. Zuengler (1982) investigated the English pronunciation of native speakers of Spanish and Greek, who were interviewed by a native American English interlocutor. It transpired that participants both converged and diverged from the native interviewer and that the direction of accommodation was a function of the strength of ethnic affiliation. More recently, Lewandowski (2012) found that German learners of English converged their pronunciation towards native English interlocutors in conversational interactions. Zając (2013a) sought to determine whether Polish learners of English accommodate their pronunciation to different accents of English. The results suggested that some of the participants converged towards their interlocutors' speech (Canadian English and Standard Southern British English speakers). Kim et al. (2011) studied phonetic convergence in conversations between subjects who had either the same or different regional dialects, and between native and non-native speakers of English. As opposed to the data obtained by Beebe (1981), Zuengler (1982), Lewandowski (2012) and Zając (2013a), Kim et al.'s results revealed that it was only the participants who shared the same language and dialect that were likely to converge.

Several studies on phonetic imitation in non-native speech were carried out recently by Rojczyk and colleagues. Rojczyk (2012a) found that Polish learners imitated a native English talker's realisation of /æ/, while Rojczyk (2012b) observed that native Polish participants imitated English VOT values. Rojczyk et al. (2013) examined immediate and distracted imitation of English unreleased plosives by native Polish speakers. Statistical analysis of the results showed that the participants imitated the phonetic feature under investigation and that distracting the informants impeded convergence to some extent.

3. Aims

The current study follows the experimental procedures used in imitation studies (i.e. eliciting and examining speech adjustments in a socially minimal setting) to investigate phonetic convergence in the pronunciation of Polish learners of English. The general aim of the study is to determine whether the phonetic imitation framework may be used to account for some of the variation present in L2 pronunciation.

Another goal is to examine whether imitation is influenced by the model talker being a native or a non-native speaker of English. On the one hand, foreign-ac-

cented speech is often viewed unfavourably by native and non-native speakers alike (e.g. Chiba et al. 1995; Dalton-Puffer et al. 1997; Gill 1994; Lippi-Green 1997). This could lead to divergence from L2 pronunciation and convergence towards native speech. On the other hand, several accommodation studies show that individuals might be more inclined to converge towards speakers with whom they share a sense of solidarity and that they may tend to accommodate more towards speakers that appear similar to them in some respects (Gregory and Hoyt 1982; Welkowitz and Feldstein 1969; Welkowitz and Feldstein 1970; Welkowitz et al. 1972). A strong sense of identification with a fellow non-native speaker could lead to greater phonetic alignment with foreign-accented speech and might induce the learner to diverge from the native speaker.

The present paper refers to a study whose results were partly discussed in an earlier article (Zając 2013b). The final aim of the current study is to expand on the findings of Zając (2013b) by examining and interpreting the previously obtained results together with the data that was not analysed in the earlier paper.

4. Variables

The phonetic variables under investigation were the duration and quality of four English front vowels (/æ e ɪ iː/), which were examined in two phonetic environments, followed by a voiced alveolar stop and a voiceless alveolar stop. In English (as in many other languages), vocalic elements tend to be considerably shorter before voiceless obstruents than before voiced obstruents (Hogan and Rozsypal 1980; Peterson and Lehiste 1960). Vowel duration in English is also one of the cues for the voicing of the following consonant (Hogan and Rozsypal 1980; Raphael 1972). In contrast, Jassem and Richter (1989) found no significant length differences between vowels preceding underlyingly voiced final obstruents and vowels preceding underlyingly voiceless final obstruents in Polish. One could therefore assume that maintaining a large enough length contrast between vowels followed by voiced consonants and vowels followed by voiceless consonants may prove problematic for Polish learners of English.

The front vowels were selected since Polish learners of English are frequently reported to struggle with their realisation. The low vowel /æ/ is often replaced by Polish speakers either with /a/ or /e/ (e.g. Gonet et al. 2010; Nowacka 2010; Sobkowiak 2001; Weckwerth 2011), which could result in the eradication of the TRAP/DRESS or the TRAP/STRUT contrast in the learner's interlanguage.

With regard to the current study, the tendency could result in the the participants merging /æ/ and /e/ into one category. The high vowel /ɪ/ is often assimilated by Polish speakers with native /i/ (e.g. Nowacka 2010; Sobkowiak 2001), which can make it difficult for Poles to maintain the KIT/FLEECE contrast in English.

5. Participants and procedure

The participants were 20 native speakers of Polish (12 females and 8 males) studying at the Institute of English Studies, University of Lodz. All of the subjects were first-year students with upper intermediate proficiency in English (approximately). The subjects participated in three experimental tasks: a written matching exercise, an auditory naming task, and a shadowing task, which was further subdivided into two phases. In the first task, the participants matched English words (the analysed tokens) to photos that represented their meanings. The purpose of this exercise was to familiarise the informants with the analysed words. In the second task, the participants saw the photos again on the computer screen and were instructed to identify them by using the words from the matching exercise and saying them out loud. The final stage of the experiment (the shadowing task) involved presenting the photos used in the earlier tasks together with a model talker's voice (a native model talker in the first section of the task and a non-native model talker in the second section). The participants' task was to listen to the voice and then identify the word represented in the photo by saying it out loud. The model talkers were two men in their mid-twenties. One of them was a native speaker of Southern British English, while the other was a native speaker of Polish, who spoke with a relatively heavy foreign accent.

6. Stimulus

The stimuli used in the shadowing task were pre-recorded monosyllabic words. The words contained the analysed front vowels flanked by /b/ and /t/ or /d/ (*bad, bat, bed, bet, bead, beat, bid, bit*). The participants could hear each of the investigated words twice, once pronounced by the native model talker and once realised by the non-native model talker.

The vowel durations used by the model talkers are presented in Table 1. The abbreviations NM and NNM stand for the native model talker and the non-native

model talker respectively. The data show that the British model talker used noticeably longer vowels before a voiced obstruent in each of the analysed pairs of words. The Polish model talker's usage of vowel duration was variable, his /æ/ and /e/ were longer when followed by the voiced obstruent, and his /ɪ/ and /iː/ were shorter when followed by the voiced obstruent.

Table 1. Vowel durations in the model talkers' productions (Zając, 2013b)

	NM		NNM	
vowel	b_d	b_t	b_d	b_t
æ	140	98	145	128
e	127	77	138	94
iː	167	145	114	118
ɪ	103	81	81	105

Figures 1 and 2 illustrate they way vowel quality was realised by the model talkers. It can be seen that the British model talker has separate categories for all four vowels. In the case of the Polish model talker, the distributions of /ɪ/ and /iː/ overlap and a similar pattern is also visible with /æ/ and /e/. This indicates that the non-native speaker merged the KIT category with the FLEECE category and the TRAP category with the DRESS category.

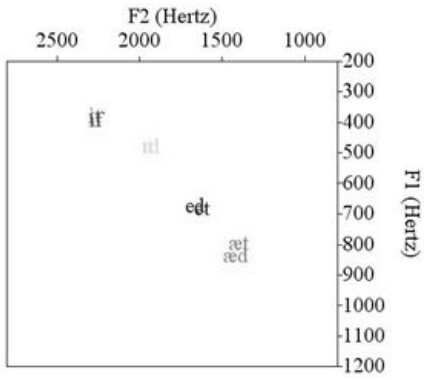


Fig. 1. Formant plot of the native model talker's vowels

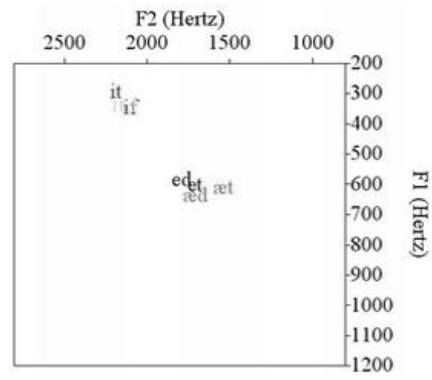


Fig. 2. Formant plot of the non-native model talker's vowels

7. Results and analysis

Table 2 presents mean durations of each of the investigated vowels in two contexts, followed by a voiced stop (b_d) and followed by a voiceless stop (b_t), and under three conditions, prior to exposure to the model talkers' speech (baseline) and following exposure to the model talkers' speech (shadowing NM and shadowing NNM). The values are given in milliseconds, standard deviation is given in brackets. The probability levels for a non-chance difference between the values were calculated with the use of one-tailed paired-samples t-tests. The results indicate that, prior to exposure to the modelled speech, the participants already shortened three out of the four investigated vowels in the context of a following voiceless obstruent. After exposure to the model talkers' pronunciation, the subjects shortened all of the investigated vowels in the context of a following voiceless stop. Interestingly, the participants shortened the vowel in *bit* after listening to the non-native model talker even though he adopted an opposite strategy (Table 1).

Table 2. Participants' mean vowel durations under three conditions (Zając, 2013b)

	baseline			shadowing NM			shadowing NNM		
vowel	b_d N=20	b_t N=20	p	b_d N=20	b_t N=20	p	b_d N=20	b_t N=20	p
æ	202 (46)	162 (38)	**0.000****	160 (31)	143 (25)	**0.001****	170 (33)	136 (23)	**0.000****
e	194 (44)	143 (25)	**0.000****	160 (40)	111 (26)	**0.000****	164 (29)	119 (19)	**0.000****
i:	205 (45)	148 (36)	**0.000****	184 (33)	141 (28)	**0.000****	162 (34)	132 (30)	**0.000****
ɪ	140 (32)	138 (42)	0.423	131 (29)	106 (21)	**0.000****	125 (22)	111 (26)	**0.031***

Table 3 shows the number of participants who exhibited a given vowel contrast under three conditions, prior to exposure to the model talkers' speech (baseline) and following exposure to the model talkers' speech (shadowing NM and shadowing NNM). Whether a particular subject maintained a given contrast or not was determined by examining the participants' vowel plots. The first and the second formants were measured at the midpoint of the vowel and a Praat (Boersma and Weenik 2014) script was subsequently used to compute the vowel plots. The results indicate that the majority of the participants failed to realise the four vowels as separate categories before listening to the modelled speech. After exposure to

the British talker's speech, the number of participants who maintained the KIT – FLEECE contrast increased slightly. However, over half of the subjects still failed to differentiate between /ɪ/ and /iː/. On the other hand, after exposure to the British model talker's speech, the majority of the informants were able to distinguish between /æ/ and /e/ and the number of participants who maintained this contrast became over three times greater than in the baseline productions. Following exposure to the Polish model talker, the number of participants who distinguished between the four vowel categories decreased slightly as compared with the baseline productions. Generally, the vast majority of the subjects failed to realise /ɪ/ and /iː/ and /æ/ and /e/ as separate categories upon exposure to the non-native talker's pronunciation. Overall, it appears that /æ/ and /e/ were differentiated by a greater number of informants than /ɪ/ and /iː/.

Table 3. The number of participants who maintained a given vowel contrast under three conditions

vowel contrast	baseline	shadowing NM	shadowing NNM
KIT – FLEECE	6	9	5
TRAP – DRESS	5	16	2

8. Discussion

The results of the study indicate that the participants adjusted vowel length in their productions after exposure to the model talkers' speech. The subjects shortened all of the investigated vowels in the context of a following voiceless obstruent in the imitation (shadowing) task, which can be interpreted as convergence towards the native English speaker and divergence from the native Polish speaker. As argued in Zając (2013b), it is possible that the participants failed to accommodate towards the non-native model talker out of a desire to sound more native-like. This interpretation seems plausible in view of the finding that some L2 speakers tend to favour native pronunciation over foreign-accented speech (Chiba et al. 1995; Dalton-Puffer et al. 1997; Forde 1995). Additionally, the informants were accompanied by the author of the study throughout the whole experimental procedure. The subjects, first-year students of English studies, most probably believed the author to be a member of the university staff. This, coupled with the formal context of the experiment, could mean that the subjects felt they should try to

diverge from the non-native model talker to create a favourable impression. Such a view is corroborated by Bell's (1984) theory of audience design, according to which speakers may sometimes accommodate to persons in their surroundings with whom they are not in direct interaction at a particular moment.

An important observation is that the subjects in the current study were found to differentiate vowel length in most of the investigated word pairs even before exposure to the native model talker's speech, which implies that this particular feature of English phonology may not be as difficult to acquire for Polish learners as previously expected. Indeed, Slowiaczek and Dinnsen (1985) observed that some vowel length differences before voiced and voiceless obstruents may also exist in Polish, which could facilitate the acquisition of this feature in English.

As regards vowel quality, the results of the current study indicate that exposure to the model talkers' speech caused some subjects to modify the spectral characteristics of their vowels, although it needs to be emphasised that the participants exhibited considerable variability in their accommodation strategies. The majority of the participants converged to the native Polish speaker by merging the two vowel contrasts after exposure to his speech. Over half of the participants diverged from the native English speaker by failing to produce a contrast between /ɪ/ and /iː/ when imitating his speech. At the same time, the majority of the subjects accommodated towards the native model talker by differentiating the TRAP and DRESS vowels. Overall, it would appear that the number of participants who accommodated towards the native Polish speaker was greater than the number of participants who imitated the native English speaker, especially in the case of the KIT/FLEECE contrast.

Taken together, the results of the current study suggest that using vowel duration as a cue for the voicing of the following consonant was a more stable element in the participants' interlanguage than differentiating between the four investigated vowels (the participants used contrasting vowel durations but mostly failed to maintain vowel quality contrasts in their baseline productions). It was also found that the participants diverged from the non-native model talker on vowel duration but mostly converged towards him on vowel quality. This could mean that the magnitude of imitation in L2 speech is more sensitive to affective factors (e.g. attitude towards foreign-accented speech) when the imitated pronunciation feature begins to function as a stable element in the speaker's interlanguage. If the imitated pronunciation feature is not yet a stable element of the interlanguage, it seems that the speaker's convergence strategies are more permeable to transfer from the L1 sound system.

9. Caveats

As referred to in the Results section, the assessment of whether a given informant distinguished between the four front vowels was made by analysing formant plots. In some cases, the selected method proved insufficiently straightforward and objective. For instance, one could interpret the vowel plot in Figure 3 to mean that speaker 13 maintained contrasts between the two vowel pairs since their distributions do not clearly overlap. On the other hand, some of the vowels appear to be very close to each other, which could be taken to mean that the speaker did not distinguish between FLEECE and KIT and TRAP and DRESS. A possible solution to this problem could be to have phonetically trained and/or native English raters listen to the participants' realisations of the word pairs (e.g. *bad* and *bed*) and ask them to decide whether the words are the same or different.

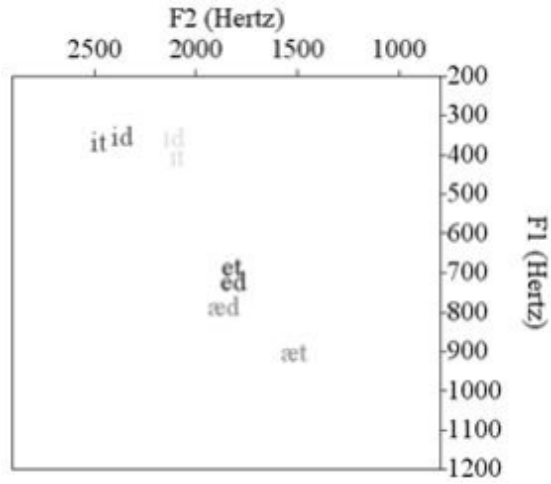


Figure 3. Formant plot of speaker 13's vowels

10. Conclusions

The results of the study indicate that exposure to the speech of different talkers may bring about variability in L2 pronunciation. The participants were found to imitate the duration and quality of four English front vowels when presented with

pre-recorded productions of single words by a native and a non-native speaker of English. The findings of the study suggest that, depending on whether or not the pronunciation feature under investigation functions as a stable element in the learner's interlanguage, the magnitude of imitation in L2 speech may be more susceptible to either the L1 sound system or affective factors such as attitude towards foreign-accented speech.

References

Anderson, J. 1987. The Markedness Differential Hypothesis and Syllable Structure Difficulty. In *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. eds. G. Ioup and, S. Weinberger, 279–291. New York: Newbury House/ Harper & Row.

Babel, M. 2009. Phonetic and Social Selectivity in Speech Accommodation. Unpublished PhD dissertation.

Babel, M. 2010. Dialect divergence and convergence in New Zealand English. *Language in Society* 39: 437–456.

Babel, M., G. McGuire, A. Nicholls, and S. Walters. 2012. Variability in imitation based on voice profile. Presentation at the 2nd Annual Workshop on Sound Change. Seeon-Seebruck, Germany.

Bayley, R. 1996. Competing Constraints on Variation in the Speech of Adult Chinese Learners of English. In *Second Language Acquisition and Linguistic Variation,* eds. R. Bayley, and R. Preston, 75–96. Amsterdam: John Benjamin Publishing Company.

Beebe, L. 1980. Sociolinguistic Variation and Style Shifting in Second Language Acquisition. *Language Learning* 30: 433–448.

Beebe, L. 1981. Social and Situational Factors Affecting the Communicative Strategy of Dialect Code-Switching. *International Journal of the Sociology of Language* 32: 139–149.

Bell, A. 1984. Language style as audience design. *Language in Society* 13: 145–204.

Benson, B. 1988. Universal Preference for the Open Syllables as an Independent Process in Interlanguage Phonology. *Language Learning* 38: 221–235.

Best, C. 1995. A Direct Realist Perspective on Cross-Language Speech Perception. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues,* ed. W. Strange, 167–200. Timonium, MD: York Press.

Bilous, F., and R. Krauss. 1988. Dominance and accommodation in the conversational behaviours of same- or mixed-gender dyads. *Language & Communication* 8: 183–194.

Boersma, P., and D. Weenink. 2014. Praat: doing phonetics by computer [Computer program]. Version 5.3.80. http://www.praat.org/. Accessed July 2014.

Bourhis, R., and H. Giles. 1977. The language of intergroup distinctiveness. In *Language, ethnicity, and intergroup relations*, ed. H. Giles, 119–136. Waltham: Academic Press.

Brouwer, S., H. Mitterer, and F. Huettig. 2010. Shadowing reduced speech and alignment. *The Journal of the Acoustical Society of America* 128: EL32-EL37.

Brown, C. 2000. The Interaction Between Speech Perception and Phonological Acquisition from Infant to Adult. In *Second Language Acquisition and Linguistic Theory.* ed. J. Archibald, 4–63. Oxford: Blackwell.

Chiba, R., H. Matsuura, and A. Yamamoto. 1995. Japanese attitudes towards English accents. *World Englishes* 14: 77–86.

Coupland, N. 1984. Accommodation at work: Some phonological data and their implications. *International Journal of the Sociology of Language* 46: 49–70.

Dalton-Puffer, C., G. Kaltenboeck, and U. Smit. 1997. Learner attitudes and L2 pronunciation in Austria. *World Englishes* 16:, 115–128.

Davidian, D., and J.E. Flege. 1984. Transfer and Developmental Processes in Adult Foreign Language Speech Production. *Applied Psycholinguistics* 5: 323–347.

Delvaux, V., and A. Soquet. 2007. The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica* 64: 145–73.

Dickerson, L., and W. Dickerson. 1977. Interlanguage phonology: current research and future directions. *The Notions of Simplification. Interlanguage and Pidgins: Actes du 5ème Colloque de Linguistique Applique de Neufchatel*: 18–30.

Dowd, J. 1984. Phonological Variation in L2 Speech: The Effects of Emotional Questions and Field-dependence/Field-independence on Second Language Performance. Unpublished doctoral dissertation.

Elliott, R. 1995. Field Independence/Dependence, Hemispheric Specialization, and Attitude in Relation to Pronunciation Accuracy in Spanish as a Foreign Language. *The Modern Language Journal* 79: 356–371.

Flege, J.E. 1988. Factors Affecting Degree of Perceived Foreign Accent in English Sentences. *Journal of the Acoustical Society of America* 84: 70–79.

Flege, J.E. 1993. Production and perception of novel, second-language phonetic contrast. *Journal of the Acoustical Society of America* 93: 1589–1608.

Forde, K. 1995. A study of learner attitudes towards accents of English. *Hong-Kong Polytechnic University Working Papers in ELT & Applied Linguistics* 1: 59–76.

Fox, R., J.E. Flege, and M. Munro. 1995. The Perception of English and Spanish Vowels by Native English and Spanish Listeners: A Multidimensional Scaling Analysis. *Journal of the Acoustical Society of America* 97: 2540–2551.

Gatbonton, E. 1975. Systematic Variations in Second Language Speech: A Sociolinguistic Study. Unpublished doctoral dissertation.

Giles, H. 1973. Accent mobility: a model and some data. *Anthropological Linguistics* 15: 87–105.

Giles, H., M. Taylor, and R. Bourhis. 1973. Towards a theory of interpersonal accommodation through language: some Canadian data. *Language in Society* 2: 177–192.

Gill, M.M. 1994. Accent and stereotypes: Their effect on perceptions of teachers and lecture comprehension. *Journal of Applied Communication Research* 22: 349–361.

Goldinger, S. 1998. Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological Review* 105: 251–279.

Goldinger, S., and T. Azuma. 2004. Episodic memory in printed word naming. *Psychonomic Bulletin & Review* 11: 716–722.

Gonet, W., J. Szpyra-Kozłowska, and R. Święciński. 2010. Clashes with ashes. In *Issues in Accents of English 2: Variability and Norm*, ed. E. Waniek-Klimczak, 213–230. Newcastle upon Tyne: Cambridge Scholars Publishing.

Gregory, S., and B. Hoyt. 1982. Conversation Partner Mutual Adaptation as Demonstrated by Fourier Series Analysis. *Journal of Psycholinguistic Research* 11: 35–46.

Gregory, S., and S. Webster. 1996. A Nonverbal Signal in Voices of Interview Partners Effectively Predicts Communication Accommodation and Social Status Perceptions. *Journal of Personality and Social Psychology* 70: 1231–1240.

Hartford, B. 1978. Phonological Differences in the English of Adolescent Female and Male Mexican-Americans. *International Journal of the Sociology of Language* 17: 55–64.

Hogan, J.T., and A.J. Rozsypal. 1980. Evaluation of vowel duration as a cue for the voicing distinction in the following word-final consonant. *Journal of the Acoustical Society of America* 67: 1764–1771.

Honorof, D., J. Weihing, and C. Fowler. 2011. Articulatory events are imitated under rapid shadowing. *Journal of Phonetics* 39: 18–38.

Iverson, P., P. Kuhl, R. Yamada, E. Diesch, Y. Tohkura, A. Ketterman, and C. Siebert. 2001. A Perceptual Interference Account of Acquisition Difficulties for Non-Native Phonemes. *Speech, Hearing and Language: Work in Progress* 13: 106–118.

Jassem, W., and L. Richter. 1989. Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17: 205–212.

Kim, M. 2011. Phonetic convergence after perceptual exposure to native and non-native speech: preliminary findings based on fine-grained acoustic-phonetic measurement. *Proceedings of the 17th International Congress of Phonetic Sciences*: 1074–1077.

Kim, M., W.S. Horton, and A.R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Journal of Laboratory Phonology* 2: 125–156.

Kuhl, P. 2000. A New View of Language Acquisition. *Proceedings of the National Academy of Science* 97: 11850–11857.

Lewandowski, N. 2012. Talent in nonnative phonetic convergence. Unpublished PhD dissertation.

Livbjerg, I., and I. Mees. 1988. *Practical English Phonetics. Ny Kontrastiv Fonetik*. Kopenhagen: Schonberg.

Lippi-Green, R. 1997. *English with an accent: Language, ideology, and discrimination in the United States*. New York: Routledge.

Llamas, C., D. Watt, and D.E. Johnson. 2009. Linguistic Accommodation and the Salience of National Identity Markers in a Border Town. *Journal of Language and Social Psychology* 28: 381–407.

Mitterer, H., and M. Ernestus. 2008. The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. Cognition 109: 168–173.

Moyer, A. 1999. Ultimate Attainment in L2 Phonology. *Studies in Second Language Acquisition* 21: 81–108.

Namy, L., L. Nygaard, and D. Sauerteig. 2002. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology* 21: 422–432.

Natale, M. 1975a. Convergence of Mean Vocal Intensity in Dyadic Communication as a Function of Social Desirability. *Journal of Personality and Social Psychology* 32: 790–804.

Natale, M. 1975b. Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills* 40: 827–830.

Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39: 132–142.

Nowacka, M. 2010. The ultimate attainment of English pronunciation by Polish College Students: a longitudinal study. In *Issues in Accents of English 2. Variability and Norm,* ed. E. Waniek-Klimczak, 233–260. Newcastle upon Tyne: Cambridge Scholars Publishing.

Oyama, S. 1976. A Sensitive Period for the Acquisition of a Nonnative Phonological System. *Journal of Psycholinguistic Research* 5: 261–283.

Pardo, J. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119: 2382–2393.

Pardo, J. 2010. Expressing oneself in conversational interaction. In *Expressing Oneself/Expressing One's Self. Communication, Cognition, Language, and Identity,* ed. E. Morsella, 183–196. New York: Psychology Press.

Pardo, J., R. Gibbons, A. Suppes, and R. Krauss. 2012. Phonetic convergence in college roommates. *Journal of Phonetics* 40: 190–197.

Pardo, J., I.C. Jay, R. Hoshino, S. Hasbun, C. Sowemimo-Coker, and R. Krauss. 2013. The Influence of Role-Switching on Phonetic Convergence in Conversation. *Discourse Processes* 50: 276–300.

Peterson, G.E., and I. Lehiste. 1960. Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America* 32: 693–703.

Piske, T., I. MacKay, and J.E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215.

Purcell, E., and R. Suter. 1980. Predictors of Pronunciation Accuracy: A Reexamination. *Language Learning* 30: 271–287.

Raphael, L.J. 1972. Preceding vowel duration as a cue to the perception of voicing of word-final consonants in American English. *Journal of the Acoustical Society of America* 51: 1296–1303.

Rochet, B. 1995. Perception and Production of L2 Speech Sounds by Adults. In *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, ed. W. Strange, 379–410. Timonium, MD: York Press.

Rojczyk, A. 2012a. Spontaneous phonetic imitation of L2 vowels in a rapid shadowing task. Poster presented at PSLLT 2012 – Pronunciation in Second Language Learning and Teaching Conference, Vancouver, Canada, 24–25 August.

Rojczyk, A. 2012b. Phonetic and phonological mode in second-language speech: VOT imitation. Paper presented at EuroSLA22 – 22nd Annual Conference of the European Second Language Association, Poznań, Poland, 5–8 September.

Rojczyk, A., A. Porzuczek, and M. Bergier. 2013. Immediate and distracted imitation in second-language speech: Unreleased plosives in English. *Research in Language* 11: 3–18.

Schweitzer, A., and N. Lewandowski. 2012. Accommodation of Backchannels in Spontaneous Speech. In the booklet of the International Symposium on Imitation and Convergence in Speech.

Shockley, K., L. Sabadini, and C.A. Fowler. 2004. Imitation in shadowing words. *Perception & Psychophysics* 66: 422–429.

Slowiaczek, L., and D. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13: 325–341.

Sobkowiak, W. 2001. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Suter, R. 1976. Predictors of Pronunciation Accuracy in Second Language Learning. *Language Learning* 26: 233–253.

Tahta, S., M. Wood, and K. Loewenthal. 1981. Foreign Accents: Factors Relating to Transfer of Accent from the First Language to a Second Language. *Language & Speech* 24: 265–272.

Tarone, E. 1982. Systematicity and attention in interlanguage. *Language Learning* 32: 69–84.

Thompson, I. 1991. Foreign Accents Revisited: The English Pronunciation of Russian Immigrants. *Language Learning* 41: 177–204.

Thompson, R. 1976. Mexican-American English: Social Correlates of Regional Pronunciation. *American Speech* 50: 18–24.

Weckwerth, J. 2011. English TRAP Vowel in Advanced Polish Learners: Variation and System Typology. City University of Hong Kong, Volume Proceedings of the 17th International Congress of Phonetic Sciences. 17–21 August 2011. Hong Kong. CD-ROM, Hong Kong, p.2110–2113 (2011)

Welkowitz, J., and S. Feldstein. 1969. Dyadic interaction and induced differences in perceived similarity. *Proceedings of the 77th Annual Convention of the American Psychological Association* 4: 343–344.

Welkowitz, J., and S. Feldstein. 1970. The relation of experimentally manipulated interpersonal perception and psychological differentiation to the temporal patterning of conversation. *Proceedings of the 78th Annual Convention of the American Psychological Association* 5: 387–388.

Welkowitz, J., M. Finklestein, S. Feldstein, and L. Aylesworth. 1972. Changes in vocal intensity as a function of interspeaker influence. *Perceptual and Motor Skills* 35: 715–718.

Yu, A.C., C. Abrego-Collier, and M. Sonderegger. 2013. Phonetic Imitation from an Individual-Difference Percpective: Subjective Attitude, Personality and "Autistic" Traits. *PLoS ONE* 8: 1–13.

Zając, M. 2013a. [ˈberɚ] or [ˈbetə]? Do Polish Learners of English Accommodate their Pronunciation? A Pilot Study. In *Teaching and Researching English Accents in Native and Non-native Speakers*, eds. E. Waniek-Klimczak, and L.R. Shockey, 229–239. Heidelberg: Springer.

Zając, M. 2013b. Phonetic imitation of vowel duration in L2 speech. *Research in Language* 11: 19–29.

Zuengler, J. 1982. Applying Accommodation Theory to Variable Performance Data in L2. *Studies in Second Language Acquisition* 4(2): 181–192.

LINGUISTICS

PHONETICS, DIALECTOLOGY, HISTORICAL LINGUISTICS

Synchronic variability in the area of phonetics, phonology, vocabulary, morphology and syntax is a natural feature of any language, including English. The existence of competing variants is in itself a fascinating phenomenon, but it is also a prerequisite for diachronic changes. This volume is a collection of studies which investigate variability from a contemporary and historical perspective, in both native and non-native varieties of English. The topics include Middle English spelling variation, lexical differences between Middle English dialects, Late Middle and Early Modern English forms of address, Middle English negation patterns, the English used by Polish immigrants living in London, lexical fixedness in native and non-native English used by Polish learners, and the phenomenon of phonetic imitation in Polish learners of English. The book should be of interest to anyone interested in English linguistics, especially English phonetics and phonology as well as history of English, historical dialectology and pragmatics.

The book is also available as an e-book

WYDAWNICTWO UNIWERSYTETU ŁÓDZKIEGO

ul. Williama Lindleya 8
90-131 Łódź

tel.: 42 66 55 863
e-mail: ksiegarnia@uni.lodz.pl

ISBN 978-83-8088-065-8