CONTENTS

LETTER FROM THE EDITORS

Dear Readers,

The present issue of VIEWS is an extensive one, both in terms of its length and in terms of the time span covered, which extends from the Old English period via Middle English and Late Modern English to the present day. As such, this issue is a truly diachronic one.

Let us look at it chronologically. Herbert Schendl investigates code-switching in Old English. In doing so, he takes up a topic he has already

explored in previous issues of VIEWS (nos. 5 & 6), more precisely code-switching in Early-Modern English and Middle English. In the meantime historical code-switching has established itself as a promising field of historical linguistics, with research, however, focusing mainly on the Middle English period. In his present contribution Herbert Schendl extends the investigation to Old English, demonstrating that code-switching was by no means alien to the Anglo-Saxons either, as is generally thought.

The Middle English period is dealt with in the paper by Ursula Lutzky, who probes into the occurrence and use of the negative prefixes *dis-, in-, mis-* and *un-*. In particular her paper, which is based on material from the *Helsinki Corpus,* focuses on the productivity of these prefixes, tracing their development in the Middle English period and investigating the competition between native and foreign ones. As a further extension of her analysis she also takes into account 'text type' as a possible conditioning factor for productivity.

The Late Modern English period is taken care of, as it were, by Stefan Dollinger and his article on problems involved in the compilation of a historical corpus. Based on his experience in transcribing early Canadian manuscripts for his *Corpus of Early Ontario English*, Stefan Dollinger discusses some of the challenges in connection with transcribing Late Modern English handwriting for machine-readable corpora. As a particular test case he focuses on the discrimination of upper and lower case letters, which is shown to depend on a variety of different factors.

Ute Smit's paper, finally, takes us up to the present day. She explores the role English plays in Viennese schools and, comparing it with official educational language policies and the actual multilingualism found in these schools, detects a number of discrepancies. She shows for instance that the dominant role accorded to English in the curriculum far from supports the officially voiced aim of a multilingual Europe, but in fact undermines it by providing for English-German bilingualism instead.

All in all, this issue – we hope – provides plenty of food for thought, thoughts that might even spark off a written reply from you. If so, we will be happy to include your comments in the next issue.

THE EDITORS

'Philological computing' vs. 'philological outsourcing' and the compilation of historical corpora: a Late Modern English test case[1]

Stefan Dollinger, University of Vienna[*]

1. Historical corpus linguistics: Corpora galore[2]

The corpus-based study of the history of English is not only a well-established, but also a highly prosperous field today (e.g. Rissanen *et al*. 1992, Raumolin-Brunberg 2002). Despite a somewhat belated start after the first machine-readable corpora of PDE were available in the early 1960s, historical corpus linguistics is alive and well, as evidenced by the entries in the *Oxford Text Archive* or the third edition of the ICAME CD-ROM collection of English corpora, which is currently being compiled.

It may therefore be time to review some of the practices of corpus compilation that seem to have become accepted in the field since the ground-breaking *Helsinki Corpus* was completed in the late 1980s. Since then, a number of historical corpora have been compiled; some of them have already been made publicly available, while others are being worked on. Adapting Kennedy's periodization (1998) of machine-readable corpora for historical corpus linguistics, we may call the younger 'post-Helsinki' corpora second generation corpora. This division is based on the second generation corpora filling temporal, geographical or sociolinguistic gaps that the *Helsinki Corpus* (*HC*) did not cover, such as the lack of 18th and 19th century data, the geographical limitation to texts from Britain, or the lack of lower class speaker data.

The *ARCHER* corpus was partly designed to remedy this lack of readily-available data, but the unsuccessful attempt of copyright clearance prevented its broader distribution. *ARCHER II* is currently being compiled to circumvent the copyright problem (Meyer 2002: 142). Regardless of unedifying legal issues like these, gaps concerning non-dominant varieties or lesser-researched periods of English are increasingly being filled (e.g. Kytö, Rudanko and Smitterberg 2000, Claridge 2001, Laitinen 2002, Dossena 2003: 91, Curzan and Hakala 2004: 146, Dollinger forthc.)[3]

It seems that the text genre of letters has gained special currency in respect to the Late Modern English (LModE) period, as can be seen by Ingrid Tieken's online database project[4], or the extension of the *Corpus of Early English Correspondence* into the 18th century (Laitinen 2002). Within the letter genre, however, research interest is shifting towards letters of the lower classes, as evidenced by Austin (1991, 2000), Fairman (2000, 2003) or Tieken (forthc.). As a result, the first machine-readable corpora of LModE lower class letters are being compiled, such as Mikko Laitinen's corpus project on 19th-century Essex pauper letters (Laitinen 2003), or the letter section in my own *Corpus of Early Ontario English* (*Conte*).

The letter section of *Conte* features some 50,000 words of early Ontario correspondence in its pre-Confederation, 1776-1850. Based on the experience gained when transcribing these early Canadian manuscript letters from microfilm, I would like to illustrate some of the challenges and problems of transcription of LModE handwriting for machine-readable corpora for linguistic studies. In section 2, the current general practice of corpus compilation will be discussed, section 3 briefly reflects on some general issues concerning LModE manuscripts (MSS). In section 4, the current practice of using edited texts for historical corpora is tested on the basis of an example from *ARCHER*. While the outcome may suggest the more frequent use of manuscripts, section 5 takes a look at the problems of transcription of upper and lower case letters from LModE MSS. While we will see that manuscripts are indeed to be preferred, the examples provided will put the meaning of the corpus compiler's goal of 'absolute fidelity to the original'

---

3 For a complete survey for up to 2001 see Meyer (2002: 142-150) and the yearly progress reports in the *ICAME Journal* for more recent contributions. It needs to be added that one corpus on the forefront of historical corpus linguistics, the *Corpus of Early American English* [1620-1720] (Kytö 1991 [1988]) is currently undergoing computerization (Kytö: personal communication).

4 http://www.let.leidenuniv.nl/hsl_shl/correspondences.htm, 9 December 2004.

(Meyer 2002: 78) in the context of LModE handwriting and its computerization.

2. Computerization of *what?* Practices of corpus compilation

Corpus compilation in general is, needless to say, a time and labour intensive task. When confronted with historical data, compilers also need to bear in mind the philological aspects of their source texts. Obviously, modern corpus compilers are not the first to discover these problems, as scholars have been editing literary texts for centuries (cf. Greetham 1994 for some general problems of editing, Walsh 1997 for early editorial practices and Shillingsburg 1996 for some challenges faced by editors in the computer age).

Editors are confronted with specific challenges for each period. For OE and ME, one is confronted with phenomena such as multiple versions of manuscripts, scribal errors or poor physical document quality. Considering these problems, the choice to resort to edited material, i.e. to 'outsource' the philological aspects of the source texts and to trust the editors of a given text, seems plausible. This was the practice of compilation of the *HC*, and the compilers are very clear about this:

> The texts keyed in at Helsinki are based on the best editions available. In most cases the use of original manuscripts has not been possible. [...] It is obvious, however, that a certain number of errors inevitably remains. (Kytö 1991: 3)

This quotation from the manual of the *HC* points to the logistic constraints compilers are often confronted with. Likewise, a similar approach was taken for the *Corpus of Early English Correspondence (CEEC)*: "As a rule, the letters were selected and scanned from edited collections" (Nevalainen and Raumolin-Brunberg 2003: 44) so that the "vast majority" (Keränen 1998a: 35) is published material, and only some letters were specifically edited for the *CEEC* (e.g. Nevala 2001).

Here, we see examples of what may be termed good practice, as the corpus user is informed about the actual 'content' of the corpus. At the same time we see the different needs of corpus compilers on the one hand and philologists on the other hand. The compiler is confronted with the task of computerization and would like to use, and in many cases due to time and labour constraints is bound to use, the work of philologists as a base. If an edition of a given text can be found, why should any time be dedicated to the transcription of texts from manuscript sources?

The 'hunt' for suitable editions has at times resulted in somewhat peculiar expectations of older editions. If an edition does not fulfill one's needs to 100%, compilers might tend to "partly [...] blame" the editors (Keränen

1998a: 35), who were, however, only editing for their own purposes, which may have been very different to what corpus compilers would wish. Keränen's statement reflects very well the idea of 'outsourcing' one's philological needs. As a consequence, compilers are willing to invest some effort into unearthing lesser-known editions of suitable texts, rather than running the risk of re-editing material (cf. Susan Pintzuk's (2004) query for printed Latin editions). As a consequence, our corpora are only as good as the editions and their editorial principles, which are not always made explicit.

This approach to data is not the best practice for many reasons, most notably because of the danger of interference on account of different editing practices at different stages in philological history. Kytö, Rudanko and Smitterberg (2000: 95) point out that

> [...] researchers should always keep in mind possible effects of the editorial process
> 19th-century texts went through when being transferred into written form.

If we take a look at another innovative project in historical corpus linguistics, the *ARCHER Corpus*, conceived in the early 1990s, we see that the practice of using editions has somehow become paradigmatic in certain areas of historical corpus linguistics. The available documentation of the corpus contains no clear statement about the use of editions and/or manuscripts (cf. Biber, Finegan and Atkinson 1994, Biber and Finegan 1995, Biber and Finegan 1997, Biber, Conrad and Reppen 1998: 251-253, 216-222 on letters, Biber and Burges 2000: 23-24).[5] However, descriptions of the sampling methods suggest that only, or at least predominantly, published material has been used, e.g. if we interpret Biber and Finegans's statement that "available bibliographies serving as sampling frames" (1997: 256) were used and these bibliographies contain no information on manuscripts (Besterman 1971: I, II, III, IV, V).[6]

In this light, it seems doubtful whether what I termed 'philological outsourcing' is good practice in corpus linguistics, as long as the compilers do not specify clearly what was done. Consequently, we might modify Meyer's

---

[5] It should be stated here that due to the problems of publication the lack of a corpus manual for *ARCHER* is understandable. It could be expected to find this kind of information there.

[6] Biber, Finegan and Atkinson (1994: 4) speak of "date of original publication or (as with diaries) composition" and of the "research articles published in a single year" (7), which may indicate that only or predominantly published material was used. The only way to clearly answer this question is to go through the corpus files one by one, checking the headers for text sources and examining these sources for their editorial principles.

statement that "the corpus compiler is concerned with absolute fidelity to the original" (Meyer 2002: 78) in the way that the corpus compiler is *de facto* concerned with absolute fidelity to the *edition*, because this is what is usually worked with, augmented by random checks against manuscripts. In doing so, the compilers may have opened the door to a phenomenon that Meurman-Solin, following Lass, refers to as the computerization of 'second-order constructs' or 'artefacts' (2001: 19). If I understand Meurman-Solin's concept correctly, she argues for the use of authentic letters only (20) and for the careful, meticulous annotation of data, if need be, even at the expense of corpus size (25, point 4). Meurman-Solin discriminates strictly between language-external and language-internal information of a text as being of a categorically different kind and identifies a problem of current corpus compilation as the problem of

> [...] marginalizing information of the first kind, i.e. that related to text histories, and highlighting information of the second kind, for instance, information about variables such as genre and text category. (Meurman-Solin 2001: 20)

We could say that Meurman-Solin would want to establish a closer link between philology and historical corpus linguistics. The view argued for in the present paper is similar to Meurman-Solin's concept of 'philological computing' (2001: 18-24). The method to be adopted also relies strongly on language-external information, albeit this kind of information is at times difficult to obtain. Because of a lack of editions for certain data, however, we are bound to unearth as much information of the 'first kind' about text and writer as possible, aiming at limiting the risk of exaggerating 'second kind' information.

Seen from this angle, the fact that for LModE manuscripts editions frequently do not exist may actually render us a good service. In the case of the *Corpus of Early Ontario English (Conte)* I am convinced that it did, although at the expense of corpus size and time coverage.

3. Handwriting in Late Modern English

Before moving on to more specific problems of LModE manuscripts, I shall here briefly demonstrate some more general issues of late 18th and early 19th century handwriting in relation to computerization.

A major problem when transcribing LModE MSS is the distinction between upper and lower case letters. We will see in section 5 that LModE confronts the compiler/editor with quite challenging problems in this respect. A second issue is the question of rendering non-ASCII characters. Although this might strike the reader as unusual for LModE and more prominent in OE

or ME, there are still some special characters that need to be transcribed and computerized, such as artistically-decorated initial letters, graphic mergers of characters or abbreviations in closing formulas and job titles. Moreover, syntacticians might be interested in 'proper' ways to interpret LModE punctuation, which is another potential problem area as sometimes sentences seem to start midway and periods and commas are, if present at all, visually hard to discriminate. All of these features are probably found in OE and ME texts.[7]

Apart from the examples above, there are also a number of more specific features and problems of LModE handwriting. Most obviously, the different sociolinguistic background in LModE times and the shorter text histories of LModE texts usually result in the existence of considerably fewer copies than in earlier times. We may even go as far as to claim for LModE letters that usually only one or two copies exist, depending on the type of letter subgenre. In *Conte* we have two subtypes: official and semi-official letters. With official letters, i.e. those written by governmental clerks, one can safely assume on the basis of letters that mention copy-books that correspondence, if not in its entirety then at least the more important pieces, was copied for later reference. In this case, we may assume that generally one or two copies exist of a single letter, i.e. the original and the copy as such. For semi-official letters, i.e. those written by citizens to the authorities, the likelihood that a copy exists is somewhat smaller. For this subgenre we can assume that more likely one than two copies exists.

In certain cases more than two copies may have been written, for instance if a copy was made at the request of a third person or a superior in order to have it forwarded. It seems clear, however, that the exception for earlier stages of the language, e.g. the unique Beowulf MSS for OE, is the norm for LModE handwriting, since printing had long before taken over the function to supply readers with multiple copies where needed.

Another point is the quality of the extant material. Although I transcribed from microfilm, occasional comparisons with the originals show that handwriting is *less* satisfactorily preserved in LModE documents than in those of earlier periods of English. It seems that with the increasingly industrialized production of writing utensils, paper, quills and later steel pens, the endurability of products was, in the long term, actually *reduced* rather than increased. Especially the ink is a case in point: while the *Beowulf*

---

[7] While problems of computerization for ME texts have been treated in Markus (1997), there appear to be similarities to LModE that will be treated elsewhere.

manuscript survived some 1000 years including a severe fire largely untouched, LModE ink has already faded after less than 200 years to an extent that the writing has become illegible. What is still legible is, as will be seen in the scans to follow, often of poor physical quality when compared to most OE manuscripts. We are therefore confronted with a paradox in the sense that progress of writing technology has actually decreased the longevity of documents for the period under survey, 1776-1850.

A final problem of this later period is the multiplicity of 'fonts'. Until LModE times, the diversification of writing systems, hands, and various calligraphic influences led to an array of new fonts and styles so that, in contrast to medieval hands, one is confronted with probably more hands than before, which require more time to single out a particular system. Strictly speaking one must start from scratch for each letter.

4. Philological Outsourcing: an example from LModE

Some problems of using editions for historical corpus compilation may be highlighted by reviewing a LModE example from the *ARCHER Corpus*.[8] As regards EModE and earlier editions of source texts, some problems have been addressed (Keränen 1998b: 219-220), so that here I will focus on the practice of using the best available editions of a given text. The test case is taken from the work of a very well documented English poet, William Blake. The editions discussed here may be considered to be products of the heyday of philological accuracy and may therefore serve as best practice examples.

In the British English section of the *ARCHER-1*, period 1800-1849, we find a letter by William Blake, which is reproduced below from the corpus:

<The letters of William Blake. 1956. Geoffrey Keynes (ed.).
  New York: Macmillan. pp. 43, 239 words>
<To William Hayley   May 6, 1800>
[1] Dear Sir.,
[2] I am very sorry for your immense loss, which is a
[3] repetition of what all feel in this valley of misery & happiness
[4] mixed. I send the Shadow of the departed Angel:
[5] hope the likeness is improved. The lip I have again
[6] lessened as you advised & done a good many other softenings

---

[7] to the whole. I know that our deceased friends are
[8] more really with us than when they were apparent to our
[9] mortal part. Thirteen years ago I lost a brother & with
[10] his spirit I converse daily & hourly in the Spirit & See
[11] him in my remembrance in the regions of my Imagination.
[12] I hear his advice & even now write from his Dictate.
[13] Forgive me for Expressing to you my Enthusiasm which I
[14] wish all to partake of Since it is to me a Source of Immortal
[15] joy: even in this world by it I am the companion
[16] of Angels. May you continue to be so more & more & to
[17] be more & more perswaded that every Mortal loss is an
[18] Immortal Gain. The Ruins of Time builds Mansions in
[19] Eternity. - I have also sent A Proof of Pericles for your
[20] Remarks, thanking you for the Kindness with which you
[21] Express them & feeling heartily your Grief with a brother's
[22] Sympathy.
[23]   I remain, Dear Sir, Your humble Servant
[24]   William Blake
[25] Lambeth. May 6. 1800

(*ARCHER-1*, <1800Blak.x5>, line numbers added for reference)

If we compare the above passage with other sources of the letter, we notice some interesting points. First of all, Keynes' letter collection, which is the source of the *ARCHER* text, has been available in its second edition since 1968, i.e. "revised and amplified" and in its third since 1980, "with related documents". The original MS was bought by Harvard University Libraries in 1956 and Keynes's transcription seems to be based on a "photostat" of the letter (Keynes 1980: 186). Since *ARCHER* is based on the first edition from 1956, we may conclude that the *ARCHER* "universe" (Biber, Finegan, Atkinson 1994: 5), i.e. the major research libraries of Southern California, did not contain the newer editions.

If we compare the *ARCHER* text with Keynes' third edition, we notice two little irregularities: first, there is no period after "Dear Sir" in line 1 and, second, line 15, "joy" is found with an upper case initial letter, i.e. "Joy". Both the first and second editions (Keynes 1956 and 1968) show these features so that we may conclude that the *ARCHER* text contains typing errors here.

What other sources for comparison do we have? Since a facsimile copy of this letter does not seem to exist (cf. Keynes 1926), we can only compare the *ARCHER* transcript to other editions than the ones by Keynes. Erdman, whose

goal was to create a "text as close as possible to Blake's own, even in punctuation" (1965: 709), gives us a variant that deviates in several points from the *ARCHER* version:

- line 1: no punctuation at all after "Dear Sir", just a line break
- line 4: no colon, but period after "Angel"; however, "hope" also starts with a lower case <h> (line 5)
- line 20: no comma after "Remarks"
- line 20: "Kindness" with lower case initial, thus "kindness"
- line 21: "brother's" without apostrophe, therefore "brothers"
- line 23: no commas: "I remain Dear Sir your humble Servant"

This example illustrates something like the maximum accuracy one may obtain by relying on philological work. Both Keynes and Erdman are established Blake scholars at the height of the philological tradition. And nevertheless we have between two and six irregularities in a text of 233 [sic!] words between the corpus text and the editions. Of the six irregularities between the editions, four concern punctuation (lines 1, 4, 20, 23). That the sentence in line 5 is started with a lower case <h> is an issue that we will pursue in section 5 in a different context, as it possibly relates to the upper/lower case difference in line 20, *kindness*. The case of line 21, however, is of more immediate interest to the syntactician. The irregularity in the phrase "with a brothers Sympathy" allows us to hypothesize on the status of possessive –*s* at around 1800 that would need to be checked against more data. It would need to be decided whether the lack of an apostrophe is merely a simple scribal error, i.e. what Fairman (2003: 266) calls, in analogy to a typo, a 'writo', or whether this variation represents a step in the evolution towards possessive –*s* in English. This is a type of question that historical linguists are interested in and that would have been missed if they had used *ARCHER* or the edition at its base.

This example has shown that even best practice editions bear potential dangers. If we go back farther in time, e.g. to 19th-century editions of letters, we will see more than the comparatively few irregularities that we have just seen.[9] One could always argue that, had we transcribed the letter ourselves from original or copy, we might have produced more than six irregularities, depending on our "surrogate competence" of Blake's handwriting (Martti

---

9 However, it needs to be added here that, according to previous research, one cannot equate older, e.g. 19th-century, editions with poor quality. Some of them are still best practice examples to the present day (Keränen 1998b: 218).

Nyman, qtd. in Nevalainen and Raumolin-Brunberg 2003: 9) and the writing systems of the period. Nevertheless, linguists transcribing for linguistic research would bring a different focus to the transcription task that would possibly render us with more reliable data in our corpora.

In the next section of this paper, I would therefore like to draw attention to the issue of upper and lower case letters that was one of the irregularities in Blake's letter, since it seems to be a recurring issue for transcribers of LModE texts.

5. Upper or lower case? Towards philological computing

Before we turn to the problems of discriminating between letters and their variants, we need to define the terminology. A *writing system* is here defined as the underlying system of graphemic representations a writer uses when writing by hand. This system is always expressed, i.e. realized, in a given person's handwriting, i.e. his or her *hand*. The modern computer user would refer to 'hand' as font. This definition of writing system is different to the way the term is used in much of the literature on handwriting (Croft 1973, Heal 1962 or Hector 1958), where usually no distinction is made between *writing system* and *hand*.

Signs on paper are *allographs* of a given grapheme. A *grapheme* is here defined as an abstraction of a sign on paper that belongs to the same class. For instance, an [A] and an [a] are allographs of the same grapheme, abstract <Aa>. There are upper case (UC) and lower case (LC) allographs. This means we have two distinct classes of allographs: a class of UC allographs, and one of LC allographs. The entirety of a writer's allographs, his or her realization of graphemes, may be called *allographic inventory*.

Capitalization is a feature that is usually consciously employed by the writer. For corpus compilers, discrimination of upper and lower case seems to be a recurring question and best practice has been to transcribe UC and LC as closely as possible, thus preserving the author's intention as closely as possible (e.g. Nevala 2001: 166 for EModE). Unless one is willing to resort to the opinion expressed in Görlach (2001: 75) that capitalization "serves the additional function of aiding legibility" in handwriting, one will have to make sense of UC and LC letters.

Osselton (1984, 1985) considers emphasis to be a major function of capitalization. For letters, Tieken (1998) identifies a correlation between capitalization practice and the relationship of the writer to the addressee. In Dollinger (2003) I tentatively examined the idea that individual writers may

have had individual allographic inventories, where not every grapheme was realized by an UC and a LC variant.[10] Here, the idea shall be pursued in which sense different practices for capitalization may have existed in several writers in early 19th-century Canadian letters. Scan 1 depicts a piece of LModE writing serving as a source for *Conte*:
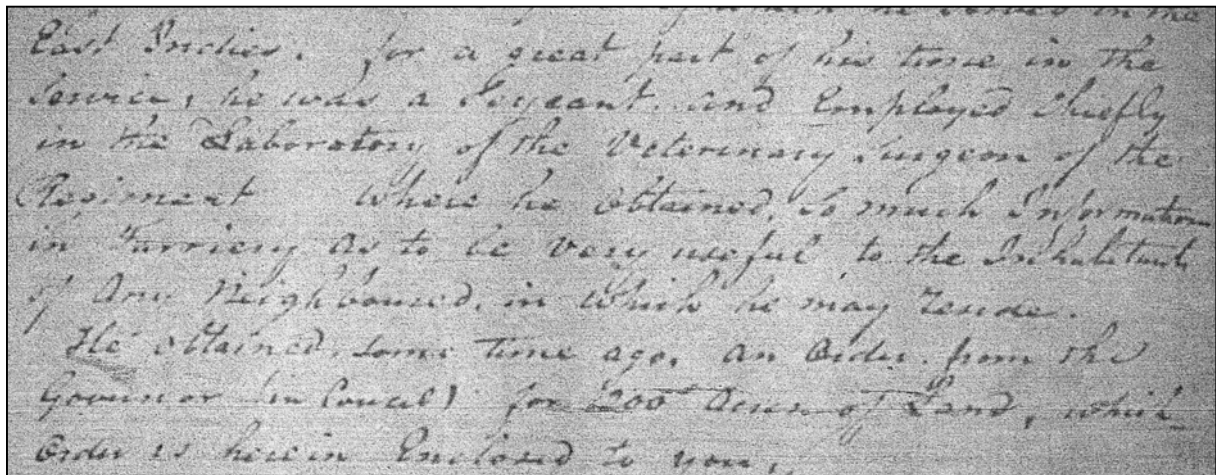
Scan 1: Detail of letter 13985, 3 December 1836, Richard Lowe to John Macaulay

A transcription from *Conte* is provided below. Line breaks have been inserted according to the original, and line numbers have been added to facilitate comparison. The line in [] is not depicted in scan 1 but is added here to put the first word in context:

[15 years of which he served in the]
1 East Indies. for a great part of his time in the
2 Service, he was a Serjeant and Employed chiefly
3 in the Laboratory of the Veterinary Surgeon of the
4 Regiment, where he obtained so much Information
5 in Farriery as to be very useful to the Inhabitants
6 of Any Neighbourhood, in which he may reside.
7 He obtained some time ago, an Order from the
8 Governor (in Council) for 200 Acres of Land, which
9 Order is herein Enclosed to you.

(*Conte*, Reel14.doc)

---

10 I referred to these inventories in Dollinger (2003) as "graphemic inventories", while the present definition of terms renders them allophonic.

If we scan the clipping of the original document and compare it with the transcription, we see that the distribution of UC and LC is somewhat surprising. Cases where one would expect a LC by today's standards are especially noteworthy, since it seems that the LModE writer used an UC, e.g. line 2, the verb *Employed* clearly appears to be an UC letter. On the other hand, in line 1, following a full stop, *for* is definitely a LC character. Could it be that the writer, Richard Lowe, systematically preferred LC [f] and UC [E] in these positions? Or did he possibly use a kind of 'neuter' form of the letter to be used *in lieu of* UC and LC allographs for <Ff> and <Ee>?

To come closer to answering these questions, the available writing by Richard Lowe was checked, which is, unfortunately, only this letter. At the same time, however, this is paradigmatic of LModE writing, especially of the less-privileged classes, since what we have is often not more than a few lines. The complete letter follows:

<13985>
<Richard Lowe to John Macauley, Surveyor General, Toronto>
<middle class>
Adolphustown, Decembr. 3d, 1836
Dear Sir,
    An old Soldier by the name of Stephen Lewis, a native of
Dublin, who had lived in my neighbourhood for about 2 years. It
appears by his discharge &c. that he had been 25 years in his
majesty's Service. 15 years of which he served in the East Indies.
for a great part his time in the Service, he was a Serjeant and
Employed chiefly in the Laboratory of the Veterinary Surgeon of
the Regiment, where he obtained so much Information in
Farriery as to be very useful to the Inhabitants of Any
Neighbourhood, in which he may reside. He obtained some time
ago, an Order from the Governor (in Council) for 200 Acres of
Land, which Order is herein Enclosed to you.
    The object of my writing for him is to acq-aint <le: acqaint,
nw: acquaint> you, that he is too old and unskilful as a farmer to
go far into the Bush. as he is a very peaceable and Steady man,
and one that can make a comfortable livelihood by his skill in
Farr-iery <le: Farriery>. You would do him very Essential
Service by locating him in or near, a thickly settled Country,
where he might have the Benefit of his land and get
Employment in his Calling and be useful to is neighbours.

> I am very respectfully
> Your very Obedt. Servant
> [sign Richard Lowe]

We are fortunate to have enough material to answer the questions pertaining to <Ee> and <Ff>. For <Ee>, as a common vowel, this is less surprising than for <Ff>, which we have in initial, word-medial, as well as word-final positions in our text. For <Ff>, we get the following distribution in table 1:

Table 1. Distribution of <Ff>

position of <Ff>	absolute count
word-internal	7
word-final	3
word-initial	8

These 10, i.e. 7 plus 3, word-internal and word-final <Ff>s are all lower case and represent no problem for interpretation. Looking at the 8 word-initial <Ff>s, we see that Lowe only capitalizes the word *Farriery* 'veterinary medicine'. As some studies have suggested (e.g. Osselton 1985), Lowe might have tried to emphasize this trade by way of capitalization in order to denote its importance. Since it is clear that Lowe also used lower-case initial <Ff>, which can be seen in line 7 in our clipping, emphasis seems to have been a likely reason for his choice, and LC [f] in line 1 could be considered a 'writo'.

The case of <Ee> seems to be less clear cut. Lower case [e] exists, but only word-medially or word-finally, e.g. line 1: *great*, line 2: *he*. Word-initially we have five allographs for <Ee> in total that all appear as UC letters, as can be seen in line 1, *Employed*, and line 9, *Enclosed*. If we interpret this from our PDE background, we see the following realization of the grapheme <Ee> in Lowe:

Table 2.1. PDE-based interpretation

position of <Ee>	PDE interpretation
word-medial	lower-case [e]
word-final	lower case [e]
word-initial	upper case [E]

If we take this inventory a bit further and check back with scan 1, we have 3 instances of <Ee> on the scan: *East Indies* (line 1), *Employed* (line 3) and *Enclosed* (last line). There does not seem to be a difference in [E], regardless

whether it appears in a noun or a verb or an adjective. We therefore could reinterpret the data from table 2.1 in the following way:

Table 2.2. LModE-based interpretation

position of <Ee>	LModE interpretation
word-medial	lower-case [e]
word-final	lower case [e]
word-initial	neutral [E]

The system depicted in table 2.2 would mean that there seems to have been a LC allograph for grapheme <Ee> in Lowe's allographic inventory. Unless we argue that Lowe used upper case allographs for verbs and nouns indiscriminately, we could interpret his initial [E] as 'neutral', i.e. a third class of allograph.

This reasoning is supported by contemporary evidence of writers who did discriminate between UC and LC word-initially. Scan 2 shows a letter by a certain Richard Callahan, also dated 1836, a transcription of which is provided below:
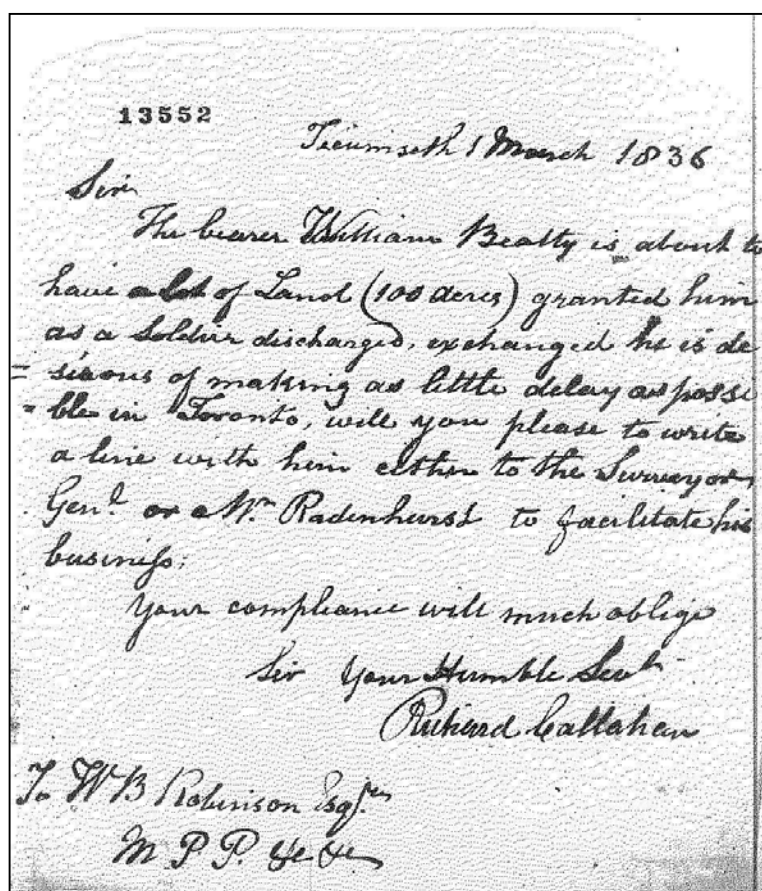
Scan 2: Richard Callahan, 1 March 1836

<13552>
<Richard Callahan to W. B. Robinson. Asks his assistance in expediting the exchange of William Beatty's lot>
Tecumseth 1 March 1836
Sir
The bearer William Beatty is about to have a lot of Land (100 acres) granted him as a Soldier discharged, exchanged he is de-sirous <le: desirous> of making as little delay as possi-ble <le: possible> in Toronto, will you please to write a line with him either to the Surveyor Genl. or Mr. Radenhurst to facilitate his business.
Your compliance will much oblige
Sir Your Humble Servt
[sign Richard Callahan]

To W B Robinson Esq.
M. P. P. &c &c

In Callahan's letter we find both [e] as well as [E] word-initially, as evidenced by the following spellings: *exchanged*, *either*, and in the penultimate line: 'To W B Robinson *Esq.*'. The writer seems to have had an UC allograph and a LC allograph for grapheme <Ee>.

Table 3. Callahan's allographs for <Ee>

position of <Ee>	LModE interpretation
word-medially	lower-case [e]
word-finally	lower case [e]
word-initially	upper case [E]

In this respect, we might say that Callahan is closer to the PDE system than Lowe, who realized grapheme <Ee> with a lower case allograph [e], word-medially and word-finally, and a neutral allograph [E] word-initially. Consequently, capitalization cannot have served as a means for emphasis here. It is also interesting to note that Callahan obviously corrected LC [w] for UC [W] in "William" in the first line of the main text, which indicates that LC is the unmarked variant in Callahan's hand.

The question remains whether Richard Lowe's allographic inventory for <Ee> is idiosyncratic or evidence for more general practice. While this question would best be answered by comparison with many other contemporary letters, whichcannot be carried out here, we may find an indicator in the letter tables set up by calligraphers. Scan 3 depicts a typographical plate that was reprinted in George Bickham's writing guide

Universal Penman, published between 1733 and 1741 (Heal 1962: 171f). Plates like these were promoted as a model at the beginning of the 19th century in North-America (Nash 1969: 9f):
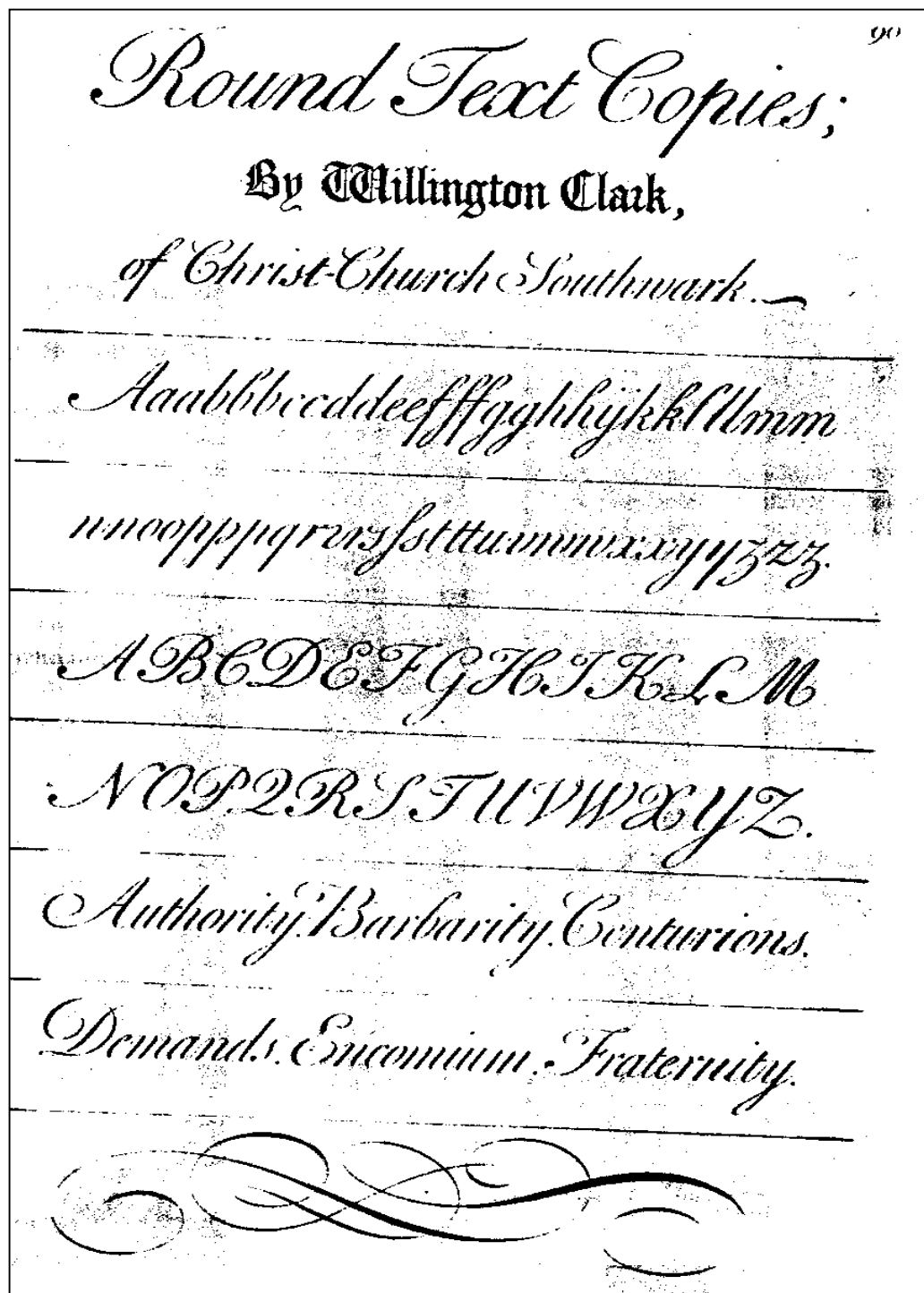
Scan 3: W. Clark's Plate no. 90, reprinted in Bickham's *Universal Penman*, as reproduced in Heal (1962: facing p. 184)

In scan 3, we see two forms of lower-case [e] and one for upper-case [E] that closely resembles our proposed neuter form. We may therefore assume that although the concept of binary realizations of letters already existed and was proposed by experts of handwriting at the beginning of the period, not all writers used it more than half a century after its presentation.

The example of <Ee> shows us that before we are able to safely transcribe upper and lower case, we would need to establish the allographic inventories for each writer and discriminate between UC, LC and neuter case. Having done that we also need to ask how we should technically incorporate this information into machine-readable historical corpora. While it is possible to annotate every word in a corpus in a way to mark-up a three-tier system, upper, lower and neutral case, this would again further add on to the task list of the corpus compilers.

6. Conclusion

This article has shown that it may be beneficial to re-adjust our focus towards more philological precision by assigning it a more central role in corpus linguistics of the LModE period. Now that the technological advances have given us more powerful computers than at the time when the *Helsinki Corpus*, or even the *ARCHER Corpus* were conceived, this seems to be a good direction to go. The message is not new, but worth being re-iterated: what we don't code in the corpus needs to be clearly indicated in the accompanying documentation. Without such manuals an electronic corpus is, if not useless, severely limited in its applicability.

The examples have shown that it may be necessary to apply Meurman-Solin's plea for philological computing to handwritten texts, which can only be reached by stopping the outsourcing of philological expertise and trying to unite two distinct domains, focussing more on the level of paleography and its sister disciplines than it used to be the case. In so doing, we may open a Pandora's Box, as apparently simple issues like capitalization will add heavily on to the task list of the corpus compiler.

If we were to follow this way, however, where would this lead to? For one thing, corpus size would rather remain constant or even decrease in comparison to today's standards. The distinction drawn in Kennedy (1998) between first-generation, e.g. Brown, London-Lund, LOB and second-generation corpora, such as the BNC and ICE, is one largely based on size and would as such not be applicable to historical corpus linguistics, as the

merits of 'philologically-computerized' corpora would lie elsewhere.[11] However, I think that even the more traditional, and therefore apparently 'safer' areas of study within corpus linguistics, morphology and syntax, might benefit from more carefully computerized data, as the example of possessive –*s* indicates.

Despite the problems involved in corpus compilation from handwritten documents, I hope to have shown that apparently unimportant questions of transcription do have a bearing on linguistic questions. I am certain that this focus would bring our data and thus our research more closely to the first-order phenomena one would ideally want to investigate.

References

Austin, Frances (ed.) 1991. *The Clift family correspondence 1792-1846.* Sheffield: University of Sheffield, Centre for English Cultural Tradition and Language.

Austin, Frances. 2000. "Letter writing in a Cornish community in the 1790s" – in: Barton, David and Nigel Hall (eds.) *Letter writing as a social practice.* Amsterdam: Benjamins (= Studies in Written Language and Literacy, 9), 43-61.

Besterman, Theodore (ed.) 1971. *A bibliography of bibliographies.* Totowa: Rowman and Littlefield.

Biber, Douglas, Edward Finegan and Dwight Atkinson. 1994. "ARCHER and its challenges: compiling and exploring a representative corpus of historical English registers" – in: Udo Fries, Gunnel Tottie and Peter Schneider (eds.) *Creating and using English language corpora.* Amsterdam: Rodopi (=Language and Computers, 13), 1-13.

Biber, Douglas and Edward Finegan. 1995. "The ARCHER Corpus". *ICAME Journal.* 19: 148.

Biber, Douglas and Edward Finegan. 1997. "Diachronic relations among speech-based and written registers in English" – in: Nevalainen, Terttu and Leena Kahlas-Tarkka (eds.) *To explain the present from the past. Studies in the changing English language in honour of Matti Rissanen.* Helsinki: Société Néophilologique (= Mémoires de la Société Néophilologique de Helsinki, 52), 253-275.

Biber, Douglas, Susan Conrad and Randi Reppen. 1998. *Corpus linguistics. Investigating language structure and use.* Cambridge: CUP.

Biber, Douglas and Jená Burges. 2000. "Historical change in the language use of women and men". *Journal of English Linguistics.* 28/1: 21-37.

Claridge, Claudia. 2001. *'Life is ruled and governed by opinion.' The Lampeter Corpus of Early Modern English Tracts. Manual of information.* Online: http://www.tu-chemnitz.de/phil/english/chairs/linguist/real/independent/lampeter/manual/pages/manual.html, accessed 19 June 2004.

---

[11] It is self-evident that corpus size remains a crucial factor in corpus linguistics.

Croft, P. J. (ed.) 1973. *Autograph poetry in the English language. Facsimiles of original manuscripts from the fourteenth to the twentieth century.* With introduction, commentary and transcripts. 2 vols. London: Cassell.

Curzan, Anne and Taryn Hakala. 2004. "XXIV International ICAME Conference. 23-27 April, 2003, Guernsey". *ICAME Journal.* 28: 144-151, online: http://helmer.hit.uib.no/icame/ij28/index.html, accessed 18 June 2004.

Dollinger, Stefan. 2003. "What the capitalization of nouns in Early Canadian English may tell us about 'colonial lag' theory: methods and problems" *VIEWS.* 12/1: 24-44, also online at: http://www.univie.ac.at/Anglistik/views/03_1/DOL_SGLE.PDF, 22 June 2004.

Dollinger, Stefan. forthc. "Oh Canada! Towards the *Corpus of Early Ontario English*" – in: Renouf, Antoinette *et al.* (eds.) *Proceedings of the 24th Conference of the International Computer Archive of Modern and Medieval English (ICAME 24). Guernsey, April 2003.* Amsterdam: Rodopi [approximate publishing information].

Dossena, Marina. 2003. "Corpus of 19th-Century Scottish Correspondence" *ICAME Journal.* 27: 91, online: http://helmer.hit.uib.no/icame/ij27/index.html, accessed 23 June 2004.

Erdman, David V. (ed.) 1965. *The poetry and prose of William Blake.* Commentary by Harold Bloom. Garden City, New York: Doubleday.

Fairman, Tony. 2000. "English pauper letters 1800-34, and the English language" – in: Barton, David and Nigel Hall (eds.) *Letter writing as a social practice.* Amsterdam: Benjamins (= Studies in Written Language and Literacy, 9), 63-82.

Fairman, Tony. 2003. "Letters of the English labouring classes and the English language, 1800-1834" – in: Dossena, Marina and Charles Jones (eds.) *Insights into Late Modern English.* Bern: Lang (= Linguistic Insights, 7), 265-282.

Görlach, Manfred. 2001. *Eighteenth-century English.* Heidelberg: Winter.

Greetham, D. C. 1994. *Textual scholarship. An introduction.* New York: Garland.

Heal, Ambrose. 1962. *The English writing-masters and their copy-books 1570-1800. A biographical dictionary and a bibliography. With an introduction on the development of handwriting by Stanley Morison. Illustrated with portraits of the masters and specimens of their hands.* Unchanged reprint of ed. of 1931. Hildesheim: Olms.

Hector, L. C. 1958. *The handwriting of English documents.* London: Arnold.

Harris, Robin S. and Terry G. Harris (eds.) 1994. *The Eldon House Diaries. Five women's views of the 19th century.* Toronto: Champlain Society and Government of Ontario.

Kennedy, Graeme. 1998. *An introduction to corpus linguistics.* London: Longman (= Studies in Language and Linguistics).

Keränen, Jukka. 1998a. "The Corpus of Early English Correspondence: progress report" – in: Antoinette Renouf (ed.) *Explorations in corpus linguistics.* Amsterdam: Rodopi (= Language and Computers. Studies in Practical Linguistics, 23), 29-37.

Keränen, Jukka. 1998b. "Forgeries and one-eyed bulls. Editorial questions in corpus work" *Neuphilologische Mitteilungen.* 99: 217-226.

Keynes, Geoffrey Kt. (ed.) 1980. *The letters of William Blake. With related documents.* 3rd ed. Oxford: Clarendon.

Keynes, Geoffrey Kt. (ed.) 1968. *The letters of William Blake.* 2nd ed. Revised and amplified. London: Hart-Davis.

Keynes, Geoffrey Kt. (ed.) 1956. *The letters of William Blake.* [1st ed.] London: Hart-Davis.

Keynes, Geoffrey Kt. 1926. *Letters from William Blake to Thomas Butts 1800-1803. Printed in facsimile*. Oxford: OUP.

Kytö, Merja. 1991 [1988]. *Variation and diachrony, with Early American English in focus. Studies on* can/may *and* shall/will. Frankfurt/Main: Lang.

Kytö, Merja (ed.) 1991. *Manual to the diachronic part of* The Helsinki Corpus of English Texts. *Coding conventions and lists of source texts.* Helsinki: Department of English, University of Helsinki.

Kytö, Merja, Juhani Rudanko and Erik Smitterberg. 2000. "Building a bridge between the present and the past: a corpus of 19th-century English". *ICAME Journal.* 24: 85-97.

Laitinen, Mikko. 2002. "Extending the Corpus of Early English Correspondence to the 18th century" *Helsinki English Studies.* 2: [no page numbers], online: http://www.eng.helsinki.fi/hes/Corpora/extending_the_corpus2.htm, accessed 19 June 2004.

Laitinen, Mikko. 2003. "Compiling an electronic corpus of the early 19th century pauper letters – linguistic evidence and methodological considerations" – in: Davis, Ceri (ed.) *ICAME 2003. The changing face of corpus linguistics* [conference abstracts]. University of Liverpool, R & D Unit for English Studies, 70.

Markus, Manfred. 1997. "Nominalization of Middle English prose: possibilities and limits" – in: Ljung, Magnus (ed.) *Corpus-based studies in English. Papers from the Seventeenth International Conference of English Language Research on Computerized Corpora (ICAME 17). Stockholm, May 15-19, 1996.* Amsterdam: Rodopi (= Language and Computers: Studies in Practical Linguistics, 20), 211-226.

Meurman-Solin, Anneli. 2001. "Structured text corpora in the study of language variation and change." *Literary and Linguistic Computing.* 16/1: 5-27.

Meyer, Charles F. 2002. *English corpus linguistics*. Cambridge: CUP.

Nash, Ray. 1969. *American penmanship 1800-1850. A history of writing and a bibliography of copybooks from Jenkins to Spencer.* Worchester: American Antiquarian Society.

Nevala, Minna. 2001. "With out any Pregyduce or Hindranc. Editing women's letters from 17th-century Norfolk" *Neuphilologische Mitteilungen.* 102: 151-171.

Nevalainen, Terttu and Helena Raumolin-Brunberg. 2003. *Historical sociolinguistics. Language change in Tudor and Stuart England.* Harlow: Pearson Education. (= Longman Linguistics Library).

Osselton, Noel E. 1984. "Informal spelling systems in Early Modern English: 1500-1800" – in: Blake Norman F. and Charles Jones (eds.) *English historical linguistics: studies in development.* Sheffield: Centre for English Cultural Tradition and Language, 123-137.

Osselton, Noel E. 1985. "Spelling-book rules and the capitalization of nouns in the seventeenth and eighteenth centuries." – in: Arn, Mary-Jo and Hanneke Wirtjes with Hans Jansen (eds.) *Historical & editorial studies in Medieval & Early Modern English for Johan Gerritsen.* Groningen: Wolters-Noordhoff, 49-61.

Pintzuk, Susan. 2004. "Linguist List 15.1394. Tue May 4 2004. Qs: Electronic versions of Latin texts" online: http://www.linguistlist.org/issues/15/15-1394.html, accessed 19 June 2004.

Raumolin-Brunberg, Helena (ed.) 2002. *Variation past and present. VARIENG studies on English for Terttu Nevalainen.* Helsinki: Société Néophilologique (= Mémories de la Société Néophilologique de Helsinki, 61).

Rissanen, Matti, Ossi Ihalainen, Terttu Nevalainen and Irma Taavitsainen (eds.) 1992. *History of Englishes. New methods and interpretations in historical linguistics.* Berlin: Mouton de Gruyter (= Topics in English Linguistics, 10).

Shillingsburg, Peter L. 1996. *Scholarly editing in the computer age. Theory and practice.* 3rd ed. Ann Arbor: University of Michigan Press (= Editorial Theory and Literary Criticism).

Stabile, Juliana M. 2002. *Toronto newspapers, 1798-1845. A case study in print.* Unpublished PhD thesis: University of Toronto.

Tieken-Boon van Ostade, Ingrid. forthc. "Eighteenth-century English letters: in search of the vernacular" [unpublished manuscript].

Tieken-Boon van Ostade, Ingrid. 1998. "Standardization of English spelling: the eighteenth-century printer's contribution" – in: Fisiak, Jacek and Marcin Krygier (eds.) *Advances in English historical linguistics (1996).* Berlin: Mouton de Gruyter (= Trends in Linguistics. Studies and Monographs, 112), 457-470.

Walsh, Marcus. 1997. *Milton, Shakespeare and eighteenth-century literary editing. The beginnings of interpretative scholarship.* Cambridge: CUP (= Cambridge Studies in Eighteenth-Century Literature and Thought, 35).

Negative prefixes in Middle English: a corpus-based study of dis-, in-, mis- & un-

Ursula Lutzky, Vienna

1. Introduction

When talking about prefixation in ME, one is referring to a situation when prefixation was no longer a very prominent means of word-formation, as many of the OE prefixes were no longer productive and the new Romance prefixes were not yet fully established. ME was thus a period of transition characterised, on the one hand, by the loss of many native prefixes[1] but, on the other hand, also by the influx of vast numbers of new lexemes from French, which subsequently gave rise to foreign prefixes entering the derivational system of the English language. The extensive borrowing during the ME period did, therefore, not only result in an enrichment of the English vocabulary but also changed the derivational morphology of the language (Cowie 2002: 410f., Baugh 1993: 177f., Burnley 1992: 445ff.; cf. also Kastovsky 1992: 377).

Even though the ME period was thus of great importance for the further development of English word-formation, research in the area of ME prefixation has so far been rather scarce. While Dalton-Puffer's (1996) corpus-based study of derivation has greatly enhanced our understanding of ME suffixation, prefixes have, apart from Burnley's (1992: 446f.) introduction into ME prefixation, not received a great deal of attention in the past. Also the semantically more restricted field of negative prefixation in ME has not been analysed in more detail so far. While Roderick McConchie's (2000) recent article on negative *dis-* forms part of the same semantic domain, McConchie concentrates on a later time span, namely on the beginning of the EModE period (1520-1620). Consequently, one can say that the time was ripe for a more detailed study of prefixation in ME.

---

[1] Many native prefixes had already disappeared towards the end of the OE period and only few of them, compared to the abundance of prefixes that existed in OE, survived into ME. (Burnley 1992: 445ff., Kastovsky 1992: 377)

For the present analysis the four prefixes *dis-, in-, mis-* and *un-* were chosen and studied with the aim of illustrating and clarifying the following points:

- How the native prefix *un-*, which had already existed in the OE period, survived into ME and met the introduction of several foreign prefixes belonging to the same semantic group.
- How the foreign prefixes *dis-* and *in-* entered the English language and slowly established themselves after the Norman Conquest.
- To what extent the productivity of the native prefix *mis-* was boosted by its fusion with French *mes-*.
- How far these four prefixes can be said to have been productive in ME.

The purpose of this article is thus to discuss the insights gained in a diachronic study of four ME prefixes. After a short introduction into the field of productivity and a description of the data, the results of the corpus-based analysis of *dis-, in-, mis-* and *un-* will be presented and interpreted.

2. Productivity

The concept of productivity, according to Aronoff (1976: 35) "one of the central mysteries of derivational morphology", has not only given rise to very diverse approaches to the topic but has always been a primary concern in the field of word-formation. The question of a word-formation process being productive or not has often been asked and attempts have been made to reach satisfying answers – the main problem being the rather fuzzy concept of productivity itself.

Basically, it is not very easy to give THE definition of productivity, especially as attempts at a definition tend to start out from various theoretical backgrounds. An example of a recent definition of productivity is provided by Bauer.

> In sum, the productivity of a morphological process is its potential for repetitive non-creative morphological coining. (Bauer 2001: 98)

Bauer's definition, though compact, remains unsatisfactory as it contains concepts that are considered to be controversial. For instance, how is the notion of potential to be accounted for? Or what is meant exactly by repetitive? Apart from that, Bauer (2001: 62ff.) himself notes that the notion of creativity cannot be clearly differentiated from productivity and, finally, also the question of the practical applicability of Bauer's definition arises. Due to these and similar controversies and due to the fact that productivity will be approached diachronically in the present study, productivity will

simply be regarded as the interplay of various conditioning factors in the following.

While morphological productivity is already a controversial matter in synchronic terms, viewing the concept from a historical perspective is even more challenging. The main problem with studying the productivity of word-formation patterns in past periods of the English language is that it can, of course, not be tested directly. Access to spoken evidence is denied and, therefore, productivity tests, coinage tests and judgement tests (Bolozky 1999: 3f.) or Aronoff's experiments based on the *Lexical Decision Task* (Aronoff 1980, Anshen & Aronoff 1981) prove unfeasible: neither can we ask native speakers of ME to form potential words that sound acceptable to them, nor can we let them judge the acceptability of a list of words, in order to discover the potential of a particular word-formation rule.

What is available for historical studies is circumstantial evidence, as, for instance, the number of new formations occurring in a text of a certain period or continued productivity in subsequent periods. Apart from that, one can search for examples of innovative use like, for instance, hybrid formations, i.e. the combination of foreign with native derivational material. Another important clue in diachronic studies is the transparency of a pattern, i.e. the formal-semantic analysability of lexemes. (Kastovsky 1992: 357f., Zbierska-Sawala 1993: 6)

Furthermore, it is, of course, possible to approach productivity in historical studies quantitatively, especially when working with corpora. Thus quantitative measures have also been included in the following analysis (cf. type and token frequencies). However, productivity measures as suggested by Baayen (Baayen & Lieber 1991: 809ff., Baayen 1992: 110ff., Baayen 1993: 190ff.) were not made use of for two reasons: on the one hand, Baayen's productivity measures (productivity in the narrow sense, global productivity, the hapax-conditioned degree of productivity) show certain deficiencies, as also pointed out by Baayen himself (cf. Baayen 1991, 1992, 1993, Plag 1999: 25ff., Van Marle 1992). On the other hand, the data and the corpus from which it was extracted were not considered to be large enough to yield reliable results.

Basically, the present study did not rely on one single previous approach to productivity, may it be quantitative or qualitative. Instead it started out from the idea that when considering productivity in historical studies, one mainly has to rely on certain conditioning factors and productivity might be regarded as the interplay of these factors. As already pointed out by Dalton-Puffer (1996: 216), what remains to be settled is in how far these factors

interact and are interdependent. In the following analysis three main points were investigated, namely type and token frequencies, hybrid formations and the occurrence of derivatives in particular text types. This final point has so far been neglected in productivity studies though, as will be seen later on, it can yield interesting results as far as the use and distribution of affixes is concerned.

3. The data

The prefixes *dis-*, *in-*, and *un-* may all express negativity and in this study they were only examined in this function, while additional meanings (e.g. reversativity as in *untie*, *disconnect*, privativity as in *unmask*, *disarm…*) they might carry were not taken into consideration. As mentioned above, the focus was on discovering how far the foreign prefixes *dis-* and *in-* established themselves during the ME period but also to what extent they coexisted (or even competed) with native *un-*. [2]

Furthermore, one prefix which in OE and ME also found use as a negative prefix but which is today usually classed as an evaluative or pejorative prefix, namely *mis-*, was included in the analysis. While *mis-* is originally a native prefix, in ME it witnessed the influx of many French loans carrying the prefix *mes-* with which it merged both formally and semantically. This shows how great the impact of the French language on English must have been, even leading to the fusion of morphological material.

The analysis of the four chosen prefixes was carried out with the aid of a computerized text corpus - *The Helsinki Corpus of English Texts.*[3] While the diachronic part of the HC contains very different types of text, which date from c 750 to c 1700, this study focused on the ME period, which is further subdivided in the corpus into the four sub-periods M1-M4. An important fact that has to be taken into consideration is the size of these sub-periods, i.e. the amount of words that can be identified for each of them.

---

[2] The negative prefix *non-* was not treated in this study as loans with *non-* only started to enter the English language towards the end of the ME period. The prefix *de-* was not included, as the data compilation showed that the number of types with negative *de-* was almost negligible compared to the number of attestations with privative and reversative *de-* in the corpus.

[3] *The Helsinki Corpus of English Texts* will from now on be referred to as HC.

Sub-period	Words	%
M1 (1150-1250)	113 010	18.6
M2 (1250-1350)	97 480	16.0
M3 (1350-1420)	184 230	30.3
M4 (1420-1500)	213 850	35.1
Total	608 570	100.0

Fig. 1. Helsinki Corpus Overview (Kytö 1991: 2)

A look at Fig. 1 shows that M1 and M2 as well as M3 and M4 do not differ greatly in size. However, there is a considerable difference in size between the first two sub-periods and the final two sub-periods. Taken together, M3 and M4 roughly make up two thirds of the data available for the ME period. In the analysis, this fact had to be borne in mind, especially when comparisons between the different sub-periods were made.[4]

After all relevant complex word forms bearing one of the prefixes *dis-, in-, mis-* or *un-* had been extracted from the HC, a database comprising ten core categories as well as a changing number of reference columns (*Ref1 – Refx*), depending on the respective frequencies with which words occurred in different text samples, was set up and filled. [5]

Word	disobeyeth	inpossible	meschief	uncu+de
Frequency	2	2	1	2
Lemma	disobeien	impossible	mischef	uncouth
Category	V	adj.	N	adj.
Base	OF	L/OF	OF	OE
Cat_Base	V	adj.	N	adj.
Base 1st	(c1390)	(a1382)/ a1400(a1325)	c1330(?a1300)/ a1400(c1303)	1121/ a1225(c1200)
Lemma 1st	a1325/ (c1375)	(c1384)/ ?a1425(c1380)	c1300/ c1325(c1300)	a1131/ c1225(?c1200)
Prefix	*dis-*	*in-*	*mis-*	*un-*
Period	M4	M3	M2	M1

Fig. 2. Structure of the database

---

[4] Although synchronic corpora are normally weighted, the general practice has been that studies based on the HC refrained from normalizing. This is why also in this analysis frequencies were not weighted to sub-period size.

[5] The structure of the database was taken over from Dalton-Puffer who used the same structure as the basis of her analysis of ME suffixes (Dalton-Puffer 1996). The tables of the database were filled with the help of the *Middle English Dictionary* and the *Oxford English Dictionary*.

The ten core categories were adopted in view of the later analysis, which was to comprise type and token frequencies (Word, Frequency, Lemma), the syntactic categories of the base and the derivative (Category, Cat_Base) and hybrid formations (Base, Base 1st, Lemma 1st). [6]

4. Analysis

4.1 Type and token frequencies

Comparing type and token frequency, token frequency has to be handled carefully as far as productivity is concerned: it does not always yield very reliable results, as small numbers of very frequent types may blur the picture. Type frequency, in contrast, counts among the major conditioning factors of productivity, i.e. the more frequently a word-formation pattern is used in the production of types, the more productive it usually is. Also the number of types that occur only once in the corpus, i.e. the so-called hapax legomena, are regarded as an indicator of productivity as they may reveal how often a prefix is used in the formation of new lexemes, i.e. of neologisms. (Cowie 2002: 426, Dalton-Puffer 1996: 217f., Plag 1999: 215f.)

If one now tries to relate these assumptions to corpus-based analysis, one can say that if a rather large number of types with low token frequencies can be identified in a corpus, this can be taken as a sign of the productivity of the affix. Furthermore, a productive word-formation pattern is also characterised by a relatively high proportion of hapax legomena.[7] On the contrary, the occurrence of a considerable number of tokens of very frequent types moves the affix closer to the 'unproductive' end on the productivity scale. (Baayen & Renouf 1996: 74, Cowie 2002: 426, Dalton-Puffer 1996: 217f.)

One should, however, also be careful not to equate frequency with productivity for two reasons. On the one hand, not every productive morphological process gives rise to many new lexemes. On the other hand, there are morphological processes which are held to be unproductive and still show a high type-frequency. Thus a particular affix may be encountered in a large number of types at a given point in time without being used very

---

[6] The discussion of syntactic categories, though forming part of the original study, will not be included in this article. For a detailed analysis see Lutzky 2004.

[7] As Baayen et al. (1999: 215f.) mention, it is, of course, important to define hapax legomena with respect to the corpus and, consequently, not every hapax legomenon is automatically a neologism. Especially when working with smaller corpora it is highly probable that hapaxes are not new formations.

frequently to coin new lexemes. Consequently, the frequent occurrence of an affix in existing lexemes may be a sign of past and not of current productivity. (Bauer 2001: 48f., Plag 1999: 22f.)

4.1.1 *Dis-*

The negative prefix *dis-*, ultimately of Latin origin, entered the English language through direct loans from Latin but especially also through French loans with the prefix *des-*. It is represented in a total number of 25 types (77 tokens) in the ME section of the HC.
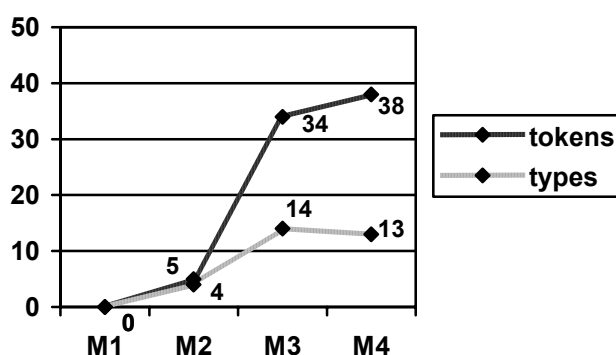

Fig. 3. Type-token figures for *dis-*

As Fig. 3 shows, the negative prefix *dis-* only emerges for the first time in M2 in the HC. While in M1 negative *dis-* is not represented at all, there was no flood of types pouring into English during M2 either: only four types (*discomfortable*, *discordaunce*, *disese*, *dishonour*) can be identified in this sub-period. From M2 to M3, a clear rise can be observed for both tokens and types and from then on the figures stay relatively stable.[8]

While in M2 type and token frequencies are still almost identical, subsequently the increase of tokens is greater than that of types. This means that in M3 and M4 we come across lexemes that occur more often than only once in the respective sub-periods, i.e. which are no hapax legomena. While

---

[8] When considering the sharp increase from M2 to M3 for both tokens and types, it should be borne in mind that the total number of words with which M3 is represented in the corpus is almost twice as large as that of M2 (see section 4) and the steep rise should thus not be over-interpreted.

in M3 about half of the total number of types is represented only once, in M4 it is only about one third. Thus in M4 one can already identify several high(er)-frequency types like, for instance, *displesen* or *disese*.

4.1.2 *In-*

The Latin prefix *in-*, which entered the English language through loans from French and Latin, is attested in 49 different types (144 tokens), i.e. its overall frequencies are considerably higher than those of *dis-*.
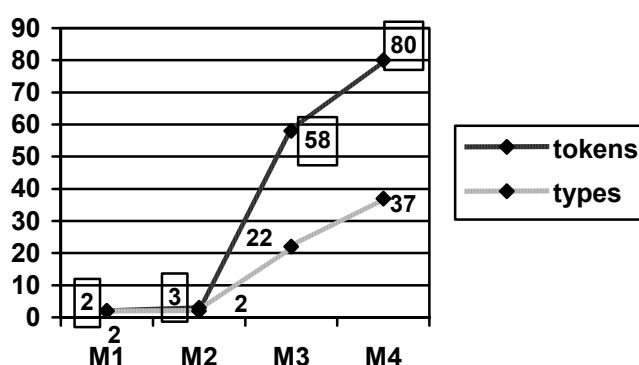

Fig. 4. Type-token figures for *in-*

As can be deduced from Fig. 4, *in-* is not very prominently represented before 1350. As opposed to *dis-*, derivatives with *in-* (*innocent* n., *inobedience*) already appear in the first sub-period of ME. However, the fact that *in-* is only represented by four different types (*innocent* n., *innocent* adj., *inobedience*, *indignacioun*) in the first two sub-periods shows that the prefix did not yet occupy a significant position in the English language.

From M2 to M3, a very sharp increase of types and tokens can be observed and this increase also continues from M3 to M4. In M4 we thus come across 26 types carrying the prefix *in-* that did not appear in the corpus before. This shows that more and more new combinations with *in-* were represented in the English language, which implies that the prefix must have become more familiar and slowly established itself.

What can be stated for both Romance prefixes is that there was a true explosion of numbers in M3 (about half of all types of both prefixes occur in this sub-period) and that M4 is the sub-period when the majority of words with *dis-* and *in-* can be identified.

4.1.3 *Mis-*

The native prefix *mis-*, which met the influx of the Old French prefix *mes-* in Middle English times, is attested in 33 types (134 tokens) in the Middle Engish period.
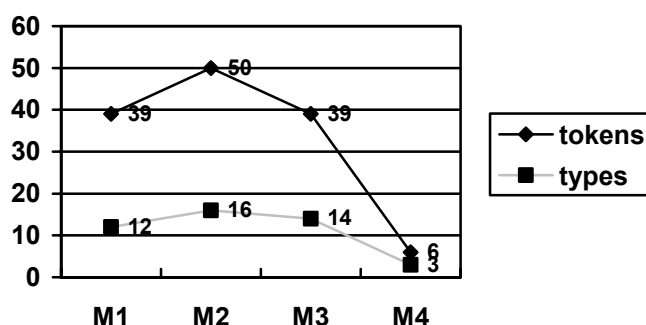

Fig. 5. Type-token figures of *mis-*

Fig. 5 shows that the prefix *mis-* increased both in terms of types and tokens from M1 to M2. As the prefix *mis-* basically stagnated in combination with native bases from M1 to M2, this noticeable increase can in fact be attributed to French influence. Thus one could say that the adoption of loans with the French prefix *mes-* had a positive effect on both the type and token frequencies in M2. Subsequently, however, a decline of *mis-* both in terms of types as well as tokens can be observed. Here one can note that it is mainly due to French influence that the drop from M2 to M3 is not very sharp. If one has a look at the data of M3, it becomes obvious that combinations of *mis-* with native bases were on the decline, while derivatives with foreign bases of coining constitute almost two thirds of the total number of types.

What can be clearly seen, however, is that numbers are diminishing from M2 onwards. In M4, the prefix is only very slightly represented and the picture that presents itself in this last sub-period would justify the conclusion that *mis-* was becoming more and more restricted towards the end of the ME period. It would, of course, be interesting to see how *mis-* subsequently developed in EModE in order to find out whether the low type and token numbers of M4 are in line with the further development of the prefix.

Based on these newly gained insights about *mis-* in ME, how should one thus approach the often voiced remark that due to French influence "a great extension of the use of the prefix [*mis-*] took place …" (OED: s.v. *mis-*)? Basically, the general assumption that the foreign prefix *mes-* strongly reinforced native *mis-* is not fully supported by the data. Even though a

certain boost that was definitely due to French influence can be identified in M2, a drop in type and token frequencies can already be observed in M3, i.e. in a period that is almost twice the size of M2. This decline in numbers consequently shows that the boost of M2 was not a lasting one. Finally, the very low numbers of M4 totally contradict Faiß's statement about the prefix *mis-*, namely that "it is not until its merger with OFr. *mes-* that it becomes really productive" (Faiß 1992: 70). In fact the term 'productive' proves rather inappropriate as far as the results of this type and token analysis are concerned. The present analysis thus shows how corpus-based approaches as well as the reappraisal and further development of the notion of productivity that has taken place over the last decades may call former conclusions into question.

4.1.4 *Un-*

Native *un-* constitutes the negative prefix that is, with a total number of 144 types (367 tokens), most strongly represented in the ME part of the corpus. This high number is especially due to the first sub-period of ME, when *un-* was 'the' negative prefix in the English language, hardly meeting any rivalry from other prefixes.
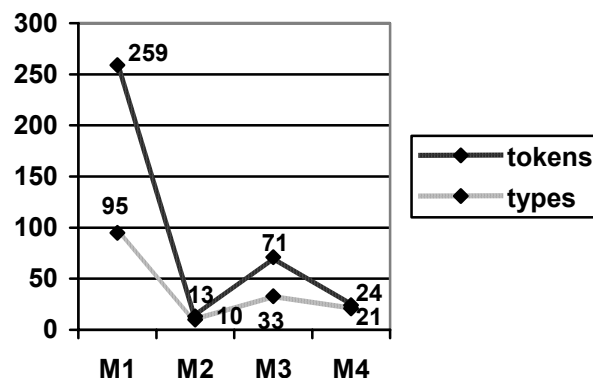
Fig. 6. Type-token figures of *un-*

While *un-* is the negative prefix that by far surpasses the other prefixes treated in this study both in terms of type and token frequencies in M1, we witness such an extreme drop in M2 that the figures make us doubt their correctness – especially as both tokens and types rise again from M2 to M3 and, finally, show a marked decrease from M3 to M4. Basically, this sudden drop in M2 does not seem to represent linguistic reality appropriately. The

sharp decline appears especially questionable as the foreign negative prefixes had not yet established themselves in the English language and one cannot yet speak of the competition of *un-* with other negative prefixes, like *dis-, in-* or *non-*. [9]

Though the type and token frequencies in M2 have to be regarded with suspicion, one may nevertheless assume that the negative prefix *un-* must have been on the decline from M1 onwards – an assumption that is also supported by the type and token numbers of M3 and M4. While many of the OE formations with *un-* had already passed out of use by the second half of the ME period, only a certain number of new formations can be identified. Here it could, of course, be argued that in these last two sub-periods the foreign prefixes *dis-* and *in-* started to enter the English language in larger numbers of loans and that the presence of these new negative prefixes might have influenced the decline of *un-*. This argument, though it probably played a role in the declining vitality of *un-* in ME, is weakened, however, by the fact that no hybrid formations with *dis-* or *in-* can be identified in the ME section of the HC. Contrarily, however, *un-* began to attach to French bases that entered the English language and the first hybrid formations of this kind can already be identified in M3.

Fig. 6 illustrates the changes in type and token frequencies of *un-* from M1 to M4. While in M1 there is an enormous difference between the number of types (95) and tokens (259), in M4 these numbers are almost identical. This shows that while several high frequency types (e.g. *uncouth*, *unhersumnes(se*, *unright* n., *unworthy*) appear in M1, in M4 the percentage of hapax legomena is considerably large (86%).

4.1.5 Summary

Finally, Fig. 7 shows the development of the type frequencies of the prefixes *dis-, in-, mis-,* and *un-* over the four sub-periods of ME. What can be clearly seen is that while the prefix *un-* is very strongly represented in M1, its dominance slowly declines and in M3 and M4 we can witness a more even distribution of the negative prefixes *dis-, in-,* and *un-*.

---

[9]   Apart from that the swift fall in type and token numbers, presumably, cannot be attributed to the size of M2 (see section 4) either, as M1 and M2 are almost identical in size and, nevertheless, huge numbers of derivatives with negative *un-* are attested in M1.

Compared to *un-*, one can say that the prefixes *dis-* and *in-* follow the opposite development. While only very few types with *dis-* and *in-* appear in M1 and M2, the prefixes occupy more prominent positions in M3 and M4. In terms of types, *in-* is even the most strongly represented prefix of all in M4. This shows again very clearly what far-reaching effects the introduction of foreign prefixes into the English language had.
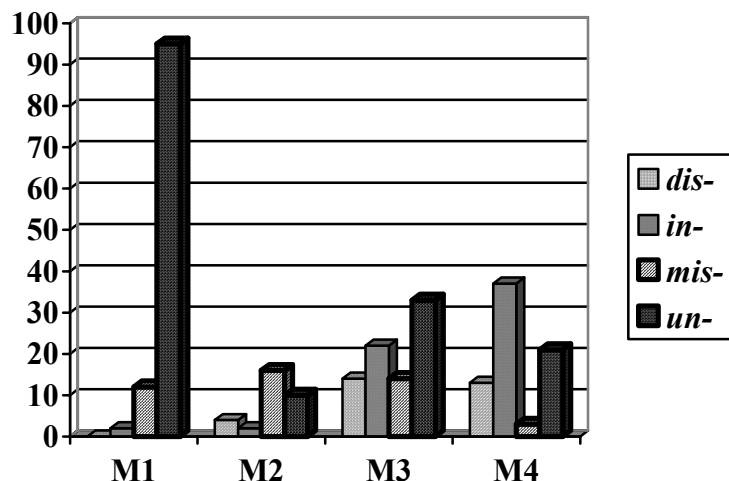
Fig. 7. Overview of the four prefixes in Middle English

4.2 Hybrid formations

Apart from type and token frequencies, a further conditioning factor of productivity was taken into consideration in the analysis of *dis-, in-, mis-* and *un-*, namely hybrid formations. Hybrid formations can have one of two shapes: either a foreign base is combined with a native affix or a native base is combined with a foreign affix. In the analysis of the ME data, both types of hybrid formations were studied.

4.2.1 Foreign prefixes

Basically, the combination of a Romance prefix with a native base is a sign of very close language contact. It usually takes some time until this type of hybrid formation emerges, while the combination of a native prefix with a foreign base can be encountered earlier. This has to do with the fact that loanwords are more quickly and more easily adopted by a language community than foreign prefixes, whose way into the recipient language is

long and complex. (Burnley 1992: 445f., Dalton-Puffer 1996: 211, Gretschmann 1994: 69f., Hansen 1982: 66, Marchand 1969: 129f., 210ff.)

In the ME section of the HC no instances of hybrid formations with *dis-* or *in-* could be identified. In order to find possible explanations for this result, one can study the representation of the negative prefixes *dis-* and *in-* over the four sub-periods of ME in the HC: one can observe that during the first two sub-periods the prefixes are not very strongly represented. While *dis-* is only attested for the first time in M2 in four different derivatives, *in-* can already be encountered in M1, but its type frequency is almost negligible during the first two sub-periods. The two prefixes are more strongly represented for the first time in M3. It was then that larger numbers of lexemes carrying these prefixes were adopted into the English language from French. If one considers that these loans first had to become familiar and grow in number before they could be analysed as complex by the speech community, it is actually not very surprising that we cannot find any hybrid formations with *dis-* or *in-* in the HC.[10]

Finally, intra-romance productivity needs to be taken into consideration. Thus it is of interest whether these foreign prefixes already formed new lexemes with foreign bases during the ME period, i.e. whether complex formations can be identified that are unattested in the donor language OF.

As far as the prefix *dis-* is concerned, only one example of a new formation is attested in M4, namely the formation *disencresynge*, where the prefix *dis-* might have been combined with the base *encresen* in analogy to the formation *decresen,* which in ME also often appeared with the prefix *dis-*, i.e. *discresen* (cf. the variation of the prefixes *des-, de-* and *dis-* in OF). However, this one example does not allow us to speak of intra-romance productivity yet. Furthermore, not a single new formation with the prefix *in-* can be encountered. This means that all derivatives with *in-* that form part of the data were borrowed from French. Consequently, also intra-romance productivity can be ruled out as far as the Romance prefixes *dis-* and *in-* are concerned.

---

[10] Generally, the negative prefix *in-* can be only found in combination with adjectives and nouns of Latin or French origin and even though it was quite vitally involved in the formation of new lexemes in the history of the English language, it did not form hybrids.

4.2.2 Native *un-*

As hybrid formations with native prefixes are said to be more easily formed and can normally be observed earlier, one might assume that hybrid formations with native *un-* may already be attested in the ME part of the HC.

This is in fact the case and the first hybrid formations with French bases already appear in M1. However, compared to the overall number of derivatives with native *un-* in this sub-period, the percentage of types showing a foreign basis of coining is very low (2.8%) and one only comes across sporadic examples like *unofserved*, *unhurt,* or *untiffed*. As the examples, both in terms of types and tokens, are not very numerous, no far-reaching conclusions can be drawn from them. What can be noted is that already before the beginning of the great influx of French loan words into the English language, sporadic examples of hybrid formations with native *un-* can be identified – a fact that speaks in favour of the productivity of the prefix.

Subsequently, a considerable number of hybrid formations is attested in M3 and M4. In M3, 9 out of 33 types have foreign bases of coining and 10 out of 21 types are hybrid-formations in M4. In both sub-periods, *un-* can be found in combination with simplex adjectives (*uncertain*, *unable*) as well as complex adjectives comprising the French suffix *-able* - a derivational pattern which was to become very productive - as in *uncovenable*, *undepartable, unmevable,* or *unresonable*. Furthermore, native *un-* already combined with participles of foreign origin during the second half of the ME period (e.g. *undouted, unoccupied, unpunished,…*).

4.2.3 *Mis-*

As far as the prefix *mis-* is concerned, it is rather difficult, if not impossible, to speak of hybrid formations. Basically, *mis-* already existed in OE but met the influx of French lexemes with the prefix *mes-* during the ME period. The two prefixes merged subsequently both formally and semantically and today, apart from loans from French, one cannot say whether *mis-* represents one or the other in a particular derivative. Consequently, it is rather doubtful whether one can really talk of hybrid formations with the prefix *mis-* especially in the ME period.

Therefore, the focus of this study was not on hybrid formations but on new formations with the prefix *mis-*: the different sub-periods were searched for combinations of *mis-* with foreign bases that had not been borrowed from French but that had been newly formed in English as well as for combinations

of *mis-* with native bases that had not yet been attested in the OE period according to the MED and the OED.

As far as formations with native bases are concerned, the analysis shows that some types already existed in the OE period (e.g. *misdede, misdon, misfare, misliken*). The remaining types which were attested for the first time in the ME period mainly have to be attributed to the time span 1150-1250. Although some of them do not surface in M1 but in later sub-periods in the HC, their dates of first attestation show that most of them already existed at the very beginning of the ME period. For example, the derivatives with native bases that occur in M3 for the first time (*misliking/e, mistrow, mistrowing*) were all formed earlier in the ME period and thus cannot be regarded as new formations of M3. All in all, the analysis of the data demonstrates that although we can attribute a certain number of new types to the first century of the ME period, the formation of new lexemes with native bases was on the decline afterwards.

Taking the formations with foreign bases of coining into consideration, it is important to find out whether complex lexemes consisting of the prefix *mes-/mis-* and Romance bases were originally borrowed from French or newly formed in ME. Consequently, dates of first attestation for each derivative as well as each base were checked in the MED[11]. To gain further certainty about the origin of complex lexemes, two OF dictionaries were consulted, namely Hindley's *Old French- English Dictionary* (2000) and Godefroy's *Dictionnaire de l'anciennce langue française et de tous ses dialectes* (1982). Apart from these, also Rothwell's *Anglo-Norman Dictionary* (1977-92) was used in order to be able to state with a certain degree of conviction whether a lexeme was, presumably, newly formed in the ME period or not.

Loans from French with the prefix *mes-* only started to enter the English language in M2 and the first new complex lexemes where *mis-* combined with a base of Romance origin can already be identified in the second half of the ME period. For instance, the verb *misavisen* (M3) was presumably a new formation on a French base. This assumption is supported by the dates of first

---

[11] If these dates coincided more or less or showed an overlap (the MED usually situates a text within a time span of 25 to 50 years), this was taken as a sign that derivative and base must have entered the English language (roughly) at the same time and could, consequently, both be regarded as loans. If, however, the complex lexeme was only attested at a later date, not overlapping with the date of first attestation of the base, it could be assumed that it might have been newly formed in ME.

attestation of the complex verb as well as its base. Furthermore, the dictionaries consulted reveal that although the base *avisen* existed in OF (*aviser*), there is no derivative from this base with the prefix *mes-* to be identified in OF and Anglo-Norman. Equally, the lexeme *misreuling* is an example of a new formation of M3. While the base *reuling* already appeared in the English language at the beginning of the ME period, the combination of this base with the prefix *mis-* has to be attributed to the second half of the ME period according to the MED.

Although new combinations of *mis-* with foreign bases do surface in M3, this trend is, however, not continued in M4. The fact that no new lexemes with either native or foreign bases can be identified in M4 does not speak in favour of the productivity of the prefix towards the end of the ME period.

4.3 Type of Text

Seeking information about the use and productivity of a prefix, it is, of course, interesting to see in what types of text it occurred in the history of the English language. Basically, text types have, up to now, been neglected more or less as a conditioning factor in historical studies on morphology and productivity. In fact, the first study in which an attempt was made to link frequency, productivity and text types seems to be Baayen's *Productivity in context: a Dutch suffix* (1997: 581ff.). Though Baayen's approach to text types differs slightly from my own, the idea that text type distribution and productivity may be interdependent is the same. I thus believe that important observations can be made when integrating text types into one's research on word-formation, as one can find out in which field a prefix was actually used: was it rather used in scientific texts or maybe in texts that were written to be spoken? Was it restricted to particular types of text only? Did it mainly appear in translations? Consequently, in order to gain further insight into the distribution of *dis-, in-, mis-* and *un-* in ME, an analysis of the types of text in which the four prefixes appear in the HC was carried out.

Analyses based on the HC are facilitated by a system of 24 reference codes which provides easy access to information about all text samples represented in the corpus. One of these reference codes, <T>, refers to the text type and allows the user of the corpus to find out quickly which type of text a particular text sample represents. In the HC, the different text types are further grouped into larger entities called "diachronic text prototypes" (Kytö 1993: 13). This grouping of text types was taken over from the HC (Kytö 1993: 10ff.), though slight adaptations were made:

- <u>Religious texts:</u> religious treatises, bible, sermons, homilies, rule
- <u>Secular instruction:</u> handbooks medicine, handbooks astronomy, handbooks other, science & medicine
- <u>Imaginative narration:</u> fiction, romances, drama – mystery plays
- <u>Non-imaginative narration:</u> history, biography (lives of saints), preface, depositions
- <u>Legal texts:</u> law, documents
- <u>Letters:</u> letters private, letters non-private
- <u>Philosophy</u>

4.3.1 *Dis-*

As could already be seen in section 5.1, *dis-* is only more strongly represented in the second half of the ME period and it is only then that a more conclusive picture presents itself as far as text type distribution is concerned.

dis-		M1	M2	M3	M4
Religious texts	religious treatises	0	2	7	11
	bible	0	1	3	0
	sermons	0	0	1	1
Secular instruction	handbooks medicine	0	0	0	1
	handbooks other	0	0	0	3
	science & medicine	0	0	2	0
Imaginative narration	fiction	0	0	9	0
	romances	0	2	0	9
	drama/mystery plays	0	0	0	3
Non-imaginative narration	history	0	0	2	1
	preface	0	0	0	1
Legal texts	documents	0	0	6	1
Letters	letters private	0	0	0	7
Philosophy		0	0	10	0
TOTAL			5	40[12]	38

Fig. 8. Dis- in different text types

Figure 8 shows that in M3 *dis-* is quite frequently attested in texts related to religion as, for instance, in Wycliffe's translation of *The Old Testament*, Chaucer's *The Parson's Tale* or the *English Wycliffite Sermons*. Apart from that, about half of all derivatives of M3 appear in fictional and philosophical texts by Geoffrey Chaucer: derivatives with *dis-* are attested in *The General Prologue to the Canterbury Tales*, *The Wife of Bath's Prologue*, as well as

Boethius and they are most prominently represented in *The Tale of Melibee*[12]. As far as *Boethius* and *The Tale of Melibee* are concerned, one may note that translation may have played an important role in the strong representation of *dis-* in these texts, as both were translated from Latin and French. [13]

While in M4 the representation of *dis-* in religious texts remains strong, the prefix's appearance in the category 'imaginative narration' is, compared to M3, significantly altered: the prefix is not at all attested in fictional texts in this final sub-period of ME. Consequently, one might conclude that it was mainly due to Chaucer that the prefix was so strongly represented in this text type in M3. In M4, however, about one quarter of all tokens occurs in romances and the prefix appears in plays for the first time, namely in the *Digby Plays* as well as *The Wakefield Pageants in the Towneley Cycle*. This is particularly interesting as drama is often said to reflect the characteristics of oral expression to a certain extent. Thus the fact that derivatives occur in this type of text in M4 suggests that they might also have been used in speech and not only in writing. Very revealing is, moreover, the representation of the prefix *dis-* in private letters, as in Clement Paston's letters to his brother or Margaret Paston's letters to her husband. This shows that towards the end of the ME period the prefix spread to more informal text types. Furthermore, due to their interactive character, letters also provide clues about the spoken language in the ME period. As private letters are of a more conversational character than other text types, the fact that derivatives with *dis-* are attested in this type of text in M4 implies that they might also have been used in oral conversation. Consequently, one can note that the prefix gradually gained more general currency towards the end of the ME period as it did not only appear in a wider variety of text types, but also in texts reflecting spoken language.

---

[12] Note that *The Tale of Melibee* is classified both as a fictional and a philosophical text in the HC. Derivatives occurring in this text sample were counted both as attestations of *dis-* in fictional as well as in philosophical text types. This leads to a certain overlap and results in a total of 40 instead of 34 tokens.

[13] While the HC lists *The Tale of Melibee* as having been translated from French, in the case of *Boethius* the HC mentions both Latin and French as possible languages of the foreign original.

4.3.2 *In-*

As far as text types are concerned, one can observe that the prefix *in-*, both in Latin quotes[14] and in already adapted loans, was restricted to the religious (and scientific) sphere in the first half of the ME period. This picture might, of course, have been altered subsequently due to the influx of many loans from French.

in-		M1	M2	M3	M4
Religious texts	religious treatises	2	0	12	11
	bible	0	3	1	0
	sermons	0	0	1	19
	rule	0	0	2	3
Secular instruction	handbooks astronomy	0	0	15	0
	handbooks other	0	0	0	9
	science & medicine	0	0	0	5
Imaginative narration	fiction	0	0	10	0
	drama/mystery plays	0	0	0	4
Non-imaginative narration	history	0	0	2	0
	preface	0	0	0	15
Legal texts	law	0	0	0	3
	documents	0	0	2	5
Letters	letters non-private	0	0	2	6
Philosophy		0	0	16	0
TOTAL		2	3	63[15]	80

Fig. 9. In- in different text types

While the prefix is still significantly represented in religious texts in M3, it also figures prominently in three other types of text, namely in fictional (Chaucer, Gower), philosophical (Chaucer's *Boethius* and *The Tale of Melibee*) and scientific texts (Chaucer's *A Treatise on the Astrolabe*). This shows that the strong representation of *in-* in M3 is mainly due to the frequent

[14] While negative *in-* is not very strongly represented during M1 and M2, one can identify several Latin words with *in-* which appear in religious and scientific texts either in Latin headings or in Latin quotes (cf. code switching). Consequently, one may assume that although not many loans with *in-* had yet been taken over from Latin and French, the fact that this derivational pattern was already represented in texts might have facilitated the subsequent adoption and integration of the prefix into the English language.

[15] Note that *The Tale of Melibee* is classified both as a fictional and a philosophical text in the HC. Derivatives occurring in this text sample were counted both as attestations of *dis-* in fictional as well as in philosophical text types which results in a total of 63 instead of 58 tokens.

occurrence of derivatives (36 out of a total of 58 tokens) in very diverse texts by Chaucer. Here one may again refer to the importance of translation as Chaucer's *Boethius* and *The Tale of Melibee* were both translated from Latin and French.

In M4, one can observe that the category 'religious texts' predominates: the dominating text type is the sermon which accounts for about two thirds of all religious texts in which derivatives with negative *in-* appear. Apart from that *in-* is also represented in the category 'secular instruction' in M4. Interestingly, the text *The Cyrurgie of Guy de Chauliac* (text type 'science & medicine') in which five tokens occur was a translation from Latin, i.e. the source language of negative *in-*. Compared to M3 when *in-* is frequently attested in fictional texts, in M4 *in-* is quite strongly represented in non-imaginative texts, namely in William Caxton's prologues and epilogues. Finally, one may note that the prefix also appears in texts of an official kind in M4 as, for instance, in legal texts, and it is attested in non-private letters by William Paston.

All in all, one can say that negative *in-* is very strongly represented in religious texts and it appears in text samples of a rather formal character, like scientific handbooks or legal texts. In contrast, the prefix's representation in more conversational text types, like drama, is rather low. Consequently, one may conclude that the generally held assumption that derivatives with negative *in-* are of a more learned, scientific character is reflected partly by the types of text in which it occurs in the ME section of the HC.

4.3.3 *Mis-*

In M1, when only derivatives with native bases of coining are attested in the corpus, the prefix *mis-* clearly prevails in the category 'religious texts'[16] - a trend which is continued in M2. However, a closer look at the data of M2 reveals that it is mainly one text that is responsible for the strong representation of *mis-* in this category. This is the religious treatise *Ayenbite of Inwyt or remorse of conscience* by Dan Michel which contains 26 out of a

---

[16] Before drawing any conclusions, it should be recalled, however, that the range of text types that is provided for M1 in the HC is not as diverse as that of later sub-periods. Religious texts constitute the majority of text samples that appear in M1 and, consequently, the numerous occurrences of derivatives with the prefix *mis-* in these types of text is not very surprising.

total number of 50 tokens with the prefix *mis-*. Interestingly, this text was a translation from French – a fact that might tempt one to assume that this might be the reason why so many derivatives with *mis-* appear in it. However, this argumentation does not hold as most of the derivatives that occur in Dan Michel's text have already been attested in M1 and all of them have got native bases of coining (e.g. *misdon, misdede, misliken*).

mis-		M1	M2	M3	M4
Religious texts	religious treatises	17	27	7	3
	bible	0	5	0	0
	homilies	15	1	2	0
	rule	0	0	2	0
Imaginative narration	fiction	0	4	5	0
	romances	0	9	0	0
Non-imaginative narr.	history	1	3	13	1
	biography	6	1	0	0
Legal texts	documents	0	0	10	2
TOTAL		39	50	39	6

Fig. 10. Mis- in different text types

In M3, the dominance of religious texts is weakened: while *mis-* still occurs quite frequently in these types of text, especially in religious treatises, it is equally strongly represented in other text types, namely in documents and historical texts. As far as documents are concerned, one might argue that the representation of *mis-* in this text type was reinforced by French influence, as all derivatives with *mis-* appearing in this text type have got French bases (e.g. *mischef, mischevous*). However, the same suggestion cannot be made for the text type 'history', as here the distribution of complex lexemes with native and foreign bases is more or less balanced.

In M4, the overall number of tokens with *mis-* is quite small. Consequently, only very tentative conclusions about text type distribution can be drawn. What can be noted, however, is that the trend which can be observed in M3 is continued: apart from religious treatises, derivatives with *mis-* are also represented in documents and historical texts in M4.

Thus one can say that while in the first half of the ME period one can still observe a prevalence of *mis-* in religious texts, subsequently a more balanced picture presents itself. However, as the data shows, the spread of the prefix *mis-* to a wider variety of text types cannot be exclusively attributed to the influx of French *mes-*.

4.3.4 *Un-*

As mentioned above, *un-* is attested with rather high type and token frequencies in M1. As far as text types are concerned, the negative prefix *un-* is not only represented in all text types in M1 but even in each individual text sample, with the category 'religious text' clearly dominating. This can be taken as a sign that *un-* was a very commonly used prefix at the beginning of the ME period.

un-		M1	M2	M3	M4
Religious texts	religious treatises	100	7	12	4
	bible	0	2	0	0
	sermons	0	0	0	3
	homilies	69	1	0	0
	rule	0	0	1	0
Secular instruction	handbooks medicine handbooks	5	0	0	0
	astronomy	0	0	1	0
	science & medicine	0	0	1	0
Imaginative narration	fiction	0	0	14	0
	romances	0	1	0	3
Non-imaginative narration	history	15	0	10	2
	biography	54	2	0	0
Legal texts	law	0	0	0	5
	documents	0	0	0	3
Letters	letters private	0	0	0	4
Philosophy		16	0	34	0
TOTAL		259	13	73	24

Fig. 11. Distribution of *un-* over different text-types

In M2, the wide distribution of the prefix in various text types is, however, significantly reduced. This is primarily due to the fact that the number of tokens drops sharply from M1 to M2, leaving M2 with a rather unrepresentative sample of derivatives with negative *un-*. Consequently, not very far-reaching conclusions can be drawn as far as the representation of negative *un-* in different text types is concerned.

In M3, *un-* does no longer occur in all text types.[17] One text type clearly dominates in M3, namely the philosophical text. The prevalence of *un-* in the

---

[17] One has to bear in mind that the number of different text types represented in this sub-period is much larger than that of M1.

text type 'philosophy' is mainly due to Chaucer's *Boethius* in which almost half of all tokens of M3 occur. Apart from philosophical texts, *un-* is also strongly represented in fictional texts, as, for instance, in several texts by Geoffrey Chaucer (*The General Prologue to the Canterbury Tales*, *The Wife of Bath's Prologue*, *The Summoner's Tale*, and *The Tale of Melibee*) and Gower's *Confessio Amantis*, as well as religious treatises and historical texts.

In M4, the prefix still appears in religious treatises and sermons, i.e. in text types in which it can be encountered from M1 onwards. Apart from that derivatives with *un-* occur in legal texts in M4, as, for instance, in *The Statutes of the Realm*. However, combinations with the prefix *un-* are also attested in less formal types of text, as, for instance, in Thomas Malory's romance *Morte Darthur*. Furthermore, the negative prefix *un-* also appears in private letters (Thomas Mull's letter to his brother in law, Thomas Betson's letters to his bride and to his mother in law) - a text type of a rather informal kind. This may be taken as an indication that derivatives with the prefix *un-* were presumably in common use.

As far as text type distribution is concerned, one can thus identify certain differences between *in-* and *un-* (cf. the categories secular instruction, letters, philosophy) which mirror the more specialised vs. the more general nature that is usually associated with these prefixes. Some results, like the fact that *in-* occurs in plays while *un-* does not or that the representation of *in-* and *un-* in legal texts is identical in M4, are, however, rather surprising.

4.3.5 Summary

In the above text type analysis several new insights could be gained. For instance, it could be shown that both foreign prefixes *dis-* and *in-* appear quite frequently in religious texts in the ME part of the HC. This is a highly interesting fact as in ME times the clergy usually had a knowledge of Latin and due to their familiarity with the foreign language they might have adopted the foreign prefixes more easily. Apart from that, it turned out that translation may also have played a role in the growing representation of the two prefixes. The occurrence of the prefix *dis-* in conversational text types like plays and private letters suggests that the prefix might not only have been used in writing but also in everyday speech and thus it might even have been in common use in ME. Also the attestation of *un-* in very diverse types of text speaks in favour of the wide distribution and usage of *un-*. On the other hand, it could be seen that *mis-* was more restricted as far as text types are

concerned and that also the influx of French *mes-* did not result in an increased use and wider text type distribution.

5. Conclusion and outlook

Taking all the evidence gained through this analysis of *dis-, in-, mis-* and *un-,* one comes to the following conclusion:

As far as the negative prefix *dis-* is concerned, it turned out that *dis-* cannot be regarded as productive in the ME period. This claim is supported by various results of the analysis: for instance, the prefix is not very strongly represented in M1 and M2 and only appears with larger type and token frequencies from M3 onwards. Furthermore, there are no hybrid formations with the foreign prefix *dis-* attested in the HC. Apart from that, the data also does not include any new formations of *dis-* with foreign bases of coining and intra-romance productivity can thus be ruled out as well.

Furthermore, this study also revealed that the prefix *dis-* is attested in very diverse types of text in the ME part of the HC: it appears in rather formal and more informal texts, in text samples of a non-interactive kind, as well as in texts of a conversational character. This wide text type distribution would actually speak in favour of the common use of the prefix. However, as one cannot identify any new formations with the prefix and as the number of different types with which *dis-* is represented in the corpus is rather low, one has to come to the conclusion that *dis-* was not a productive prefix in the ME period.

As in the case of *dis-*, also for the prefix *in-* higher type and token frequencies can only be observed from M3 onwards and neither hybrid formations nor new formations with foreign bases which were not borrowed from French or Latin are attested in the data (cf. Dalton and Mettinger 1993). The analysis of text types showed that *in-* features in several different types of text in the HC but not in informal texts of a conversational kind. However, derivatives with *in-* occur quite frequently in religious and philosophical texts as well as in handbooks and formal texts, like statutes and documents. Taking all these factors into consideration, one might conclude that the prefix cannot yet be called productive in ME. However, there are signs that the prefix was on its way to becoming productive in its restricted sphere and, consequently, an analysis of the further development of the prefix in EModE might yield interesting results.

The analysis of the prefix *mis-* showed that combinations of *mis-* with native bases were generally on the decline during the ME period. In M2,

however, combinations with foreign bases appear in the data. A steady increase of these derivatives can be observed until M3, when also new formations with foreign bases appear. Nevertheless, a swift fall follows from M3 to M4 and type and token frequencies are very low in this final sub-period. An analysis of the representation of the prefix in EModE times might reveal how far this sudden drop fits the further development of the prefix.

As far as text type distribution is concerned, it has to be noted that *mis-* is the most restricted of the four prefixes analysed in this study. In M1 and M2, it mainly appears in religious texts and also in the second half of the ME period it primarily occurs in three types of text, namely in religious and historical texts as well as in documents.

Thus the picture that presents itself is two-fold: on the one hand, new formations with both native and foreign bases can be identified, though their number decreases towards the end of the ME period. On the other hand, type and token frequencies are on the decline from M2 onwards and the prefix is hardly represented at all in M4. Consequently, I would tentatively conclude that *mis-* was no longer fully productive towards the end of the ME period.

Finally, for native *un-* it could be observed that while the prefix is still represented very strongly in a large number of OE formations in M1, subsequently, type and token numbers drop sharply. However, what the data also shows is that *un-* was, nevertheless, used in the formation of new lexemes, though many of these new formations with *un-* went out of use again and are marked as obsolete in ModE. This, however, also shows the dynamic character of the prefix which was freely attached to bases of both native and foreign origin. Some of them only had a temporary existence, but the vitality with which the prefix was employed can, nevertheless, be observed.

As far as text type distribution is concerned, it can be noted that at the beginning of the ME period, *un-*, being represented in all text samples of M1, was very widely used. But also towards the end of the ME period, *un-* is attested in several different text types. What is particularly noticeable is the diversity of texts in which *un-* occurs: on the one hand, it appears in very formal texts but it also occurs in rather informal texts, like private letters.

If one takes all the evidence together, one can come to the conclusion that although the prefix was not as strongly represented in terms of type and token frequencies towards the end of the ME period as at the beginning, it was, however, vitally involved in the production of new lexemes, among them also a considerable number of hybrid formations. Consequently, this analysis clearly shows the productive nature of native *un-* in ME.

The empirical analysis of the four prefixes *dis-, in-, mis-,* and *un-* thus yielded many interesting results, and further insights into the particular nature of these prefixes in the ME period were gained. Some generally held assumptions (e.g. the semantic restriction of negative *in-* or the continuous vitality of *un-*) found support, while others (e.g. the increased productivity of *mis-* after its merger with French *mes-*) were refuted. What remains to be investigated is the further development of the four prefixes *dis-, in-, mis-* and *un-* in the EModE period. However, also a more comprehensive study comprising other ME prefixes still needs to be conducted in order to broaden our understanding of historical English word-formation.

References

Anshen, F. & Aronoff, M. (1981). Morphological Productivity and Phonological Transparency. *The Canadian Journal of Linguistics*, 26:1, 63-72.

Aronoff, Mark (1976). *Word-formation in Generative Grammar*. Cambridge, Massachusetts: MIT Press.

Aronoff, M. (1980). The relevance of productivity in a synchronic description of word formation. In J. Fisiak (ed.) *Historical Morphology*. The Hague: Mouton Publishers, pp.71-82.

Baayen, H. & Lieber, R. (1991). Productivity and English derivation: a corpus-based study. *Linguistics*, 29.5, 801-843.

Baayen, H. (1992). Quantitative aspects of morphological productivity. In G. Booij & J. van Marle (eds.) *Yearbook of Morphology 1991*. Dordrecht: Kluwer Academic Publishers, pp. 109-149.

Baayen, H. (1993). On frequency, transparency and productivity. In G. Booij & J. van Marle (eds.) *Yearbook of Morphology 1992*. Dordrecht: Kluwer Academic Publishers, pp. 181-208.

Baayen, H. & Renouf, A. (1996). Chronicling the *Times*: Productive lexical innovations in an English newspaper. *Language*, 72.1, 69-96.

Baayen, H. & Neijt, A. (1997). Productivity in context: a case study of a Dutch suffix. *Linguistics*, 35.3, 565-587.

Bauer, Laurie (2001). *Morphological productivity*. Cambridge: CUP.

Baugh, A.C. & Cable, T. (1993). *A History of the English Language*. 4th ed. London: Routledge.

Bolozky, Shmuel (1999). *Measuring productivity in word formation. The case of Israeli Hebrew*. Leiden: Brill.

Burnley, David (1992). Lexis and Semantics. Word-formation. In N. Blake (ed.) *The Cambridge History of the English Language 1066-1476. Volume II*. Cambridge: CUP, pp.439-450.

Cowie, C. & Dalton-Puffer, C. (2002). Diachronic word-formation and studying changes in productivity over time: Theoretical and methodological considerations. In V. Diaz & E.

Javier (eds.) *A Changing World of Words. Studies in English Historical Lexicography, Lexicology and Semantics*. Amsterdam: Rodopi, pp. 410-437.

Dalton-Puffer, C. and E. Mettinger-Schartmann. (1993). Frenglish? Sur la productivité de la morphologie francaise dans le moyen anglais. *Travaux de linguistique et de philologie* 31, 183-193.

Dalton-Puffer, Christiane (1996). *The French Influence on Middle English Morphology. A Corpus Based Study of Derivation*. Berlin: Mouton de Gruyter.

Faiß, Klaus (1992). *English Historical Morphology and Word-Formation*: *Loss versus Enrichment*. Trier: Wissenschaftlicher Verlag Trier.

Godefroy, Frédéric (1982). *Dictionnaire de l'ancienne langue française et de tous ses dialectes du IXe au XVe siècle*. Genève: Slatkine.

Gretschmann, Ursula (1994). The French component in Modern English word-formation: diachrony within synchrony. In R. Bacchielli (ed.) *Historical English Word-Formation. Papers read at the Sixth National Conference of the History of English*. Urbino: Quattro Venti, pp.67-76.

Hansen, Barbara (et al.) (1982). *Englische Lexikologie. Einführung in die Wortbildung und lexikalische Semantik*. Leipzig: VEB.

Hindley, Alan (et al.) (2000). *Old French – English Dictionary*. Cambridge: CUP.

Kastovsky, Dieter (1992). Semantics and vocabulary. Word-formation. In R.M. Hogg (ed.) *The Cambridge History of the English Language. The Beginnings to 1066. Volume I*. Cambridge: CUP, pp. 355-400.

Kurath, Hans (et al.) (eds.) (1956-) *Middle English dictionary*. Ann Arbor: University of Michigan Press.

Kytö, Merja (1991). *Manual to the Diachronic Part of the Helsinki Corpus of English Texts. Coding Conventions and Lists of Source Texts*. Helsinki: Helsinki University Press.

Kytö, M. & Rissanen, M. (1993). General Introduction. In M. Rissanen (et al.) (eds.) *Early English in the Computer Age. Explorations through the Helsinki Corpus*. Berlin: Mouton de Gruyter, pp. 1-17.

Lutzky, Ursula (2004). *Prefixation in Middle English: dis-, in-, mis- & un-*. Unpublished M.A. Thesis, University of Vienna.

Marchand, Hans (1969). *The Categories and Types of Present-Day English Word-Formation*. 2nd rev. ed. München: Beck.

McConchie, Roderick (2000). Fashionable idiolects? The use of the negative prefix *dis-* 1520-1620. In D. Kastovsky & A. Mettinger (eds.) *The History of English in a Social Context: A Contribution to Historical Sociolinguistics*. Berlin: Mouton de Gruyter, pp. 199-225.

OED = The Oxford English Dictionary. Second Edition. On CD-ROM. Version 3.0 (2002) Oxford: OUP.

Plag, Ingo (1999). *Morphological Productivity. Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter.

Plag, I. *et al*. (1999). Morphological productivity across speech and writing. *English Language and Linguistics*, 3.2, 209-228.

Rothwell, William (et al.) (1977-92). *Anglo-Norman Dictionary*. London: The Modern Humanities Research Association.

Van Marle, Jaap (1992). The relationship between morphological productivity and frequency: a comment on Baayen's performance oriented conception of morphological productivity. In G. Booij & J. van Marle (eds.) *Yearbook of Morphology 1991*. Dordrecht: Kluwer Academic Publishers, pp. 151-163.

Zbierska-Sawala, Anna (1993). *Early Middle English Word Formation*. Frankfurt am Main: Lang.

'Hec sunt prata to wassingwellan': Aspects of code-switching in Old English charters

Herbert Schendl, University of Vienna[*]

1. Introduction

In 1996 and 1997 my first two papers on historical code-switching were published in *VIEWS*, dealing with the phenomenon of language mixing in historical English texts in general and in so-called 'macaronic poems' in particular. Based on a paper delivered at the LACHE conference in 1994, I claimed that such early language mixing should be seen as written instances of historical code-switching and should be analysed using the theoretical models developed by current code-switching research.[1] Since then, historical code-switching research has established itself as a promising field of English historical linguistics and an increasing number of studies have been devoted to the topic.[2] However, for a number of reasons, this research has so far mainly focused on the Middle English period: the extensive multilingualism in post-Conquest England, at least among literate people, provided the ideal background for code-switching, and there is a large number of texts from a wide range of text types and genres which mix Latin, English and/or French (e.g., business accounts, sermons, letters, year books, poems, medical texts, etc., see Schendl 1996). As shown in some of these studies, the syntactic forms as well as the functions of Middle English switching are quite diverse and partly vary in the different text types. So far, however, there has been hardly any published research on language-mixing in the Old English period. This may be due to the widespread view that the Anglo-Saxons did not mix languages, and the fact that different scripts were used for writing Latin and

[1] For surveys of different theoretical models of code-switching see, e.g., Romaine 1995, Muyskens 1995.

[2] For a survey of research up to 2000 see Schendl 2002a. The phenomenon is not restricted to Britain, but is found all over medieval and early modern Europe; see Schendl 2004 for a comparison of English historical mixed texts with those from other European countries and languages, including further references. The present papers, however, will only deal with the situation in Anglo-Saxon England.

Old English (Caroline vs Anglo-Saxon minuscule) has been quoted as further support for the strict separation of languages in Old English (see Howlett 1997).

In this paper I would first like to show that code-switching was quite frequent in certain texts and text types from the Old English period, even though it was certainly less widespread than in the Middle English period and may often have been done consciously, as the use of different scripts seems to indicate; secondly, I will briefly discuss and illustrate the main syntactic types of and functional reasons for switching, though this must remain rather sketchy in the present paper, which tries to survey the whole Old English period.

There are at least two Old English text types in which code-switching between Latin and Old English is attested: these are, firstly, two well-known mixed or 'macaronic' poems, the *Phoenix* and the so-called *Macaronic Poem*. The *Phoenix*, after more than 600 alliterative lines in Old English, has a bilingual coda of eleven lines, with the first half-line in Old English, the second half-line in Latin; both half-lines are connected by alliteration, which supports the integration of the two languages, see (1).[3]

(1) *Phoenix*:   Hafað us alyfed   **lucis auctor**
                          Þæt we motum her   **mereri**,
                          goddædum begietan   **gaudia in celo**.

                          ['(He) has granted us   **The author of light**
                          that we may here   **merit**,
                          with good deeds obtain,   **the joys in heaven**.']

A similarly regular pattern of switching between the Old English first half-line and the Latin second half-line is found in the 31 lines of the Old English *Macaronic Poem*; as in the *Phoenix*, the two half lines are also linked by alliteration.

The present paper will, however, focus on a second group of texts, in which code-switching is attested on a large scale, namely the legal and administrative documents commonly called charters. Charters are short legal

---

[3] In the quoted examples, code-switched passages are indicated by bold print, while bold plus italics refers to cases where no unambiguous decision on the language can be made, such as the preposition *in*, which can be both Latin and Old English, see example (5).

documents typically recording a grant or lease of land or certain privileges to individuals or institutions. They first appeared in England in the seventh century and developed both formally and linguistically to the end of the Old English period (see Keynes 1999, s.v.). Various subclassifications of these documents have been proposed. I will follow here Whitelock's (1955: 343) classification, who uses 'charter' as the superordinate term, which she subdivides into two classes, namely royal and private charters; the 'royal charter' in turn is subdivided into the so-called 'diploma' (often also referred to as 'charter' in the narrow sense of the word) and the 'writ'.[4] The diploma is a highly formal document with a clear structure, predominantly in Latin, while writs are in the vernacular. The non-royal documents often follow the royal ones quite closely and have equally been subdivided into 'writs' and charters in a wider sense.

The majority of the royal 'diplomas' show a distinctive structure, with, among others, an invocation, followed by the proem (i.e., preamble), a dispositive section (with immunity clause, statement of powers, etc.), the sanction, the boundary clause (or 'perambulation'), the dating clause, and the witness list; an endorsement may summarise the substance of the charter (for a complete list of elements see Keynes 1999: 99; see also Whitelock 1955: 344; Chaplais 1965); however, not all typical elements have to be present in every single document.

The 'writ', on the other hand, is "a letter on administrative business to which a seal was appended", addressing a variety of issues and from its first occurrence typically written in Old English; its opening clauses names the sender of the letter and the person(s) to whom it is addressed (Harmer 1952: 1; see also Whitelock 1955: 345f.).

About 2,000 writs and charters have survived from the Anglo-Saxon period, of which more than half are diplomas. They date from the last quarter of the 7th century up to the Norman Conquest, with a concentration around the middle of the 10th century (cf. Clanchy 1993: 1). About 300 of the diplomas are originals, many are later copies preserved in cartularies, i.e. later compilations of documents, and there is also quite a number of certain or possible forgeries; such forgeries were produced to support a (genuine or fraudulent) claim, but are not based on an original charter issued by the person named in the respective document. The linguistic value of forgeries may be doubtful, while the later copies generally show a high degree of

---

[4] For a different classification of the various types of 'charters' see Keynes 1999, s.v.

linguistic reliability (cf. Kitson 1995); therefore we will also assume that copying has not levelled out or introduced code-switching into a text to any large degree, though we can neither prove nor disprove this assumption and have to take the surviving data in their recorded form.

2. Code-choice and code-switching in the Old English charter

This section will analyse and illustrate some aspects of language choice and language mixing, i.e., code-switching in the various types of charters, especially in regard to the function of switching.[5] There will be no systematic analysis of the syntactic switching patterns, though some syntactic aspects will be pointed out briefly (for a discussion of possible syntactic constraints in historical code-switching see Schendl 2000b). It should be emphasized right from the start that code-switching does not occur in all types of charters and that the form and function of switching seems to differ depending on the specific subtype and period.

The writs (both royal and non-royal) seem to have been exclusively monolingual in the vernacular: none of the 120 writs in Harmer's edition (1952) shows any Latin elements (see also Whitelock 1955: 346), though, on the other hand, some of the (often later) Latin versions of these vernacular writs have the legal terms for the granted privileges in Old English, cf. the early 12th century Latin translation of an 11th vernacular writ by King Edward under (2):[6]

(2) *Latin translation of a writ by King Edward, c 1052-66* (Harmer 1952, nr 34)

Ego Eadwardus gratio Dei rex Anglorum omnibus episcopus, ducibus, comitibus ... amicabiliter salutem. Notum vobis facio me concessisse Stigando archiepiscopo 7

---

[5] In Schendl forthcoming, I have analysed in some detail the more than 70 leases issued by St Oswald of Worcester in the second half of the 10th century, relating code-switching to variables such as time of composition and addressees of the individual charters. – It goes without saying that language choice and mixing in the charters has been commented upon in the extensive traditional literature on Anglo-Saxon charters (see, e.g., Whitelock 1955, Stenton 1955, etc.) and there are also linguistic studies based on the Old English boundary clauses, such as Kitson 1995; but, to my knowledge, mixed charters have not been studied as manifestations of Old English code-switching, even less so from a (socio)linguistic point of view.

[6] For a detailed discussion of Anglo-Saxon legal terms see, e.g., the Introduction in Harmer 1952.

monachis ecclesie Christi Cant' omnes terras quas habuerunt tempore patris mei 7 omnium antecessorum meorum. Et **saca 7 socne on stronde 7 streame, on wode, on felde, tolles an[d] teames, griðbreche 7 hamsocne 7 forstalles 7 infangeneþefes 7 flemenfremthe** supra suos homines infra urbes 7 extra, in tantum 7 tam plenarie sicuti proprii ministri mei exquirere deberent, et etiam super tot *thegnes* quot habent. [continued in Latin]

['I, Edward, by the grace of God king of England send friendly greetings to all my bishops and all my earls and all my sheriffs ... And I inform you that I have granted Archbishop Stigand and the community at Christ Church, Canterbury, that they be entitled to all the lands that they had in the time of my predecessors and in my time. And **to sake and soke, on strand and in stream, in woodland and in open country, to toll and to team, to grithbreach and to hamsocn, to foresteall and to infangenetheof and to flymenafyrmth**, over their own men within boroughs and without, as fully and completely as my own officers would exercise it, and over as many *thegns* as they have.']

While code-switching is not attested in the vernacular writs, there are quite a number of switches to Latin in predominantly vernacular non-royal 'charters'. About 10% of the 135 texts in Robertson's vernacular *Anglo-Saxon Charters* (2nd ed. 1956) have at least one Latin phrase or sentence, some even show a considerable amount of Latin. Some typical Latin passages in otherwise vernacular documents are listed under (3), though even these elements are frequently found in the vernacular. They occur typically, though not exclusively, in the invocation, dating clause and witness lists and are predominantly of a formulaic nature.

(3) *Selected formulaic Latin sentences in Old English vernacular writs and charters*

Invocation:
In nomine patris et filii et spiritus sancti
In nomine domini (et saluatoris nostri Ihesu Christi)
Regnante in perpetuum domino nostro Ihesu Christo

Dispositive section:
Disponente regi regum cuncta cæli secreta necnon quæ sub cæli culmine apud homines notantur miro ordine gubernante cujus incarnationis humanæ anno DCCCCXXII indictione x hæc donatio quæ in ista cartula saxonicis sermonibus apparet confirmata ac donata erat.
['In the year 922 of the incarnation of the King of kings, who governs all the secret things of heaven and likewise controls in marvellous order those which are recorded among men under the dome of heaven, and in the tenth Indiction, this grant which appears in this charter in the Saxon speech was confirmed and given.'] (Robertson 1956: 42, nr. 21)

Dating clause:

Anno uero dominice incarnationis .DCCCLII. Indictione .xv. hoc factum est

['This has been done in the year of the incarnation of the Lord 852 and in the fifteenth Indiction']

<u>Witness list</u>:

hii sunt nomina et uocabula qui in synodali concilio fuerunt congregati;

['Here are the names and designations of those who were assembled at the council meeting']

Other typical switches into Latin consist of quotations from the bible or of a statement that the land being granted is free from any burdens except the church dues.

A more balanced use of Latin and Old English is, however, found in grants or leases of land issued by the bishops of Worcester, especially by St Oswald, dating from the second half of the 10th century. The example under (4) shows a slight predominance of Old English over Latin, with an obvious functional distribution of the two languages: the conditions for the lease and the boundary of the estate are in Old English, as is a – possibly later – statement by the receiver of the grant, Ælfric, while the more formulaic elements are in Latin. Except for the place names, all the switches are intersentential, i.e., they occur between sentences or independent clauses.

(4) *Lease of three lives by Oswald, Bishop of Worcester, of land at Cotheridge, Worcestershire, A.D. 963* (MS Cotton Tiberius, A. xiii, f. 63; Robertson 1956, nr 35)

Ego Osuuald ergo Christi krismate pręsul iudicatus dominicę . Incarnationis anno . DCCCLXIII annuente regi Anglorum . EADGARO . Ælfereque Merciorum comite nec non et familiae Wiogornensis ecclesiae . quandam ruris particulam unam uidelicet mansam in loco qui celebri a solicolis nuncupatur **ÆT CODDAN HRYCCE** uocabulo cuidam ministro meo nomine . Ælfric . perpetua largitus sum heredidate et post uite suę terminum duobus tantum heredibus immunem derelinquat quibus defunctis ecclesiae dei in **Uuigorna ceastre** restituatur.

On þæt gerad þe he ælce geare of þam lande ge erige twegen æceras 7 þær on his circ sceat gesawe 7 þæt æft ge ripe 7 in ge bringe . 7 ic an him ælce geare on minum wudu . XII . foþre wudas butan ceape .

Ðis sindan þa land ge mæru to coddan hrycge . ærest up ón temedan and lang biscopes gemæres norð rihte ón ætinc weg of ætinc wege in coddan hrycges bece and lang beces tó bricge burnan fordes þanan and lang stræte þæt hit cymeð beneoþan oban treow þanan suð rihte and lang þære hege ræwe in rixuc andlang rixuc ón hihtes gehæg þanan suð rihte in þa stræte and lang stræte þ in bregnes ford up and lang temedan þæt eft on biscopes gemære.

Scripta est hæc carta is testibus consentientibus quorum inferius nomina notantur. [List of 15 witnesses]

Ic Ælfric cyþe minan leofe hlaforde þæt ic ón Æþelsige minan suna þæs landes þe ic to gearnode æfter minan dæge to habbanne his dæg 7 æfter his dæge to syllanne þæm þe him leofast seo 7 þæt sio on þa spere hand.

['I, Oswald, having been ordained bishop by the chrism of Christ, in the year of our Lord's incarnation 963, with the consent of Edgar, king of the English, and of Ælfhere, ealdorman of the Mercians, and also of the community of the church of Worcester, have granted in eternal inheritance a certain portion of land, namely one hide in the place which is called by the inhabitants by the well-known name of **Cotheridge**, to a certain thegn of mine by name of Ælfric; and after the end of his life he is to leave it unburdened to two heirs only, and when they are dead it is to be restored to the church of God in **Worcester**.

On condition that each year he plough two acres of that land and sow therein his church-scot, and afterwards reap and garner it. And I grant him each year twelve loads of wood in my woodland without payment.

These are the boundaries to Cotheridge: First up the Teme and due north along the bishop's boundary to the Atchen way; from the Atchen way to Cotheridge stream, and along the stream to Bridgeburnan ford, thence along the paved road until it comes below Oba's tree, thence due south along the Hedgerow to the Rixuc, along the Rixuc to Hihtesgehaeg, thence due south to the paved road, and then along the paved road to Bransford, up along the Teme, then back to the bishop's boundary.

This charter was written with these witnesses consenting whose names are noted below.

I, Ælfric, make known to my dear lord that I grant to my son Æthelsige the land which I acquired of you, after my death to have for his lifetime and after his death to give to whom may be most pleasing to him, and that is to be in the male line.']

So much for a brief illustration of some tendencies of code-switching in writs and non-royal charters. Now to the royal charters (or 'diplomas'), which are often claimed to be monolingual in Latin (see Keynes 1999: 99), but which are linguistically much more complex than usually claimed. The first of these clearly structured royal documents, which normally grant land to a person or religious house, date from the last quarter of the 7th century and have survived from various parts of England. Even in these early documents we sometimes find a brief description of the boundaries of the granted land, and up to the end of the 8th century, these boundary clauses were, like the rest of the charter, generally in Latin. However, around the middle of the 8th century, "detailed perambulations begin to appear, particularly in Wessex, **set out at first in Latin and English**" (Stenton 1955: 56, my emphasis).

Stenton's dating of the beginning of switched Old English boundary clauses has, however, to be antedated by more than half a century, since this type of switching is – though rarely - already attested from the late 7th century onwards. King Cædwalla's charter from AD 680 under (5), though surviving only in a 10th century copy, shows intrasentential code-switching even *within* the boundary clause: phrases indicating direction such as 'from the west', 'eastward', 'from that place', etc. are in Latin, while the actual identification of the location is predominantly in the vernacular. Quite generally, charters use the Old English forms of place names even in Latin texts, but often identify them as vernacular forms by a preceding *qui dicitur*, *ubi ... nominatur*, *quod nostra lingua ... nominamus*, 'which is / where it is called (in our language) ...' etc. This convention of naming locations in the vernacular may be regarded a particular, though linguistically minor, function of Old English code-switching, while the preceding *qui dicitur*, etc. can be seen as a kind of 'flagging' device to indicate the change of language.

Of particular interest in (5), however, is the switching of the governing prepositions in (or before) such place names and names of landmarks. Here we find both instances where the whole PP, i.e. preposition plus dependent noun, are in Old English, such as *on fleot* 'to the stream', *on loxan leage* 'to Loxley', *to hleap mere* 'to Leapmares', etc., but sometimes also cases where the preposition is in Latin, while the dependent noun is in Old English, as in **ad** *pecgan ham* 'to Pagham'. With the preposition *in* an unambiguous classification of language is not possible. However, we also find complete Latin PPs in this naming function, such as *per viam puplicam ad terram heantunensem ad angulum circianum* 'along the highway to the Easthampnett land, to the bend' or *dirigitur in mare* 'runs to the sea' (as against *on fleot* 'to the stream', see above). While the use of the vernacular for the place names can be explained as a conventional strategy to make identification easier, there is no discernible functional motivation for the use of Latin forms here; however, the impossibility to find functional explanations for every single instance of switching is a well-known phenomenon in current code-switching research and not untypical of the use of switching as a particular mode of discourse.[7]

---

[7] For an analysis of the language patterns of PPs in Middle English mixed texts see Schendl 2000b: 78, for code-switching as a particular mode of discourse see Poplack 1980: 614, Schendl 1997: 61ff.

(5) *Grant by Cædwalla, king of Wessex, to Bishop Wilfrid of land at Pagham, &c. A.D. 680* **(**BL Cottton MS Augustus ii, 86, 10th c. copy; Barker, 1947, nr 1)

Si quis vero quod absit contra hæc decreta firmiter statuta contraire et ea solvere conatus fuerit, noverit se ante tribunal examinis christi rationem redditurum et habere partem cum iuda traditore domini nostri ihesu christi in inferno inferiore. hæc sunt territoria ad **pecgan ham** pertinentia primitus ab occidente **uedringmutha**. per illum portum ad locum qui dicitur **holan horan fleot** et sic ducitur *in* **lang port**. inde ad aquilonem **to unning londe**. sic ad orientem **on fleot** super illud quod dicitur **in ufes ford**. inde in locum qui dicitur **cynges uuic**. et sic ad locum qui dicitur **langan ersc**. inde **on loxan leage**. ... inde in pontem **thelbrycg**. et sic ad aquilonem juxta palustria loca. super hoc ad locum qui dicitur **hylsan seohtra** et sic ad orientem *in* **uuærmundes hamm**. Hinc *in* **uuadan hlæu**. Ab illo loco *in* **fisc mere**. et sic *in* **brynes fleot**. sicque dirigitur in mare. Sed et hi sunt termini pertinentibus ad **tang mere**. primitus **of hleap mere** per viam puplicam ad terram heantunensem ad angulum circianum. inde in locum **horsa gehæg**. ... hinc ad **ælrithe**. ab ipso rivo ad fraxinum unum. et sic ad locum **cealc mere**. hinc ad **headan screaf**. ab illo loco **to lulan treouue**. et sic *in* **tatan ham**. sic ad **risc mere**. ab illo loco **to hleap mere**. Et sunt pascua ovium in meos dune pertinentia ad **tangmere**.

['And (may it not be) if anyone attempts to go against and nullify this firmly established decree, let him know that he will have to render account before the judgement seat of Christ, and will take his place with Judas, the betrayer of our Lord Jesus Christ, in the depths of Hell. These are the lands belonging to **Pagham**, firstly from the west of **Withering**, by that harbour to the place which is called **Honer stream**, and thus it leads *to* (the) **long village**, thence northwards **to Unna's land**, so eastward **to the stream** and over it at the place called (*in*) **Ufa's ford**, thence to the place called the **king's farm**, and thence to the place called **Lagness**, thence **to Loxley**, ... thence to the bridge (at) **Elbridge**, and thus northwards beside the marshy places, over this to the stream called (**Aldingbourne Rife**) and thence east *to* **Wærmund's enclosure**, thence *to* **Wada's barrow**, from that place *to* **the fish pond**, from there *to* (**Ryfebank Rife**), and so the line runs to the sea. And these are bounds belonging to **Tangmere**, firstly **from Leapmares** along the highway to the Easthampnett land, to the bend, thence to the (place) **horse paddock**, ... thence to (the) **eel brook**, from that brook to the single ash-tree, and thus to the place (called the) **chalky pool**, thence to the (?) **deep cave**, from that place **to Lula's tree**, and thus *to* **Tata's farm**, thus to (the) **rushy pool**, and from there **to Leapmares**. And there are sheep pastures belonging to **Tangmere**.']

This mixture of Latin and the vernacular within the boundary clause is also frequently found in West Saxon charters from the late 8th century (cf. also Stenton 1955: 26), as in Cynewulf's grant to Bica from the year 778 under

(6). The general insensitivity to this type of code-switching in traditional research is illustrated by the fact that the language of this charter is given as Latin in Sawyer's list of charters (Sawyer 1968, nr 264), i.e., no explicit mention is made of the Old English prepositional phrases denoting direction *to* the various places, such as *to bradan leage* 'to Bradley' in the otherwise Latin text (this is the more interesting since the citation form or nominative of place names often has the form *æt* + noun, e.g. *æt bradan leage*).

(6) ***Grant by Cynewulf, King of the Saxons, to Bica, Earl, of land at Bedwyn, A.D. 778*** (Birch, nr 225)

[…] et sic in affricum vergens in longum illius septi tendit ad **peadan stigele** deinde per iddem sept[um] *in* **filiðleage** . australem partem inde in longum prædicti septi in quoddam vallu *in* . **haradene** . ... et extenso tramite ejusdem septi . **to holhrygc gete** . et eodem septo **to hadfeld geate** . et eodem septo **to baggan gete** . et sic in illud septum . **to bradan leage** . transitque . illo septo **bradanleage** intrans *in* . **standene** . et in longum ejus in quoddam vallum ejusque . valli serie *in* . **puttan ... ealh** . et sic in longum aggeris **to bulcan pytte**.

Such mixed boundary clauses seem to represent an intermediate stage between the early monolingual Latin texts and the later monolingual vernacular ones, which first appear in the early part of the 9th century. However, there was evidently some temporal overlap between the various types: even in the middle of the 9th century we still find examples like the one under (7): in this original charter issued by Ethelberht of Kent, code-switching also occurs within the boundaries, with a number of Old English prepositional phrases denoting place names inserted into the Latin text; furthermore, some of the other goods and privileges granted in the charter are also given in the vernacular within the otherwise Latin text, such as *I . sealtern 7 II . wena gang mid cyninges wenum to blean ðem wiada 7 . IIII . oxnum gers mid cyninges oxnum* 'and one salthouse and [the right for] two wagons to go with the king's wagons to Blean wood, and pasture for four oxen with the king's oxen'. Other grants, however, are in Latin, such as *II.que molina ad illam eandem terram pertinentia* 'and two mills belonging to the same land', etc. That switching here (as in general) is not due to lack of words is illustrated by the alternate use of *ab aquilone* and *be norðan*, both meaning 'in the north'. (For this type of variation in the choice of language see also the discussion of PPs in example (5).)

(7) *Grant by King Ethelberht of Kent to Wullaf, of land at Wassingwell, in exchange for land at Merham, A.D. 858* (MS Cotton, Augustus II, 66; original charter; Birch, nr 496)

Regnante in perpetuum domino Deo nostro omnipotenti sabaot ego EÐELBEARHT rex cum consensu ac licentia meorum secularium optimatum divinorumque personarum liventi animo dabo et concedo meo fideli ministro **WULLAFE** aliquam partem terre juris mei hoc est . V . aratra in illa loco ubi **WASNGWELLE** nominatur in bicissitudinem alterius terre hoc est **et mersaham** hanc terram supranominatam **et wassingwellan** ego Eðelbearht ab omni servitute regali operis eternaliter liverabo sicut ante fuerat illa prenominata terra **et mersaham** hec sunt etenim marisci que ad eandem terram rite ac recte pertinent quos 'h'ega ante abuerat id est **an wiwarawic** quae ante subjecta erat **to wii** 7 **to leanaham** 7 **et febresham** . I . **sealtern** 7 . II . **wena gang mid cyninges wenum to blean ðem wiada** 7 . IIII . **oxnum gers mid cyninges oxnum an wiwarawic** . XXX . statera kasei et item . X . statera in alia **wiwarawic** 7 . XX . **lamba** 7 . X . **fehta** hec autem terra suprascripta **et wassingwellan** his notissimis terminibus antiquitus circum jacentibus ab occidente **cyninges folcland** quod abet wighelm 7 wulflaf ab aquilone **cuðrices dun heregeðeland** ab oriente **wighelmes land** a meritie **biscepes land to cert** . II.que molina ad illam eandem terram pertinentia una **an wassinwellan** alia **an hwiteceldan** hec sunt pascua porcorum quot nostra lingua *denbera* nominamus hoc est **lamburnanden orricesden teligden stanehtandenn** et illa silva **sandhyrst** nominatur que pertinet **to wassingwellan** hancque livertatem huic eodem agel`l´ulo illoque **wullafe** similiter **et wassingwellan** cum consensu ac licentia meorum optimatum liventer largitus sum ut omnium regalium tributum et vi exactorum operum et penalium rerum principali dominatione furisque conprehensione et cuncta seculari gravidine absque expeditione sola et pontium structura et arcium munitionibus secura et inmunis permaneat ...

Hec sunt prata **to wassingwellan stocmed healf be norðan hegforde be sturemeda sue ðer to limpað**. [List of 23 witnesses]

['Our Lord Almighty God of Hosts reigning for ever.

I, King Ethelbert, with the consent and permission of my secular nobles and religious dignitaries, with willing heart will give and concede to my faithful thegn **Wulflaf** some portion of land of my rightful possession, namely five ploughlands in the place which is called **Wassingwell**, in exchange for other land, namely **at Mersham**. I, Ethelbert, will free eternally this above-mentioned land **at Wassingwell** from all liability to royal service just as the afore-mentioned land **at Mersham** was before. These indeed are the marshes which duly and rightly belong to the same land, which marshes Hega had before: i.e. **one dairy-farm of the people of Wye**, which before was subject **to Wye** and **to Lenham**, and one **salthouse at Faversham**, and [the right for] **two wagons to go with the king's wagons to Blean wood**, and **pasture for** four **oxen with the king's oxen**; **in the dairy-farm of the people of Wye** 30 weys of cheese, and also 10 weys in the other **dairy-farm**

of the people of Wye, and 20 **lambs** and 10 **fleeces**. And the above-written land **at Wassingwell** [has] from of old these well-known boundaries lying round it: in the west, **the king's folkland**, which Wighelm and Wulflaf hold; in the north, **Cuthric's down**, **Heregetheland**; in the east, **Wighhelm's land**; in the south, (the) **bishop's land at Chart**; and two mills belonging to the same land, one **in Wassingwell**, the other **in Hwitecelde**. These are the swine-pastures which we call in our language **denbera**, namely **Lamburnanden**, **Orricesden**, **Tilden**, **Stanehtanden**, and the wood called **Stanhurst** which belongs **to Wassingwell**. And I have willingly granted this privilege to this same piece of land **at Wassingwell** and likewise to the said **Wulflaf**, with the consent and permission of my chief men, that it may remain free and immune from all royal tribute, and services exacted by force and penal matters, from the domination of the ealdorman and the capturing of a thief and every secular burden, except military service only, and the building of bridges and fortification of fortresses. ...

These are the meadows **belonging to Wassingwell: half Stocmead, north of Hegford, by Stour mead, as belong thereto**.']

By the 10th century, the monolingual Old English boundary clause had become a typical feature of charters. Thus we see a diachronic development from the monolingual Latin charter to a charter showing switching within a subtext of the charter, namely the boundary clause, and finally to charters showing switching between the Latin main text of the charter and the fully vernacular boundary clause. As mentioned above, there is, however, a certain amount of overlap in this development, a fact which is also attested in the diachronic development of other text types in the Middle English period, such as letters (see Schendl 2002b) and sermons (see Wenzel 1994). Furthermore, other elements of the charter may also be switched to a certain degree, which is again compatible with current code-switching research.

Example (8), a charter by King Æthelstan from the early 10th century, is a good illustration of a vernacular boundary clause embedded in a Latin text. After the Latin *Hii sunt termini huius prefate terre*, 'These are the boundaries of the aforesaid land', the text switches into Old English. (A further vernacular sentence in this charter is the summarising endorsement '*Dis sind þa land þe Æethelstan ...*') Thus the two languages are rather consistently separated, with a clear functional motivation for the switches. The land boundaries as a central piece of information in the charter are described in the language in which these landmarks and boundaries were locally known; this obviously facilitated their clear identification – an important aspect of such a legal document (see also the discussion of place names and PPs above). A similar reason may apply to the summarising endorsement. This clear

discourse function of the vernacular was evidently a deliberate, conventional choice of the authorities issuing these charters at that time. This is supported by the frequent use of different scripts for the two languages mentioned in the introduction.

(8) *King Æthelstan confirms five hides at Chalgrave and Tebworth, Bedfordshire, to Ealdred, minister, A.D. 926* (BL Cotton Claud. B. vi, 23v-24r: copy s. xiii; Kelly 2000: 88f.)

In nomine Domini nostri Iesu Christi. ... Quapropter ego Æthelstanus Angulsaxonum rex non modica infulatus, sublimatus dignitate, superno instigatus desiderio, fideli meo Ealdredo ministro terram que nuncupatur **Cealhgræfan** et **Teobban wyrþe** .v. manentium quam propria condignaque pecunia, id est .x. libras inter aurum et argentum, a paganis emerat, iubente Eadwardo rege necnon et duce Æþeredo, cum ceteris comitibus atque ministris, in iuris hereditarii libertatem concedens donabo, habendam possidendamque quamdiu uixerit et post obitum suum quibuscumque sibi placitis heredibus dare uoluerit. Hii sunt termini huius prefate terre.

Ðær se dic sceot in Wæclinga stræte, anlanges Wæxlinga stræte ðæt on ðane ford, þæt anlang broces in þanne oðerne ford, þonne of ðæm forde up on þane welle, 7 þanan in ðæt dell, þanan of ðæm delle in ðone dic, of ðæm dice in ðone oþerne dic, þone of ðæm dice in þone broc, þonne of þam broce to Cynburge wellan, þanne anlang dices to east coten, þæt þanan in þane ealdan broc, up of þæm ealdan broce on æfem ðæt riþig, þæt up rihte in ðiod weg, æftær ðiod wege in þone dic, æfter dice in Wæxlingga strate.

Si autem hec prefata terre donatio ab omni seculari honore libera, preter expeditionem et arcis pontisue construccionem, pro competenti pecuni quam ego accepi, id est .cl. mancas de puro auro. Si quis uero hanc largitionis munificentiam, arrepto procacitatis stimulo, infringere uel mutare aut minuere temptauerit, sciat se in illa magni examinis die cum poli cardines terreque fundamenta simul et infernorum ima pauitando contremescent latibula, qua uniuscuisque patebit opus et conscientiam siue bonum siue malum quod gesserit, si non prius satisfaccione emendauerit. Anni ab incarnatione Domini nostri Iesu Christi .dccccxxvi., indiccione .xiiii.

Þis sind þa land þe Æthelstan cyng gebocade Ealdred wið his clæne feo on ðas gewitnesse þe her on sind.

+ Ego Æþelstanus rex fundamine sancte crucis subarraui [List of 22 witnesses]

['In the name of our Lord Jesus Christ. ... Therefore I, Athelstan, king of the Anglo-Saxons, adorned and elevated with no small dignity, prompted by desire from on high, will grant to my faithful thegn Ealdred the land of five hides which is called **Chalgrave** and **Tebworth**, which he bought with sufficient money of his own, namely ten pounds of gold

and silver, from the pagans by the order of King Edward and also of Ealdorman Ethelred along with the other ealdormen and thegns ; conceding with it the freedom of hereditary right, to have and possess as long as he lives, and to give after his death to whatever heirs, acceptable to himself, he shall wish.

These are the boundaries of the aforesaid land: **Where the dyke runs into Watling Street, along Watling Street to the ford, then along the brook to the other ford, then from that ford up to the spring, and thence into the valley, thence from the valley to the dyke, from the dyke to the second dyke, then from that dyke to the brook, then from the brook to Kimberwell, then along the dyke to Eastcote, then thence to the old brook, up from the old brook parallel with the little stream, then straight up to the highway, along the highway to the dyke, along the dyke to Watling Street.**

And the donation of the aforesaid land is to be free from every secular burden except military service and the construction of bridges and fortresses, in return for an adequate sum of money which I have received from him, i.e. 150 mancuses of pure gold.

If anyone, indeed, incited by impudence, shall try to infringe or change or diminish this generous munificence, let him know that, on the day of the great Judgment, when the hinges of the pole and the foundation of the earth as well as the deepest dens of hell shall quake and tremble, on which each shall reveal his work and conscience, what he did, good or ill, if he have not previously made emends with compensation.

In the year of the incarnation of our Lord Jesus Christ 926, the fourteen indiction.

These are the estates which King Athelstan granted by charter to Ealdred in return for his pure money, in the witness which is herein.

+ I, King Athelstan, have inscribed below with the sign of the Holy Cross. [Translation Whitelock 1955 : 503f.]

In (8), the introductory sentence before the actual boundaries is in Latin (*Hii sunt termini huius prefate terre*), though it may also be in the vernacular, having a form like *Ðis sindon þa land gemæru* 'These are the land boundaries'. In a few cases, the statement is first given in Latin and then paraphrased in the vernacular, as in (9):

(9) Istis terminibus predictum cingitur nemus. **Þis syndan ðæs dennes landgemæru to hæsel ersc**.

['The aforesaid grove is surrounded by these bounds. **These are the forest boundaries of Hazelhurst**.']

This repetition may be due to the fact that land boundaries sometimes seem to have been inserted into the Latin text only later, possibly on the basis of notes written by local scribes (cf. Kitson 1995: 48); though this is an

interesting aspect of the actual *process* of texts production, it does not affect the *result*, namely the fact that for the readers of such charters they show code-switching in the same way as texts produced in one continuous writing process.

3. Conclusion

The above discussion necessarily had to remain sketchy and there are quite a number of additional points relating to the question of code-choice and code-switching in Old English charters which will have to be the subject of further research. However, even this brief survey has clearly shown that code-switching is quite frequent in Anglo-Saxon legal texts. In these texts, code-choice and code-switching were evidently influenced by various extralinguistic variables, such as the time of text production, the particular type of document, as well as the structural elements of the text. In regard to the latter, there seems to have been a tendency to prefer Latin in the formulaic parts of the documents, though this is by no means obligatory. Furthermore, switching seems to have been more frequent and extensive in predominantly Latin texts, while texts in which the base language is the vernacular tend to show switching less frequently and often restrict it to certain formulaic phrases, especially at the beginning or end of the document. These results are, however, only tentative and have to be supported by further research in progress on a larger, ideally the complete corpus of Old English charters.

The sociolinguistic situation and the number of languages used in England was certainly quite different before and after the Conquest, but there was an obvious continuity of code-switching in administrative texts from the Old to the Middle English period, a fact generally overlooked in current research, which has so far mainly focused on the Middle English period. Finally, the number of text types showing switching was much higher in the Middle English period, while, as far as I know, only two types used switching in the Old English period. Most importantly, the Old English documents discussed in the present paper prove that we should antedate the use of code-switching in early written English by about five centuries.

References

Barker, E. 1947. "Sussex Anglo-Saxon charters, part I". *Sussex Archaeological Collections* 86: 42-101.
Birch, W. de Gray. 1885-1899. *Cartularium Saxonicum*. 3 vols. and index. London.

Chaplais, Pierre. 1965. "The origins and authenticity of the royal Anglo-Saxon diploma". *Journal of the Society of Archivists 3*(2): 59-61.

Clanchy, M. T. 1993. *From memory to written record. England 1066-1307*. 2nd ed. Cambridge, Ma. and Oxford: Blackwell.

Harmer, F.E. (ed.) 1952. *Anglo-Saxon writs*. Manchester: Manchester University Press.

Hooke, Della. 1990. *Worcestershire Anglo-Saxon charter-bounds*. [Studies in Anglo-Saxon History II] Woodbridge: The Boydell Press.

Howlett, D. 1997. "A polyglot glossary of the twelfth century". In: Gregory, St. and D.A. Trotter (eds.), *De mot en mot. Essays in honour of William Rothwell*. Cardiff: University of Wales Press, 81-91.

Jenkyns, J. 1999. "Charter bounds". In: Lapidge *et al*., 97-99.

Kelly, S.E. (ed.) 2000. *Charters of Abingdon Abbey*. Part 1. (Anglo-Saxon Charters VII) Oxford: Oxford University Press.

Keynes, S. 1999. "Charters and writs". In: Lapidge *et al*., 99-100.

Kitson, P. 1995. "The nature of Old English dialect distributions, mainly as exhibited in charter boundaries. Part I: Vocabulary". In: Fisiak, J. (ed.), *Medieval dialectology*, Berlin: de Gruyter, 43-135.

Lapidge, M. *et al*. 1999. *The Blackwell encyclopedia of Anglo-Saxon England*. Oxford: Blackwell.

Muysken, P. 1995. "Code-switching and grammatical theory": In: Milroy, L. & P. Muysken (eds.), *One speaker, two languages. Cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press, 177-198.

Poplack, S. 1980. "Sometimes I'll start a sentence in English y termino en espanol: toward a typology of code-switching". *Linguistics* 18: 581-618.

Robertson, A. J. (ed.) and transl. 1956 [1939]. *Anglo-Saxon charters*. 2nd ed. [Cambridge Studies in English Legal History] Cambridge: Cambridge University Press.

Romaine, S. 1995. *Bilingualism*. 2nd ed. Oxford: Blackwell.

Sawyer, P.H. 1968. *Anglo-Saxon charters: An annotated list and bibliography*. London: Offices of the Royal Historical Society.

Schendl, H. 1996. "Text types and code-switching in medieval and Early Modern English". *Vienna English Working PaperS* (*VIEWS*) 5: 50-62.

Schendl, H. 1997. "To London fro Kent / Sunt predia depopulantes': Code-switching and medieval English macaronic poems". *Vienna English Working PaperS* (*VIEWS*) 6: 52-66.

Schendl, H. 2000a. "Linguistic aspects of code-switching in medieval English texts". In: Trotter, D. (ed.), *Multilingualism in Later Medieval Britain*. Cambridge: D.S. Brewer, 77-92.

Schendl, H. 2000b. "Syntactic constraints on code-switching in medieval texts". In: Taavitsainen, I., T. Nevalainen, P. Pahta, and M. Rissanen (eds.), *Placing Middle English in Context*, Berlin: de Gruyter, 67-86.

Schendl, H. 2001. "Code-switching in medieval English poetry". In: Kastovsky, D. and A. Mettinger (eds.), *Language contact in the history of English*, Frankfurt/M.: Lang, 305-335.

Schendl, H. 2002a. "Mixed-language texts as data and evidence in English historical linguistics". In: Minkova, D. and R. Stockwell (eds.), *Studies in the history of the English language. A millenial perspective*. Berlin: de Gruyter, 51-78.

Schendl, H. 2002b. "Code-choice and code-switching in some early 15th century letters". In: P. Lucas and A. Lucas (eds.), *Middle English from tongue to text.* Frankfurt/M.: Lang, 247-262.

Schendl, H. 2004. "English historical code-switching in a European perspective". In: Dabelsteen, C.B. and J.N. Jørgensen (eds.), *Language and language practices* (Copenhagen Studies in Bilingualism, 36). Copenhagen: University of Copenhagen, 194-208.

Schendl, H. forthcoming. "Beyond boundaries: Language choice and code-switching in the leases of Oswald of Worcester"

Stenton, F. M. 1955. *The Latin charters of the Anglo-Saxon period*. Oxford: Clarendon.

Wenzel, S. 1994. *Macaronic sermons. Bilingualism and preaching in late-medieval England*. Ann Arbor: University of Michigan Press.

Whitelock, D. 1955. *English historical documents c. 500-1042*. Oxford: Oxford University Press.

Language policies in Viennese schools: the roles English plays

*Ute Smit, Vienna**

1. Introduction

The profound political changes of the 1990s (e.g. Austria joining the European Union, the fall of the Iron Curtain) had their impact on language policies in Austria in general and Vienna in particular: European multilingualism became 'our' issue, too, and the countries in our geographical vicinity re-appeared as actual neighbours on our mind-maps. In this light, politicians called for more language learning in Austrian schools (cf. Simpson 2004), which allowed the Viennese school administration to start various initiatives (see below), some of which have been integrated into mainstream schooling by now.

As the main focus of these innovations in school-based language learning and use is on English, it is the purpose of this paper to consider how the roles English has been made to play fit in with the actual 'language-scapes' in Viennese schools and the local, national and European educational language policies. In other words, my aim here is to throw light on the discrepancies I detect between, firstly, the actual multilingualism found in Viennese schools, secondly, the vaguely formulated idea of multilingual European citizens persistently voiced as ultimate aim of language policies, and, thirdly, the actual language (learning) policies in Viennese schools which foster bilingualism in German and English. In order to reach this aim I will first give a brief overview of the Viennese language situation in general (see below), against which the past and present language policies in Viennese schools can be sketched and delineated as they stand in relation to European language policies, on the one hand, and, on the other, the background of local historical developments (see 2.). I will then describe the roles English plays in the state school system (see 3.), which will lead to a discussion of the partly competing underlying language planning tendencies of Viennese mainstream schooling

* Author's email for correspondence: ute.smit@univie.ac.at

and of the roles English plays in them (see 4.). In my concluding analysis (see 5.), I will focus on the diverse perceptions people have of the status English holds presently in our schools and what implications this might have for the presently so wide-spread bilingual language (learning) policy.

As the Austrian capital of many centuries, Vienna has always been a centre of migration, attracting speakers of many different languages. This has not changed in its recent history, which means that Vienna's population today is far from monocultural and monolingual (cf. e.g. Fritz ed. 2001). The last census that included a question on language(s) of communication (see Table 1) revealed a section of about 20% of the population that uses other languages than German in their daily lives. The numerically strongest of these language minorities are the L1 speakers of Bosnian/Croatian/Serbian (BCS) and Turkish. This is not surprising considering the two numerically strong migrations from the respective countries of origin: the first was the economically motivated one in the 1960s till '80s – the so-called 'guest-workers' – and the second took place in the 1990s because of the war in former Yugoslavia. The remaining groups fall into those of Western and Eastern European origins and reflect the geo-political situation of Vienna as, on the one hand, city of international organisations, which is, on the other hand, situated in the geographical vicinity of former Eastern Europe. This explains the relatively numerous group of English speakers as well as the strong representation of Slavic languages.

Table 1. Languages of communication in Vienna; self-reported (cf. Statistik Austria 2001)

census 1991	population	with Austrian nationality
German	1 288 430	1 269 417
autochthonous languages* (+ German)	39 989	22 604
B/C/S** (+ German)	80 973	11 285
Turkish (+ German)	45 648	3 811
Englisch (+ German)	17 340	10 516
Polish (+ German)	16 709	6 509
Slovenian (+ German)	3 172	1 825
other (+ German)	48 287	17 228
total	1 539 848	1 343 196

* = The languages of the legally acknowledged indigenous minorities; i.e. Croatian, Czech, Hungarian, Slovak, Romanes

** = Bosnian/Croatian/Serbian; referred to as 'Serbo-Croatian' at the time of the census

2. Educational language policies

In view of the fact that language policies are always embedded in cultural and therefore historical processes, I will give a brief overview of the language policies of the last 150 years before I turn to the present situation.

2.1. Historical overview

While the Austrian-Hungarian monarchy was an, also officially declared, multilingual state, 19th century nationalism soon saw to it that the territorially defined language rights of the respective ethnic groups were strictly applied (cf. Rindler-Schjerve ed. 2003; Wascher 1999). With regard to education this meant that the language of instruction of a certain school was the same as the official language of the respective area. Bilingual schooling was thus made impossible, as were schools catering for language minorities in other areas of the state. Vienna, as located in the German-speaking part, could thus only host German-medium schools. This stood in an extremely stark contrast to the population of that time: In 1880, for instance, 65% of all people living in Vienna were migrants of diverse language background and in the ensuing decade 200 000 labourers from Bohemia moved to Vienna and surroundings (cf. Burger 1995). But even this numerically strong minority did not manage to change the monolingually German educational language policy and had to found their own private schools, which did not, however, gain the right to administer publicly accepted final exams.

The fundamentally monolingual ideology of nationalism was again very relevant during the peace negotiations after the First World War (cf. Posch 1994). In the eyes of the Austrian as well as the other delegations, Austria was seen as German speaking. While the Peace Treaty of St. Germain included minority clauses, it left the term 'minority' open to interpretation, which paved the way for those in power to not only offer very little support for minorities, but to also oppressively pursue assimilatory policies. With regard to educational matters, Czech-medium public schools in Vienna are a case in point. While they became only possible due to a bilateral agreement with the Czech Republic, the Viennese authorities tried their utmost to make them unattractive to the generally bilingual clientele by permitting German neither as medium of instruction nor as subject (cf. Posch 1994).

A similar language planning approach was followed after the Second World War (cf. de Cillia 1998: 139-145; Wakounig 1994): the State Treaty of 1955, further specified in the 1976 Law on Ethnic Groups, includes a clause on minority rights, granting explicitly named autochthonous groups official language rights, such as basic education in their languages. But as these rights are still territorially restricted and based on the hotly debated numerical principle, the "minority question" still causes debate, especially with regard to bilingual place-name signs (cf. *Der Standard* 15 Dec. 2001 & 21 Oct. 2004).

The autochthonous groups historically present in Vienna are the Czechs and Hungarians. While both languages do not make any visible appearances in public life, they feature in specific schools: there is still one Czech school, which is by now run bilingually (cf. http://www.komensky.at/ for "Schulverein Komensky"), and Hungarian interest groups have managed to instigate bilingual schooling programs in a few Viennese primary schools. (cf. http://www.lehrerweb.at/ssr/europabuero/hungaricum.html for "Projekt Hungari-cum").

2.2. Present situation

2.2.1. Austrian vs. European educational language policies

Similar to most other European nations (cf. Gogolin 2002) and illustrated by the historical overview given above, Austrian educational language policies have been deeply steeped in a monolingual ideology for a long time. On a political level, this assimilatory approach has not been challenged fundamentally until now (cf. de Cillia 1995b & 2003): the school as a "monolingual habitus" (Gogolin 1994) educating all children in the country's national language is still the norm. Based on widely unquestioned, ideologically heavily loaded concepts of 'mother tongue' and 'cultural identity' (cf. Radtke 2002), multilingual children are usually pressed into a monolingual system, in which success rests mainly on language abilities in German. Other L1s are generally taken as irrelevant to educational success and belonging to the individual's private sphere. Thus classified 'monolingual' speakers of German are then expected to learn a foreign language, which is presently mainly English (see 3.1. below).

While this is basically still the frame of Viennese schools, the 1990s have brought some important and long lasting changes: On the political level, Austria joined the EU in 1996, and in the late 1990s Austria's neighbouring countries to the East (i.e. close to Vienna) became EU applicant countries and

recently member states. This started off an increased interest in, amongst other aspects, their languages (cf. e.g. Europäisches Jahr der Sprachen 2001; Haidinger 2002). With regard to language (learning) policies, the White Paper on *Teaching and Learning: Towards the Learning Society* (Cresson 1996) spelt out the aims of the European Commission. These are, amongst others, "the promotion of community language learning" (p. 32), which can be understood as a call for every European to gain "proficiency in several Community languages" (Cresson 1998: 2).

The White Paper thus formulated as a political agenda what language learning research and teaching initiatives had been pointing to for some time already (cf. Morgan 2001): the European multilingual nature, accepted and respected in the EU plurilingual language policy, must also become the basic principle of educational language policies. As Rudolf de Cillia (1995a: 38) puts it

> [d]as wichtigste Prinzip einer europäischen Sprachenpolitik ist das der Diversifizierung, der sprachlichen Pluralität, d.h. dass möglichst alle großen europäischen Sprachen unterrichtet und gelernt werden sollen [the most important principle of a European language policy is the one of diversification, of linguistic plurality, which means that as many major European languages as possible should be taught and learnt; my translation].

When translated into language learning practice this means that schools should prepare their pupils to function in a variety of multilingual situations by teaching not only a broad range of languages, but also a range of language competencies with a high degree of applicability (cf. Christ 1998; Morgan 2001: 41). Added to this, a special focus should be placed on intercultural and -lingual communication skills and abilities (cp. Gogolin's (2002) call for "Interkulturelle sprachliche Bildung").

As will be discussed in more detail below, some of these ideas, such as intercultural education or language proficiency in more than one language, have found their way into mainstream curricula, while others, like true multilingualism, have not made any impact on Viennese language (learning) policies so far.

2.2.2. Viennese vs. Austrian educational language policies

In view of the fact that Vienna has been a meeting point for many different language groups throughout its history, it does not come as a surprise that the proportion of pupils with L1s other than German is considerably higher than in the rest of Austria (see Table 2). While this is true of all school grades and types, the percentages range from just about 20% in grammar schools and

vocational colleges to more than double that in primary and secondary modern ones. This is symptomatic of the Austrian educational system, in which the upper secondary levels are, in practice, only frequented by those GSL (German as a second language) pupils whose German is native-like (cf. de Cillia 1995b).

Table 2: GSL speaking pupils (in primary and secondary education) in 2002/03 (cf. BMBWK. 2004)[8]

type of school	Vienna		Austria
	no	%	%
primary (grades 1-4)	26 061	40,9	16,4
secondary modern (grades 5-10)	19 595	43,0	13,8
special needs (grades 1-9)	1 334	36,5	22,0
grammar (grades 5-12)	10 202	19,2	7,2
vocational college (grades 9-13)	4 523	19,5	7,6

As the Viennese school authorities ('Wiener Stadtschulrat') have gained a lot of experience in dealing with pupils from very different language backgrounds, their language policies are somewhat different from the standard Austrian ones (cf. Weidinger Walter 2000 & 2001). In brief, Viennese schools offer an integrative learning environment to all pupils in so far as all of them are placed into mainstream classrooms irrespective of their German language proficiency levels. Whenever feasible, special German tuition is offered in class by specifically trained teachers who work in teams with the respective class teachers. The German language learning program is based on a 6-year GSL curriculum, which all GSL-pupils follow according to their individual language learning needs (cf. Fleck 2001 for a more detailed account). While the GSL program is well thought-through in theory, it is not always realised to its full extent. Due to various adverse circumstances, a good many of the GSL pupils do not get their full share of GSL tuition, while

---

8   Interestingly enough, in the official statistics given out by the Austrian Ministry of Education 'GSL' does not include all pupils with an L1 other than German, but only those whose school career in Austria has been shorter than or equal to 6 years. Once pupils start their 7th year of schooling in Austria, they are statistically taken out of the 'GSL' group and put into the German L1 majority group. While the relevant documents do not give any explanation for this procedure, I presume that this period of 6 years is linked to the GSL curriculum, which is designed for six years as well.

others do not have enough time to gain the level of proficiency in German they would need for more advanced education. (For more detailed criticism of the present situation cf. de Cillia 1995a & 1998.)

In addition to extra German lessons, GSL-pupils have a right to L1 instruction, which is, however, much less intensive than GSL training and usually offered in separate lessons. That these L1-lessons are considered less important is also apparent (a) in their marginal status within the curriculum, (b) in the teachers' temporary status of employment (1-year contracts only) (cf. Pinterits 2002), and (c) by the fact that very few pupils insist on their right once they reach upper secondary level; in 2000, for instance, only 1% of all GSL pupils took advantage of the offer (cf. Waldrauch 2001: 18-22). It thus seems that the higher up the pupils go within the system, the less they feel like developing their L1s. Summing up these regulations, one can say that mainstream schools in Vienna try to integrate and not merely assimilate GSL-speaking pupils, especially in the lower classes, but offer transitional bilingualism in the sense that pupils are expected to become academically fully proficient in German only.[9]

Besides this integratively introduced transitional bilingualism leading to German, Viennese pupils are also expected to learn foreign languages. In recent years, the afore-sketched European language policy is being cited more and more and schools try to give themselves multi- or, at least, bilingual descriptions (Simpson 2004). Compulsory foreign language (FL) learning now starts in grade 1 and continues up to school-leaving age (cf. Österreichisches Sprachen-Kompetenz-Zentrum 2001, for an overview). Added to this, a number of innovations have been started to offer intensified FL learning, the use of FL as medium of instruction and special classes on European Studies (on primary and secondary levels) (see 3.2).

Based on this increase in multilingual programs, the Viennese school authorities have called Vienna "the biggest language school in Europe" (S. Simpson, 25 June 2004, personal conversation). While these initiatives are definitely laudable, the kind of 'multilingualism' offered should be looked at in more detail, especially because of the overwhelming dominance of English.

---

[9] Transitional bilingual education, as Baker (1996: 173) defines it, "aims to shift the child from the home, minority language to the dominant, majority language."

3. The roles English plays

In contrast to Austrian school legislation, which does not specify which foreign languages should be taught as subjects or used as medium of instruction (cf. BMBWK 2003), reality shows that it is English which has developed into the default option.

3.1. English as FL

As the overview of foreign-language learning regulations (Table 3) shows, all Austrian pupils learn at least one FL during their nine compulsory school years, and one more if they continue up to "Matura" (school leaving exam).

Table 3. Compulsory FL-learning and FLs chosen (percentages from 1991/92) (cf. Österreichisches Sprachen-Kompetenz-Zentrum 2001)

primary level (grades 1-4)	introduction to 1 FL; grade 4 pupils: approx. 98% English
lower secondary (grades 5-8)	1 compulsory FL; from gr. 7: ~50% of grammar schools: 2 FLs grade 8 pupils: approx. 75% English only
upper secondary (grades 9-12/13)	1 further FL grade 12 pupils: approx. 50% English only

While the law does not favour any language, the percentages in Table 3, even if more than ten years old, illustrate quite clearly the preponderance of English. Even after the sociopolitical changes of the 1990s with the ensuing re-evaluation of the relevance of neighbouring languages (see above), the overall numbers and percentages do not reflect any considerable change in this regard so far: in 1998, for instance, only 0.06% of pupils took Hungarian and 0.025% Czech as FLs (cf. Krumm 1999: 12)

In view of this highly persistent, one-sided distribution, heavily weighted towards English as only sizeable FL in Austria (and Vienna), Rudolf de Cillia's repeatedly voiced point of criticism (e.g. 1997, 1998 & 2003) is still highly valid: the present setup of learning only one FL from grades 1 to 9 (or 12) practically negates the European plurilingual language policy. With so many teaching sessions reserved for English only, the likelihood that other FLs will get some sizeable share is close to nil. In other words, the present FL learning system supports a "Leitsprachenpolitik" (i.e. policy of a leading language), positioning English (as global language and lingua franca) as

apparently only relevant FL to learn (cf. also Krumm and Portmann-Tselikas 1999).

The present situation gains in diversity when one turns to new teaching initiatives, which fall into two groups: either they belong to the broad category of 'CLIL' (content and language integrated learning) and aim for the use of another language but German as medium of instruction (see 3.2.); or they go beyond the generally found 'monolingual habitus' and offer bilingual schooling (see 3.3.).

3.2. English as 'L' (language) in CLIL (content and language integrated learning)

A legal provision allowing schools to choose their mediums of instruction has made it possible for languages other than German to be used for this purpose (cf. BMBWK 2003). Since the early 1990s and the start of political support for the internationalisation of the Austrian schools, this provision has been used again and again to allow the implementation of content-and-language-integrated learning in Austria (c.f. e.g. Abuja ed. 1998). Under the label of 'FsAA' ('Fremdsprache als Arbeitssprache', i.e. FL as working language), non-language subjects, as, for instance, geography, history or business economics, can now be taught in English, French or Italian. This teaching approach has not only benefited from the support of the Ministry of Education and its institutions (e.g. Zentrum für Schulentwicklung in Graz), but has also become highly popular in the last 10 years. An accurate description of its numeric distribution is, unfortunately, not available as the most recent figures date from 1996/97, when a questionnaire was handed out to the school-heads of a sample of lower secondary, grammar and vocationally-oriented schools (cf. Gierlinger 2002; Kettemann et al. 1997). At that stage, about a third of the upper secondary schools used 'FsAA', and less than 10% at lower secondary level.[10] Added to this, most of the schools reported on the generally positive experience they had with this teaching method.

Since then, more and more schools have introduced this teaching methodology, for various reasons. 'FsAA' is bureaucratically very flexible, it can be used for complete subjects, or only parts thereof, for a whole year or

---

[10] These percentages have increased rapidly since then: As part of an M.A. research project, the same questionnaire was handed out again to all 48 Lower Austrian secondary schools in January 2000 (cf. Zeller 2000: 69-70). While 6 schools indicated they used 'FsAA' in 1996, the number had risen to 18 in 2000.

only a few sessions. In theory, it is similarly flexible with regard to languages used, but, as already indicated by the abbreviation generally employed – 'EAA' instead of 'FsAA' – most schools and teachers stick to one language only, and that is English. This is in part due to the participants' language proficiency; most pupils and teachers would simply not be competent enough in any other FL. However, attitudinal factors might also be of relevance here as most teachers, parents and pupils seem to single out English for this role (Abuja ed. 1998; Kettemann et al 1997). One can speculate about the reasons; it might be because of the preponderance of English in what people associate with international communication, but it might also be a result of the present school system with its primary and partly sole focus on English (see 3.1. above).[11]

Next to 'FsAA' as a teaching methodology in mainstream schools, a few more recent initiatives have been started that use FLs as mediums of instruction. The best-known among them are the 'European Schools', covering all 12 school years (Simpson 2004). They started as an EU-supported cross-border project, catering for Austrian, Czech, Slovak and Hungarian pupils, who would follow a synchronised curriculum accepted in the respective countries. At lower secondary level, the pupils would attend school in Vienna, follow most of their academic subjects in their respective L1s, the more practical ones in German and a new academic subject called 'European Studies' in English. In addition, they can choose each others' L1s as FL subjects.

Along similar lines, the Viennese authorities started a schooling initiative for mainstream Viennese children (Simpson 2004), aiming at international or global education. It begins at primary level (Global Education Primary School), and continues at lower and upper secondary levels (Modern Language Secondary School and Junior High School for Science, Arts and IT).[12] At all levels, pupils have more FL lessons and a clear focus on English,

---

[11] Despite its increasing popularity, EAA has not received the research interest it would deserve. Questions such as its consequences on the pupils' language proficiency, the methodology used in teaching through English (Abuja ed. 1998) but also more fundamentally, what kind of language use EAA classes offer still await intensive treatment. At least with regard to the latter question, a detailed study is underway (cf. Dalton 2002).

[12] It should also be mentioned that next to the much more numerous English programs, similar ones are run for French and Italian, but only at one school each (cf. Simpson 2004).

which is also used as medium of instruction for a subject specifically developed for this school, namely 'Global Studies'.[13]

As these initiatives are fairly unique in Europe, the Viennese authorities receive regular visitors from similar institutions from other cities (S. Simpson 25 June 2004, personal conversation). While the people responsible for them are rightly proud of their work, especially so because they are non-fee paying state schools, which are situated all over Vienna, these initiatives reveal the same lop-sided focus on English as the apparently only relevant FL to be learnt and used. It seems still far from the European ideas of multilingualism, of language proficiency in at least 3 community languages for every European and of diversification in the FL offer at schools.

3.3. English as medium of instruction in bilingual programs

As mentioned above, Viennese schools are monolingual, with German as default medium of instruction. But, as so often, there is an exception to the rule: the VBS (Vienna Bilingual Schooling), which is a truly bilingual schooling program. It offers education in German and English from grades 1-12, following the Austrian curricula (for a more detailed description cf. Weidinger, Wiltrud 2001). It commenced in 1992 as a response to the growing English-speaking community in the city. The VBS follows the two-way bilingual model (cf. Baker 1996: 186) of having two groups with different L1s, i.e. German and English, in the same class with the aims to (a) familiarise each group with the respective L2 and (b) gain literacy in both of them. For this purpose, VBS classes have German and English speaking teachers who teach some subjects in teams and some in the pupils' L1s only, decreasingly so the older the students are.

The VBS administration considers it a very successful program for two reasons: Firstly, more and more schools start a VBS strand[14] and, secondly, accompanying research has shown that pupils' proficiency levels in English surpass the ones of Austrian pupils in monolingual programs (cf. Hüttner 1997: 88-99 for a detailed account). Next to language proficiency, the VBS is also intended to support intercultural learning, because its design rests on the assumption of two lingual-cultural groups meeting in the classroom. Here, the

---

[13] The materials developed for this subject are freely downloadable at http://www.cernet.at.

[14] The VBS runs in one class per form of, at the moment, 7 primary, 8 lower secondary and 6 upper secondary schools (Simpson 2004).

question is though, which cultures are being considered. The German-speaking groups attract German-speaking Viennese children and are therefore relatively homogeneous. The English-speaking groups, on the other hand, draw their pupils from the group of approximately 20 000 English speakers in Vienna (cp. Table 1).[15] That this is a much more heterogeneous conglomerate becomes already apparent when considering a few basic factors: many English speakers come to Vienna because of work commitments; they often stay for limited time periods; they represent many different nationalities and are often bi- or multilingual in various languages. For VBS this means that their English-speaking groups are multilingual (often including some German, too) and definitely multicultural in themselves. In other words, they are much less tightly knit than the German-speaking ones. This fact is, however, not explicitly recognised; neither in the VBS brochures nor, as Hüttner (1997) describes in her qualitative research, in actual school life where GSL pupils were, willingly or unwillingly, pressed into a homogeneous English role. This might have been happening so as to keep up the notion of biculturalism in support of the VBS concept as representing two-way bilingual education.

4. Discussion

The language policies found in Viennese mainstream schools at present in my view reflect the complex interrelationships of three different language planning approaches. There is, firstly, the age-old and ideologically highly entrenched approach of assimilation, which has always seen Viennese schools as culturally homogeneous and monolingually German. Quite clearly, this was the one and only approach during and after the monarchy right into the 1960s. But it did not stop abruptly then, either; in fact, it is still 'simmering on'. A case in point is the afore-mentioned systemic distinction between compulsory and academically-oriented education with regard to their degree of linguistic diversity, where the latter is conspicuously less open to GSL pupils. Similar traces can also be found in recent publications, even when pointing out the relevance of multiculturalism in Viennese schools. For

---

[15] This numerical imbalance in group-size between English- and German-L1 speakers is also reflected in VBS classes: In contrast to the assumed 50:50 split in VBS classes, Gräll (1999: 96), for instance, observed a VBS class with more than twice as many German-speaking than English-speaking pupils. One reason for this imbalance could simply be that there are too few English L1 speaking children in Vienna to fill the many VBS classes (cp. figures in Table 1).

example, in a brochure describing the multilingual mix of Viennese pupils (Weidinger, Walter 2000: 7), all of whom, it is stressed, should have equal opportunities to receive proper education, the term "Wiener Schüler" (Viennese pupils) is used ambiguously, partly referring to German L1 speakers only. This, I would claim, indicates that monolingualism in the form of German as language of Viennese pupils is, at least implicitly, still assumed as the norm.

The second language planning approach that co-shapes Viennese language policies is integration. Even if it is a much more recent phenomenon – its beginnings date back maybe 30 years – it is highly visible as most official documents and brochures stress integration as main aim at Viennese schools (e.g. Gröpel 2001: 231 f.; Weidinger, Walter 2001). But it is not only apparent on paper, it has also made its mark in actual school life (cf. BMBWK 2003; Olechowski et al. 2002), such as in integrating GSL speaking pupils right from the start and offering them extra German input as part of normal classes. Similarly, all GSL students have a right to classes in their L1s with the idea that their L1 competence should develop continuously. Another case in point is the recent initiative to offer bilingual literacy programs in Turkish and German in a few primary school classes with a high percentage of Turkish L1 speakers (cf. Weidinger, Wiltrud 2001: 131). As repeatedly stated in policy documents, however, the main aim of integration is not linguistic, but cultural. Since 1991 "interkulturelles Lernen" (intercultural learning) has been a general principle of teaching and learning, with the aim to contribute

> einen Beitrag zum besseren gegenseitigen Verständnis bzw. zur besseren gegenseitigen Wertschätzung, zum Erkennen von Gemeinsamkeiten und zum Abbau von Vorurteilen [to an improved mutual understanding and respect and to an increased dismantling of prejudices; my translation] (BMBWK 2003: 21)

This means that the linguistic aspects of integration, i.e. mutual respect for L1s and support to develop one's L1 according to age and educational level, have been seen as less crucial, and especially the second aspect – developing one's L1 – is therefore still partly up to the individuals to do themselves.

The third language planning approach is the most recent and presently very vibrant one of what I would like to call internationalisation. The idea to equip all Austrians with the linguistic means to reach out to other nations was first put into practice in the 1970s when foreign language learning became compulsory for all school types, and gathered momentum in the early 1990s when political and socio-economic forces actively supported it. With the political agenda of joining the European Union, on the one hand, and, on the other, the increased relevance of international companies, the people

responsible for educational language policies placed their prime focus on proficiency in foreign languages. This led to most of the new school initiatives described above (see 3.2. and 3.3). As all of these aim at individual and, finally also societal multilingualism, one is reminded of the EU's White Paper and its call for proficiency in at least three EU languages (see 2.).

But is this really the aim of the present Viennese language policies? After all, Viennese pupils are required to learn only one foreign language, more than that can be done on a voluntary basis only. The neighbouring languages, such as Czech, Slovak or Hungarian, seem to attract more and more interest, but again, they are on offer in relatively few schools only. Contrary to that, most of the new bi/multilingual school initiatives described above (i.e. CLIL, 'European' and bilingual schools) offer increased input in one language only, and that is English. It is the first foreign language most pupils study at primary level and remains that during the lower secondary years. If pupils continue their school career, they will again have English as first and partly only foreign language until their 'Matura' (school leaving exam). And if they take part in CLIL classes in-between, they can reckon with them being taught in English, too. For the few students who attend one of the special school types (e.g. VBS, EMS), it will again be English with which they will most likely get into contact. It is because of this overwhelming preponderance of English as first and foremost foreign language, that it seems unrealistic to me to expect true multilingualism as outcome of the present language policy situation. Instead, I would argue, the aim seems to be societal bilingualism in German and English. In this conception German is presumably seen as national language with all the relevant functions (cf. Cooper 1989, ch. 5) and English as international one, which fits in well with what Wright (2004: 172) describes as the "need [we feel] for a global language".

To sum up, it thus seems fair to interpret the present educational language policies as aiming at what could be called 'globalised bilingualism' for the society on the whole, while keeping true multilingualism for individuals to pursue and explore. This is possible because of the interplay of assimilation, which forces every pupil to gain their highest proficiency level in German, integration, which recognises the cultural diversity of the population and allows individual multilingualism, and internationalisation, which fosters proficiency in English throughout all levels of society.

5. Conclusion

As mentioned repeatedly in this paper, the uniquely privileged status English enjoys in our school system has aroused criticism from applied linguists, for theoretical and pragmatic reasons (e.g. de Cillia 1995a, 1995b & 1998; Krumm 1995 & 1999). Theoretically, the presently advocated 'globalised bilingualism' squares badly with the European policy of multilingualism as it will not transform the coming Austrian generations into speakers of various, and diverse, community languages. More pragmatically, the large percentage of language (learning) lessons taken up by English gives fairly little time and leeway for other languages to be learnt to higher levels of proficiency. The practical realisation of the curricula thus makes it very difficult for the average Austrian pupil to attain (upper) intermediate proficiency in a second, or even third FL.

These, in my view, rightly voiced concerns seem to contrast quite sharply with the opinions and attitudes found amongst people other than applied linguists, very few of whom actually worry much about the present preponderance of English. On the contrary, the school authorities are proud of the present system; teachers – willingly or unwillingly – play along and offer more and more classes in 'EAA'; and many parents seem to favour an even earlier start of learning English and pay for extra English classes at pre-school institutions. But these positive evaluations are also predominant in circles outside education proper. In informal discussions with non-educators[16], I have met more who whole-heartedly support, rather than question, the basic thrust of 'globalised bilingualism' as they feel strongly that English should be privileged over all other FLs taken together. The reasons given are usually instrumental ones – English as the main language 'to make your way' while travelling, studying or working abroad, but also as main language of communication with non-Austrians in Austria or Europe, be it for business or pleasure. What is interesting, though, is that, when given the time to talk about this issue in more detail, many people readily acknowledge the relevance and value of being proficient in other languages as well. Everyone usually has a story to tell that illustrates how helpful or even crucial proficiency in languages other than German or English can be. While in the

---

[16] Here, I draw on impressions gained in conversations with friends and acquaintances, but also in a discussion round on "Europäische Sprachentwirrungen"(= on detangling the European languages), organised by the "Österreichisches Institut für Raumforschung" (Austrian Institute for Regional Studies and Spatial Planning) in Salzburg, 16 Oct. 2003.

past this used to concentrate on Western European languages only, our neighbouring languages to the East have gained ground by now, too.

So, what I would like to read out of these comments is that Austrians outside the educational/academic sector support 'globalised bilingualism' not to the exclusion of other FLs, but simply because they experience and see this as the demand of our times. In other words, proficiency in English has become a 'must' and is seen as only second in importance to basic literacy and numeracy[17], while proficiency in other languages has kept the value and status it has always had; people want to learn, or have their children learn, those languages they perceive as economically viable or socially attractive.

I would therefore like to argue that the gap between the rather negative applied linguistic and rather positive general assessment of 'globalised bilingualism' could be bridged by clarifying what 'English proficiency' and the learning process to get there actually entail. If applied linguists found a way to convince policy makers and recipients (in this case, educators, parents and pupils) that English does not need to play the roles of main FL, 'L' in CLIL and only medium of instruction next to German in the present manner and time-allotment for societal bilingualism to be achieved, curricular time and capacity would be free for other foreign languages; and our educational language policies could thus move closer towards true multilingualism.

References

Abuja, Gunther ed. 1998. *Englisch als Arbeitssprache: Modelle, Erfahrungen und Lehrerbildung*. (Berichte Reihe III, Nummer 4). Graz: Zentrum für Schulentwicklung III.

Baker, Colin. 1993. *Foundations of bilingual education and bilingualism*. Clevedon: Multilingual Matters.

Clevedon: Multilingual Matters.

BMBWK. 2003. "Gesetzliche Grundlagen schulischer Maßnahmen für SchülerInnen mit einer anderen Erstsprache als Deutsch. Gesetzt und Verordnungen. (= Informationsblätter des Referats für Interkulturelles Lernen Nr. 1/2003).

BMBWK. 2004. "SchülerInnen mit einer anderen Erstsprache als Deutsch. Statistische Übersicht 1993/94-2002/2003." (= Informationsblätter des Referats für Interkulturelles Lernen Nr. 2/2004).

Burger, Hannelore. 1995. *Sprachenrecht und Sprachengerechtigkeit im Österreichischen Unterrichtswesen 1867-1918*. Wien, ÖAW (Österreichische Akadmie der Wissenschaften).

---

[17] Cp. Grin's (2001) assessment of English proficiency as "norm and part of social habitus in the broad sense" (p. 75), similar to what literacy used to be in the early 20th century.

Christ, Herbert. 1998. Bildungspolitik für Mehrsprachigkeit. In Ingrid Gogolin, Sabine Graap and Gunther List eds. *Über Mehrsprachigkeit*. Tübingen: Stauffenberg Verlag, 337-360.

Cooper, Robert L. 1989. *Language Planning and Social Change*. Cambridge: Cambridge University Press.

Cresson, Edith. 1996. *Teaching and Learning: Towards the Learning Society*. Strasbourg: European Commission.

Cresson, Edith. 1998. *Multilingualism in Vocational Training*. Brussels: European Commission.

Dalton, Christiane. 2002. Content and language integrated learning in Austrian classrooms: applied linguistics takes a look. *VIEWS (Vienna English Working Papers)* 11/1&2: 4-26.

de Cillia, Rudolf, Gero Fischer and Grete Anzengruber eds. 1997. *Lehren und Lernen fremder Sprachen in Österreich*. Schulheft 88. Wien: REMAprint.

de Cillia, Rudolf. 1995a. Österreichs Schulsprachenpolitik – eine "Sprachenpolitik für das Jahr 2000?. In Ruth Wodak and Rudolf de Cillia eds. *Sprachenpolitik in Mittel- und Osteuropa*. Wien, Passagen Verlag, 37-51.

de Cillia, Rudolf. 1995b. Höhere Schulen – ausländerfrei? SchülerInnen mit nichtdeutscher Muttersprache an höheren Schulen. In Ruth Wodak and Rudolf de Cillia eds. *Sprachenpolitik in Mittel- und Osteuropa*. Wien, Passagen Verlag, 253-263.

de Cillia, Rudolf. 1997. Bilingualer Unterricht und Fremdsprachenunterricht. In Rudolf de Cillia, Gero Fischer and Grete Anzengruber eds. *Lehren und Lernen fremder Sprachen in Österreich*. Schulheft 88. Wien: REMAprint, 67-79.

de Cillia, Rudolf. 1998. *Burenwurscht bleibt Burenwurscht. Sprachenpolitik und gesellschaftliche Mehrsprachigkeit in Österreich*. Klagenfurt/Celovec: Drava Verlag.

de Cillia, Rudolf. 2003. Tendenzen und Prinizipien europäischer Sprachenpolitik. In Hans-Jürgen Krumm ed. *Sprachenvielfalt. Babylonische Sprachverwirrungen oder Mehrsprachigkeit als Chance? Ringvorlesung der Geistes- und Kulturwissenschaftlichen Fakultät der Universität Wien*. Innsbruck: StudienVerlag, 27-40.

Der Standard. 2001. Ortstafelstreit wieder voll entbrannt. (printed version from 15 Dec. 2001). http://derstandard.at/?id=807153 (27 September 2004).

Der Standard. 2004. Kärntner Ortstafeln: Vorerst keine weitere Konsenskonferenz (printed verson from 21 Oct. 2004) http://derstandard.at/?id=1831636 (9 November 2004).

Europäisches Jahr der Sprachen 2001. BMBWK. http://www.bmbwk.gv.at/ministerium/04/Europaeisches_Jahr_der_S5673.xml. (27 September 2004).

Fleck, Elfie. 2002. Der mutterspachliche Unterricht. *Erziehung & Unterricht 9/10*, 1-13.

Fritz, Thomas ed. 2001. *280 Sprachen für Wien*. Wien: Edition Volkshochschule.

Gierlinger, Erwin. 2002. Fremdsprache als Arbeitssprache in Österreich: Situation und Perspektive. (Comenius 2.1. Projekt) htttp://www.pze.at/memo/ (14 July 2004)

Grin, François. 2001. English as economic value: facts and fallacies. *World Englishes* 20/1: 65-78.

Gogolin, Ingrid. 1994. *Der monolinguale Habitus der multilingualen Schule*. Münster/New York: Waxmann.

Gogolin, Ingrid. 2002. Interkulturelle sprachliche Bildung. In Walter Weidinger ed. *Bilingualität und Schule 2. Wissenschaftliche Befunde*. Wien: öbv&hpt, 98-107.

Gräll, Karin. 1999. *The Vienna Bilingual School-Middle School. The language use of bilingual school-starters. University of Vienna:* M.A. thesis.

Gröpel, Wolfgang. 2001. Kindheit, Migration und Schullaufbahn. In Walter Weidinger ed. *Bilingualität und Schule. Ausbildung, wissenschaftliche Perspektiven und empirische Befunde*. Wien: öbv&hpt, 162-245.

Haidinger, Martin. 2002. Positive Bilanz zum 'Jahr der Sprachen'. http://science.orf.at/science/news/35904 (27 September 2004)

Hüttner, Julia Isabel. 1997. Patterns of Code choice and Codeswitching in a Bilingual Primary School Setting. M.A. thesis, University of Vienna.

Kettemann, Bernhard, Isabel Landsiedler, Markus Kerschbaumer and Markus Cossée. 1997. Zur Situation des Fremdsprachenunterrichts in Österreich. Die Ergebnisse der *Language Education Study* der IEA. In Rudolf de Cillia, Gero Fischer and Grete Anzengruber eds. *Lehren und Lernen fremder Sprachen in Österreich*. Schulheft 88. Wien: REMAprint, 117-130.

Krumm, Hans-Jürgen and Paul R. Portmann-Tselikas. 1999. Sprachenpolitische Beiträge zu Deutsch als Fremdsprache. In Hans-Jürgen Krumm and Paul R. Portmann-Tselikas eds. *Theorie und Praxis. Österreichische Beiträge zu Deutsch als Fremdsprache. Schwerpunkt: Sprachenpolitik in Österreich*. Innsbruck/Vienna: Studien Verlag, 11-17.

Krumm, Hans-Jürgen. 1999. Nachbarsprachen – eine Herausforderung für die monolinguale Einstellung der Bildungssysteme und den traditionallen Fremsprachenunterricht. In Hans-Jürgen Krumm ed. *Die Sprachen unserer Nachbarn – Unserer Sprachen. Chancen zur Diversifizierung des Sprachenangebots im Zuge der EU-Erweiterung*. Vienna: eviva, 115-126.

Krumm. Hans-Jürgen. 1995. Das Erlernen einer zweiten oder dritten Fremdsprache im Rahmen von Mehrsprachigkeitskonzepten. In Ruth Wodak and Rudolf de Cillia eds. *Sprachpolitik in Mittel- und Osteuropa*. Vienna, Passagen Verlag, 195-208.

Morgan, Carol. 2001. Multilingualism and multilingual language learning. In Walter Weidinger ed. *Bilingualität und Schule. Ausbildung, wissenschaftliche Perspektiven und empirische Befunde*. Wien: öbchpt, 41-49.

Olechowski, Richard; Günter Hanisch, Tamara Katschnig, Gabriele Khan-Svik, Elisabeth Persy. 2002. In Walter Weidinger ed. *Bilingualität und Schule 2. Wissenschaftliche Befunde*. Wien: öbv&hpt, 8-65.

Österreichisches Sprachen-Kompetenz-Zentrum 2001. Fremdsprachenunterricht in Österreich (abbreviated version). Graz. http://www.sprachen.ac.at/links/pdf.php. (23 July 2004)

Pinterits, Manfred. 2002. Mehrsprachige Kinder an Wiener Schulen – die Vielfalt akzeptieren. In: Walter Weidinger ed. *Bilingualität und Schule. Ausbildung, wissenschaftliche Perspektiven und empirische Befunde*. Wien: öbv&hpt, 22-29.

Posch, Herbert. 1994. "Drei Aspekte zur Situation von AusländerInnen in Öästerreich in der Ersten Republik. Lehrerfortbildungsseminar *Vom Ziegelböhm zum Zeitungskolporteur. Leben und Arbeiten in der Fremde*. Unpub. MS.

Radtke, Frank-Olaf. 2002. Politik, Sprachen und Schule. In Walter Weidinger ed. *Bilingualität und Schule 2. Wissenschaftliche Befunde*. Wien: öbv&hpt, 88-97.

Rindler Schjerve, Rosita ed. 2003. *Diglossia and Power. Language Policies and Practice in the 19th Century Habsburg Empire*. Berlin/New York: Mouton de Gruyter.

Simpson, Stuart. 2004. Vienna – the largest language school in Europe. Europa Büro, Wiener Stadtschulrat.

Statistik Austria. 2001. *Statistisches Jahrbuch 2001*. http://www.statistik.at/jahrbuch/deutsch/start.shtml (16 March 2002).

Vienna Bilingual Schooling. Wiener Schulführer 2001 Infoseite. http://www.schulfuehrer.at/information/bilingual/bilingual.htm. (14 March 2002)

Wakounig, Vladimir. 1994. Minderheitensprachen in Österreichs Bildungspolitik. In Andreas Paula ed. *Modelle für den Umgang mit Sprachen und Kulturen*. Klagenfurt/Celovec: Drava Verlag, 199-217.

Waldrauch, Harald. 2001. Muttersprachliche Lehrerinnen und Lehrer in Österreich. Statistische Auswertung für das Schuljahr 2000/2001. (= Informationsblätter des Referats für Interkulturelles Lernen Nr. 5/2001).

Wascher, Manuela Claire. 1999. Mehrsprachigkeit in Schule und Unterricht. M.A. thesis, University of Vienna.

Weidinger, Walter ed. 2001. *Bilingualität und Schule. Ausbildung, wissenschaftliche Perspektiven und empirische Befunde*. Wien: öbchpt.

Weidinger, Walter ed. 2002. *Bilingualität und Schule 2. Wissenschaftliche Befunde*. Wien: öbv&hpt.

Weidinger, Walter. 2000. "Wiener Kinder sind ...". Wien: Stadschulrat (Informationsblatt).

Weidinger, Walter. 2001. Bilingualität und Schule – Zahlen, Fakten, Trends am Beispiel der Wiener Pflichtschulen. In Walter Weidinger ed. *Bilingualität und Schule. Ausbildung, wissenschaftliche Perspektiven und empirische Befunde*. Wien: öbv&hpt, 12-21.

Weidinger, Wiltrud. 2001. Bilingualität und Schule aus erziehungsiwssenschaftlicher Sicht. In Walter Weidinger ed. *Bilingualität und Schule. Ausbildung, wissenschaftliche Perspektiven und empirische Befunde*. Wien: öbv&hpt, 80-161.

Wright, Sue. 2004. *Language Policy and Language Planning* . Basingstoke, Hampshire: Palgrave Macmillan.

Zeller, Albert. 2000. *English across the curriculum: English as the language of instruction at Lower Austrian AHS-secondary schools*. University of Vienna: M.A. thesis.

How to contact us:

VIE🅔🆆[z]

c/o
Institut für Anglistik & Amerikanistik der Universität Wien
Universitätscampus AAKH, Spitalgasse 2, Hof 8
A – 1090 Vienna; Austria

fax   **(intern.) 43 1 4277 9424**
eMail   **gunther.kaltenboeck@univie.ac.at**
W³   http://www.univie.ac.at/Anglistik/views.htm
                          (all issues available online)