

# Middle English Phraseology: Problems Identifying the Source(s)

---

MELANIE BORCHERS  
University of Duisburg-Essen  
melanie.borchers@uni-due.de

**ABSTRACT.** The following paper is based on findings of a pilot study dealing with the problem of determining which Middle English phraseologism might be based on native patterns as opposed to those multi-word lexemes that are due to foreign influences—in our case from Old and Middle French. With the help of two exemplary analyses the difficulties in identifying the source language(s) of Middle English phraseologism will be illustrated.

## 1 Introduction

The Middle English (henceforth ME) period is a period marked by language contact which has interesting consequences for its phraseology. This work will have a closer look at the source languages Old English (henceforth OE) and Old/Middle French<sup>1</sup> (henceforth OF/MF), disregarding the phraseological influences that Old Norse (henceforth ON) might have exerted on English and excluding the highly problematic OF/MF loans that themselves are Latinisms in French. This means that such lexemes that are borrowings from another language have been excluded for reasons of ruling out third language influences.

## 2 Phraseological Loan Strategies

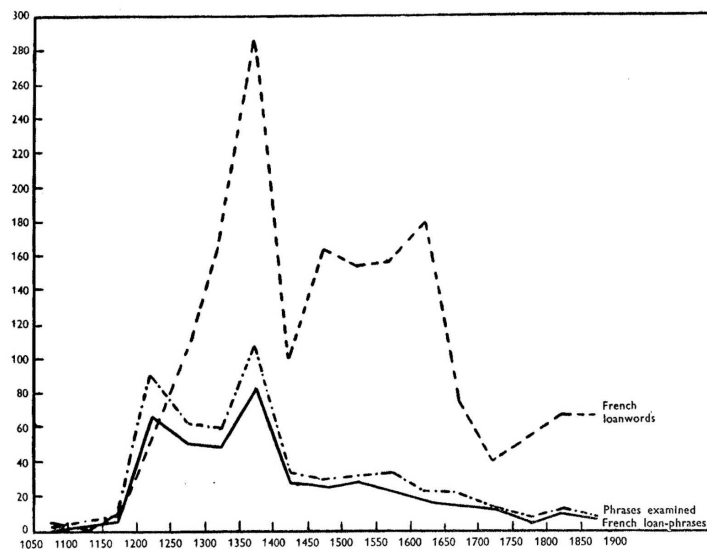
While the pilot study provided positive proof for the hypothesis that there seems to be a correlation between the lexical and the phraseological OF influence (cf. Diagram 1) and thereby manifesting the importance of the second period of French influence (1250 to 1450), this paper will claim the following:

The OF phraseological influence, as an extension of the lexical one, concerns all levels of phraseological loan strategies. To clarify the three different strategies of

---

<sup>1</sup> OF and MF will have to be differentiated even further into at least two different regional varieties that exerted their influence on English, namely Norman French, the French that the conquerors brought with them from Normandy, and Central or Parisian French, which had its main period of influence later. I will only mention the doublets *catch* vs. *chase* to illustrate the enormous value of such a distinction.

borrowing that are evidenced in the data, Haugen's (1964) terminology<sup>2</sup> has been adapted for the phraseological level. Subsequently, (i) loan phrases, (ii) phraseological loan translations and (iii) phraseological loan blends have revealed to have either been adopted from the foreign language, or to have strengthened and enriched an already existent English phraseologism by the foreign idiom.



Frequency of first entries. The horizontal axis indicates the time by half-centuries. The vertical axis the corresponding number of first entries.

**Diagram 1: Prins' subjoined graph of French loan words as well as loan phrases**

### 3 The Corpora for the Present Study

Turning to the corpora upon which this study is based, the various employed corpora will be delineated in the following paragraphs. The ME data have been taken from two different sources. First, the ME section of the *Helsinki Corpus of English: Diachronic and Dialectal (HC)*<sup>3</sup> was taken into account. This corpus comprises 91 text files containing approximately 650,000 words covering the time span from 1150 to 1500. Secondly, the *Corpus of Early English Correspondence Sampler (CEECS)*<sup>4</sup> comprises a total of 1147 letters, consisting of about 450,000 words and covering the time span from 1418 to 1680, thus

<sup>2</sup> According to Haugen (1964) there are three diverse loan strategies: (i) loan words; i.e. morphemic importations from a donor language, (ii) loan translations, i.e. morphemic substitutions without the importation of form, and (iii) loan blends, which are an intermediary borrowing showing morphemic substitution as well as importation

<sup>3</sup> Compare Merja Kytö (1996).

<sup>4</sup> Compare Arja Nurmi (1998).

extending beyond the data relevant for this paper. The extracted data, i.e. excluding the EModE<sup>5</sup> data, amounts to 126,000 words, taking into account less than a third. The *CEECs* might be considered at least to some degree semi-oral data. Though there is no question that letters are medially of written quality, the data could however be regarded as to be prone to some extent to the oral use of language, too. In summary, the resulting ME corpus that this study is based on consists of approximately 776,000 words. Realising that only a restricted quantity of available data for the ME period is taken into consideration, the author is aware of the fact that due to the limited surviving data the expected findings may be open to criticism.

Having formed a ME corpus, the next step is to examine the OE and OF corpora with which the ME findings could be compared. The OE data are again taken from the respective section of the *Helsinki Corpus of English: Diachronic and Dialectal (HC)*. This section dates from the earliest sources until 1150 and consists of about 410,000 words. In addition to the data from the *HC*, this study takes the *OEC, The Old English Corpus*<sup>6</sup> into account. Except for some variant forms, this corpus consists of all surviving OE written material, thus encompassing over 3,000 different texts and a total of approximately 3,845,000 words. The material, however important, will solely be viewed as a means of comparison for the ME findings.

Having established a corpus of the OE material for comparison, it is time to turn to the corpus of OF. The *Base de Français Médiéval (BFM)*<sup>7</sup> offers the desired French material. The database comprises 74 complete OF and MF texts amounting to about 3,000,000 words and thus assuring diversity of texts.

## 4 Choice of Keywords

The choice of keywords is as essential as the choice of the corpora and thus this section is dedicated to provide the criteria why some keywords are more appropriate than others. Though the pilot study was based on more keywords, three will suffice to exemplify the decision process.

There are three different categories that the keywords might be classified into: native lexemes, Romance borrowings and thirdly lexemes of hybrid origin. Lexemes are taken as native lexemes if they are of Germanic origin, Romance

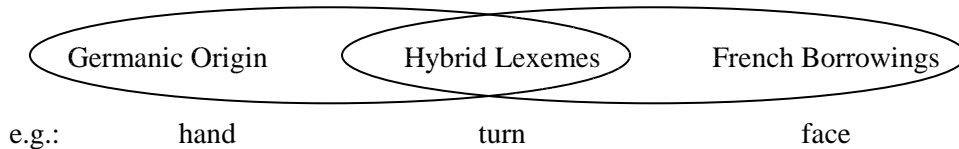
---

<sup>5</sup> The present study sticks to the periodisation the *The Cambridge History of the English Language* proposes, which dates the EModE period from 1476 to 1776.

<sup>6</sup> Full details of the sources used in compilation of the *Old English Corpus* may be found in the following publication: Frank, Roberta & Angus Cameron (eds.). (1973). *A Plan for the Dictionary of Old English*. Toronto: University of Toronto Press.

<sup>7</sup> For further information on the BFM project see its website at <URL <http://bfm.ens-lsh.fr/>>.

borrowings are lexemes which have been morphologically integrated from French<sup>8</sup> and lexemes of hybrid origin are the ones for which both is true. Consequently, findings containing native lexemes might show an influence from French phraseological units while Romance borrowings might display their original phraseological constructions in the integrating language. For lexemes of hybrid origin both characteristics might apply; i.e. they constitute their intersection (cf. Diagram 2).



**Diagram 2: Three-parted distinction of keywords**

## 5 Findings

In the following paragraphs two exemplary analyses will be presented in order to demonstrate the highly problematic process of identifying *the* source of a ME phraseologism.

### 5.1 FACE TO FACE

One of the chosen items is *face*, which mainly occurred as a noun in ME, whereas in PDE the verb seems to be used about as frequently as the noun. While the noun was borrowed from OF as early as the 14<sup>th</sup> century, *face* appeared as a verb<sup>9</sup> about a century later. The *OED* dates its first occurrence as 1440.

In agreement with *The Oxford Dictionary of English Etymology* I reached the same conclusion that *face* superseded early ME *onsene* (< OE *ansi* ← *en*). The OE nouns *nebb* and *ondwlita* were also replaced by the OF loan word *face*. While some remnants of *nebb* can still be found in the ME corpus, there is not a single ME occurrence of *ondwlita*.

---

<sup>8</sup> In this study Romance is taken as solely referring to the French language.

<sup>9</sup> Although the consulted OF dictionaries do not contain a French verb *facer*, there must have been some corresponding verb since the *BFM* delivers many occurrences of a verbal form of *face*. Another factor that might have fostered the verbalisation of the nominal form *face* in ME might be the fact that *face to face* often co-occurs with the verb *to see*. While Brinton (1996) presents evidence of a modern attitude to segmentalise verbs in her article “Attitudes Toward Increasing Segmentalization”, *to face* might represent some kind of ME counter effect to the modern development. Whereas today we would rather use *to take a walk* than the simple verb *to walk*, one could hypothesize that *to see face to face* was desegmentalised with *to face* as the outcome.

Beside two insignificant alliterative co-occurrences, the most prominent example of twin formulas in the preceding pilot study comprises the noun *face*, a twin formula that is manifested three times in the ME material.

- (1) And Moises seide to the Lord, Egipcians here not, fro whos myddil thou leddist out this puple, and the dwelleris of this loond, whiche herden that thou, Lord, art in this puple, and art seyn *face to face*, (...). (HC CMOTEST)
- (2) Je en avrai une coronne es ciex , plus que les angres , qui le voient *face a face* , par quoy il convient que il le croient . " (BFM JOINVILLE)

This twin formula evidently goes back to French (cf. (2)). Not only was the noun itself a loan word from French but it was also integrated into the English language by taking its twin formula. This, however, is an obvious though erroneous conclusion that might only be drawn when taking the numerous French data into consideration. To be conclusive on the origin of this phraseologism it does not suffice to have a look merely at the French data. The following OE hit delivers invaluable information, too.

- (3) (...) & eall hire wille hire is geforðed on godnyse, for heo gesihð hire Sune eall ealswa he is, *neb wið nebb*, & wuneð on ecere blisse mid him. (HC BE000507)
- (4) for ha # seoð him as he is. *nebbe to nebbe*. (HC CMSAWLES)
- (5) to weorin o þi meidhad. & secheð erst up-on hire *nebbe to nebbe*. (HC CMHALI)

*Nebbe* as the OE word for ‘face’ already occurred in such twin constructions before the Conquest (cf.(3)) and was still existent in the ME period. (4) and (5) manifest the two relicts that survived to the ME period. While all the OE formulas indicate *with* to be the combinatory preposition, the two ME hits reveal that *wið* had been replaced by the preposition *to*<sup>10</sup>. This new preposition has remained the combining element within this twin formula from then onward. Since these two matches that appear in the same timeframe (i.e. early ME) as well as dialect (i.e. West Midland) and none appear for the rest of the ME period, neither geographically nor temporarily, they might be seen to be the last to resist the French phraseological borrowing *face to face*.

From a quantitative point of view, it is interesting to compare the frequency of both formulas. (3) represents the sole example of this construction in the OE material. Even if one includes the two remnants in the ME data a frequency of only 0.58:1,000,000<sup>11</sup> is obtained. This stands in remarkable contrast to the twin

---

<sup>10</sup> Compare Hornero Corisco (1997: 43) for further details.

<sup>11</sup> Since the two survivals of the ME material are taken into account as well, the number, though open to criticism, shows the maximum number of all occurrences in the OE as well as early ME material, which means that the actual frequency of this OE twin formula has to be regarded as even smaller than 0.58:1,000,000.

formula *face to face* which evidences a frequency of 3.86:1,000,000 in the ME data. This increase in use might be due to the French influence.

In summary, the ME phraseological unit *face to face* should be regarded as a French borrowing on the one hand and an OE survival on the other. The latter of these most likely facilitated the integration of the loan phrase from French into Medieval English usage.

## 5.2 CRY AND WEEP/WEEP AND CRY

The *OED* traces the verb *cry* back to a French origin, to the still existent French verb *crier*. From this a noun was formed (French *cri*) which again was adopted into English; thus, the ME noun *cry*. The first instances of its use in English can be dated back to the 13<sup>th</sup> century.

Amongst others, the verb *cry* replaced the OE verbs *gre* ← *tan*, *re* ← *otan* and *we* ← *pan* while *cry* as a noun replaced *cierm* or *hrem*. According to Ogura (1996: 130) “OE cleopian was replaced by callen in the sense ‘to call’, and by crien (< OF crier) in the sense ‘to cry’. Once borrowed, *crien* begins to be used in the senses ‘to cry’, ‘to weep’, ‘to yell’, and ‘to beg’”.

The corpus reveals four instances of the verbal combination *cri* ← *en* and *we* ← *pen*, denoting an intensifying shared meaning, ‘to mourn’. (6) represents a pure instance of an uninterrupted phrase, (7) and (8), however, have been adapted to syntactical demands in that one verb is placed in the initial and the second verb is positioned into the additional main clause. The clauses are then connected by the conjunction *and*. Thus, these examples hint at the possibility that the combination may extend past the phrasal level.

- (6) And also a com be a touret, þat was in þe castel iset, A herde *wep*e and *cri*e; þederward he gan him hie. (HC CMBEVIS)
- (7) þe bischop also *w*eop on him : and on him *cri*de faste, (...). (HC CMSELEG)
- (8) (...) þe Chanones *w*epen with him al-so : and *cri*den cristes ore; (...). (HC CMSELEG)

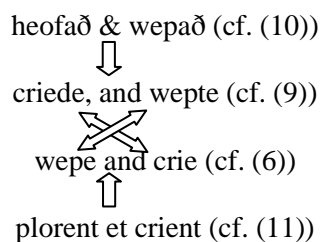
Whereas the preceding examples suggest that the two verbs always combine in the same word order, i.e. first *we* ← *pen* and then *cri* ← *en*, (9) provides an instance where the contrary is true. This implies that the constituents are interchangeable. Though this is the sole phenomenon in the data, it is rather likely that ME usage combined them both ways.

- (9) Therfor al the cumpeny *cri*ede, and *w*epte in that nyzt, and alle the sones of Israel grutchiden aȝens Moises and Aaron, and seiden, We wolden that we hadden be deed in Egipt, and not in this waast wildirnesse; (...). (HC CMOTEST)

But how could this implementation of word order come into existence? An analysis of the OE data suggests the first possible answer: the occurrences of the verb *héofian*, meaning ‘to lament’, prove the existence of a comparable combination already in OE (cf. (10)). This instance suggests that the French verb solely replaced the OE constituent *héofian*. A thorough investigation of the OF corpus material elucidates the situation even further. (11) and (12) indicate the co-occurrence of the two verbs that had already existed in OF usage.

- (10) (...), He cwæð: Wa eow þe nu hlihhaþ, for þan þe ge eft *heofað & wepað*.  
(OEC BE000464)
- (11) (...); les fames forment le maudient et de pitié *plorent et crient*; (...).  
(BFM THEBES2)
- (12) Et orent li pluisseur pité des hommes, fenmes et enfans qui *ploroient et crioient* a haus cris, et les laissoient passer et widier la ville legierement; (...).  
(BFM FROISSART1)

In conclusion, the merger of OE *wepað* with its OF synonymous verb *cry* immersed as a fused pair in ME.



**Diagram 3: illustration of the merging process of ME *weep and cry/cry and weep***

## 4 Conclusion

The preceding two analyses demonstrate the difficulty of identifying ME phraseological sources. Though the etymological background of the investigated lexemes might be clarified as to whether it is of either native or foreign origin or even an intersection of the two, the loan strategy poses bigger problems. The terminology offers a three-parted distinction into (i) loan phrases, (ii) phraseological loan translations and (iii) phraseological loan blends, but the phraseologism *weep and cry* illustrates how these distinctions need to be treated as both permeable and combinable. The two examples might indicate a much more complex picture of phraseological influence of French on English during the ME period. The potential that lexemes of hybrid origin bear is expected to surpass the complexity that the few example analyses have exhibited.

## References

- Brinton, Laurel. (1996). Attitudes Toward Increasing Segmentalization in English: Complex and Phrasal Verbs in English. *Journal of English Linguistics* 24: 186-205.
- Haugen, Einar. (1964). The Analysis of Linguistic Borrowing. *Language: Journal of the Linguistic Society of America* 25, 210-231.
- Hornero Corisco, Ana M<sup>a</sup>. (1997). French Influence on English Prepositions: A Study of *Ancrene Wisse*. *Studia Anglica Posnoniensia* 32, 33-45.
- Kytö, Merja. (1996<sup>3</sup>). *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts* [online]. Helsinki: Helsinki University. Available at:  
〈URL: <http://khnt.hit.uib.no/icame/manuals/HC/INDEX.HTM>〉 [Accessed 1 May 2006].
- Nurmi, Arja. (ed.). (1998). *Manual for the Corpus of Early English Correspondence Sampler. CEECS* [online]. Helsinki: University of Helsinki. Available at:  
〈URL: <http://khnt.hit.uib.no/icame/manuals/ceecs/INDEX.HTM>〉 [Accessed 1 May 2006].
- Ogura, Michiko. (1996). *Verbs in Medieval English. Differences in Verb Choice in Verse and Prose*. Berlin: Mouton de Gruyter.
- Oxford English Dictionary*. [online]. (2002-). Oxford: OUP. Available at:  
〈URL: <http://dictionary.oed.com>〉 [Accessed 4 February 2006].
- Prins, A.A. (1952). *French Influence in English Phrasing*. Leiden: Universitaire Pers Leiden.

## Data Basis

- La Base de Français Médiéval* [online]. (2005). Lyon: UMR ICAR / ENS-LSH. Available at: 〈URL: <http://bfm.ens-lsh.fr>〉 [Accessed 2 February 2006].
- The Corpus of Early English Correspondence. Sampler*. 1999. (ICAME Collection of English Language Corpora, 2<sup>nd</sup> edition.)
- The Helsinki Corpus of English Texts. Diachronic Part*. 1999. (ICAME Collection of English Language Corpora, 2<sup>nd</sup> edition.)
- The Old English Corpus*. 2004. (*Dictionary of Old English Project*, Centre for Medieval Studies, University of Toronto.)