Loan verb accommodation: a comparison of Old Norse and French in Middle English[1]

W. J U L I A N E   E L T E R ⬤
University of Mannheim

and

M A R L I E K E   S H A W ⬤
KU Leuven

Recent research shows that, even under direct insertion, loan verbs are subject to constraints: for instance, they enter non-finite categories more readily than finite categories. To deepen our understanding of such LOAN WORD ACCOMMODATION BIASES we investigate two contact situations to test whether biases hold in contact between closely related languages. A corpus study on Norse and French loan verbs entering Middle English compares the proportions of their finite and non-finite usage to gauge the impact of etymology and temporal distance to direct contact on loan integration. We identify significant bias towards non-finite use for both etymologies, but it is stronger for French than for Norse loan verbs. This suggests that biases are stronger in some contexts than in others: they are more prominent at a smaller temporal distance to direct contact and in contact between languages that are less closely related.

Keywords: historical linguistics, contact linguistics, loan verb integration, Old Norse and French loans, Middle English

1   Introduction

Many linguistic studies have focused on language contact and borrowing (e.g. Haspelmath & Tadmor 2009; Durkin 2014), in particular synchronic and diachronic cases of contact with English. Recent research has started investigating morphosyntactic constraints on loan word accommodation in French (F) loan verbs entering Middle English (ME) (Shaw & De Smet 2022): it has been revealed that speakers are biased to use loan words more in certain grammatical structures (e.g. non-finite verb forms) than in others (e.g. finite verb forms). This phenomenon, called 'loan word accommodation biases', will be explained in detail in section 2. Although this finding is innovative, Shaw (2022: 241) has indicated that the morphosyntactic integration of loan words may

---

benefit from additional research, as the research by Shaw & De Smet (2022) and Shaw (2022) focused on two specific language contact situations and cannot be generalised.

The present study aims to gauge how typological closeness of languages in contact and temporal distance to the period of contact impact the presence and strength of accommodation biases.[2] To this end, the present study compares the integration of loan verbs from French[3] and Norse-derived[4] verbs into ME. Both contact settings are of similar intensity and have the same replica language, English, into which lexical material is integrated; however, the contact settings differ considerably with regards to the typological closeness with the English language and the time distance to the period of contact (see section 3). These two factors can, therefore, reliably be used as the main points of comparison in the analysis. By doing so, this study may deepen our understanding of the nature of constraints on loan word accommodation in other diachronic contact situations than the French–Middle English contact. As such, the present study aims to contribute significantly to research on the morphosyntactic integration of loan verbs and on constraints on loan word accommodation.

In what follows, section 2 will expand on existing research on loan words as well as their accommodation to the replica language, or the language in which linguistic material is integrated. Section 3 will compare the Old and Middle English contact situations with Old Norse and French respectively. In the next sections, the focus will be on the case study at hand, first by formulating research questions and hypotheses (section 4), second by means of a detailed discussion of the data and methodology used for this study (section 5), and third by presenting the findings (section 6). The seventh section will offer a discussion of the findings as well as a conclusion and some avenues for further research.

2 Loan words and their accommodation

2.1 Loan word accommodation

During the Old and Middle English periods, the English language borrowed words from various languages, such as Latin, French and Old Norse (e.g. Ingham 2012; Pons-Sanz

---

[2] Typological closeness is taken to be indicated by overall similarities of morphosyntactic structures between languages (Thomason & Kaufman 1991: 72; Thomason 2001: 76). See section 3 for an overview of the genealogical relationships between the languages in contact which are taken to be the main source of difference in typological distance in the two contact situations under investigation. Additional related factors like the status of bilingualism and mutual intelligibility will also be discussed in sections 3 and 4. For a discussion of the role of relatedness in language contact, see Bowern (2013).

[3] The term 'French' or 'French-origin' is used as a simplified umbrella term for the following two historical varieties: Anglo-French, or the variety of French as spoken in medieval England (see Rothwell 1996), and Old French, or the variety of French as spoken in Paris at the time (see Einhorn 1975: 1). The only situation in which 'Old French' is used in this article is when referring to original Old French etyma forms so as not to confuse Old French with Contemporary French.

[4] The term 'Norse-derived' is used to indicate the high level of nuance necessary when characterising ME lexis influenced by or copied from Old Norse (see Dance, Pons-Sanz & Schorn 2019; Dance 2003, 2011, 2012, 2018). See section 3 for the reasoning behind this use of terminology.

2013; Durkin 2014). Borrowing is taken to mean the integration of new linguistic material into a language, often called 'replica language', on the model of linguistic material from another language, often called 'model language', with which the replica language has been in contact in some way (Weinreich 1953). Frequently investigated examples of model languages are Old Norse (ON) and French (Finkenstaedt & Wolff 1973; Grant 2009; Durkin 2014), which have added to the English language with verbs such as ME *taken* ('to take') and *travailen* ('to travail, work') respectively.

The number and nature of borrowings resulting from a contact situation depend on the intensity of contact (Thomason & Kaufman 1991), but also on, for instance, the morphological complexity of the borrowable categories (Matras 2020: 191f.). Therefore, linguistic closeness of the languages in contact may favour the borrowing of more complex categories (e.g. Meillet 1921; Moravcsik 1975; Johanson 2002; Winford 2003: 51ff.).

The present study concentrates on loan verbs.[5] The morphosyntactic implications of loan integration of verbs are generally focused on less in models of borrowability and loan integration (cf. Thomason & Kaufman 1991; Matras 2020) than the formal or semantic implications of the loan integration of other lexical categories, or they are presented as a constraint on lexical borrowing (cf. Winford 2003). Wohlgemuth's (2009) work on loan verb accommodation is seminal and identifies four accommodation strategies based on his typological research: 'direct insertion', 'indirect insertion', the 'light verb strategy' and 'paradigm insertion'. Direct insertion describes a loan verb integration pattern where replica language inflections are added directly onto the borrowed stem, while in indirect insertion an additional verbalising affix is added to the word stem of the copy before it can be inflected in the replica language (Wohlgemuth 2009: 87, 94). Under the light verb strategy, Wohlgemuth (2009: 102) classifies all patterns where a copied verb is integrated as part of a complex predicate in combination with a dedicated light verb which carries inflections. Finally, paradigm insertion shows copied verbs continuing to carry their source language inflections in the replica language (Wohlgemuth 2009: 118, 119). Direct insertion is the most frequent strategy cross-linguistically and has, moreover, been identified as the most prominent strategy for loan verb insertion into English in both the Scandinavian and French contact situations (Wohlgemuth 2009: 338). Examples of direct insertion are Old French *comander* and Old Norse *reisa*, which are implemented in Middle English as *commaund-en* and *reis-en* respectively (cf. Lewis *et al.* 1952–2001). Both loan verbs are used with the native English infinitival *-en* marker (cf. native English *find-en*). Since inflection cannot be avoided under direct insertion, Wohlgemuth's (2009: 291) subsequent argument is that loan verb integration is not as constrained by inflection as much as often assumed (e.g. Harris & Campbell 1995: 135; Sijs 2005: 56–7).

When entering a language through direct insertion, loan verbs are integrated grammatically into their replica language system (Poplack, Sankoff & Miller 1988; Muysken 2000), which

---

[5] See Hug (1987: 7ff.) and Dekeyser (1986: 261) for an assessment of the role of verbs in the overall lexical contribution of French and Old Norse to the Middle English lexicon.

means that they adopt replica language inflections and can be used in all morphosyntactic categories in which model language verbs can be used.[6] In the case of Norse-derived ME *taken* and French-origin ME *travailen* in examples (1)–(2), both taken from *The Penn–Helsinki Parsed Corpus of Middle English* (PPCME2) (Kroch, Taylor & Santorini 2000–), the verbs are inflected as past third-person singular forms. *Taken* is used in the past strong form *tooke* ('took'), and *travailen* in the past weak form *trauaylde* ('worked').

(1) Marie Magdeleyne **took-e** [ON: taka] an alabaustre box of precious oynement
     Mary   Magdalene   take.PST-3SG   an alabaster box of precious ointment
     'Mary Magdalene took an alabaster box of precious ointment.'
     (*Aelred of Rievaulx's De Institutione Inclusarum*, CMAELR3,44.540)

(2) He **trauayl-de** [OF: travaillier] forto bryng man ynto euerlastyng reste
     He work-PST.3SG   to   bring man into everlasting rest
     'He worked to bring man into everlasting rest.' (*Mirk's Festial*, CMMIRK,2.20)

That loan verbs *taken* and *travailen* can be used in inflected forms such as past forms points to them functioning as fully integrated verbs in ME.

2.2 Loan word accommodation biases

More recent research confirms that loan verbs can be used just like native verbs. However, even under direct insertion, loan verbs are subject to constraints on inflection and are biased towards some morphosyntactic usage categories, a phenomenon referred to as LOAN WORD ACCOMMODATION BIASES (De Smet 2014; Shaw & De Smet 2022). Specifically, French loans in late ME have been found to occur disproportionately more frequently in uninflected forms than in inflected forms when compared to native verbs (Shaw & De Smet 2022). Additionally, loan verbs are disproportionately more frequent in non-finite forms (i.e. infinitive, past participle, present participle) than in finite forms (i.e. imperative, past, present) when compared to native verbs. An example of the non-finite usage of French loan verb *maintenen* ('to maintain') is provided in (3), where the verb is used in its past participle form, *meigtened* ('maintained'). The non-finite usage of this French loan verb is contrasted with the finite usage of native English verb *komen* ('to come') in (4), where *kom* ('came') is a third-person singular form in the past tense. Both examples have been taken from the same text sample from *The Helsinki Corpus of English texts* (Rissanen *et al.* 1991).

(3) if he had ben mair, I wot wel he wold-e   haue
     if he had been mayor I know well he would.PST.SBJV-3SG have.AUX.PRS
     meigten-ed [OF: maintenir] all his ordinances
     maintain-PST.PTCP   all hys assignments
     'If he had been mayor, I know that he would have maintained all his assignments.'
     (Helsinki Corpus, *Thomas Usk's Appeal* l.125–6)

---

[6] The present article only addresses morphosyntactic integration. Phonological integration is not considered.

(4) Also, that day that Sir Nichol Brembre was chose mair, a-non after mete
also that day that Sir Nicholas Brembre was chosen mayor soon after meal
kom John Norhampton to John Mores hows […]
come.PST.3SG John Northampton to John Moore's house […]
'Also, the day that Sir Nicholas Brembre was chosen as mayor, soon after lunch John
Northampton came to John Moore's house […].' (Helsinki Corpus, *Thomas Usk's Appeal*
l.135–6)

The usage of finiteness in (3) and (4) above exemplifies the dominant distribution of French loan verbs used in non-finite forms as compared to native English verbs, which are more prevalent in finite forms.

Biases of loan words towards specific morphosyntactic categories have been found not just for verbs, but also for adjectives, both in the historical French contact outcome and for Modern English loan words in Dutch (Shaw 2022; Shaw & De Smet 2022). However, this article will focus on loan verb accommodation biases in historical contact situations. As suggested by Shaw (2022: 241), the morphosyntactic integration of loan verbs requires further investigation, particularly across different contact situations, as the research by Shaw & De Smet (2022) has only focused on French loans in Middle English and Modern English loans in Modern Dutch. Therefore, the present article deepens our understanding by comparing the accommodation of French and Norse-derived loan verbs into ME. The factors of temporal distance to the period of direct contact and typological relation to ME are the main points of comparison. We compare loan verbs from both French and Norse regarding their morphosyntactic accommodation into the English replica language system by operationalising accommodation biases as a measure of loan verb integration.

3 A comparison of French and Norse-derived loan verbs in Middle English

Considering the nature of the contact situations under comparison, the question is whether the biases previously attested for French verbs in ME still hold in the contact between typologically and lexically closer ON and ME. The contact situations are comparable regarding their intensity, but they differ in other characteristics.[7] Firstly, Old French and Middle English belong to different language families, namely the Romance and Germanic branches of the Indo-European languages respectively, while Old English and Old Norse are both Germanic languages. The closer genealogical connection between Old Norse and Old English is reflected in a higher structural and lexical closeness, which resulted in adequate mutual intelligibility of the languages in contact for monolingual speakers (cf. Townend 2002), which was not the case for speakers during contact with French.

Secondly, while the contact with Scandinavian in England roughly spans 787–1042 CE and spreads from the northeast to cover the area that becomes known as the *Danelaw* (Pons-Sanz 2013: 6f.; cf. Thomason & Kaufman 1991: 280–2), French contact spreads

---

[7] For detailed accounts, see Townend (2002, 2006), Durkin (2014) and Ingham (2012, 2020), among others.

from the south to cover all English territory, starting in 1066 CE, and lasts until *c.*1500 CE (e.g. Rothwell 1983: 259–60). This difference in topological progression and overall spread of these contact situations reflects potential differences in their intensity and lasting impact on different dialects across England, which will be explored in section 6.2.

Thirdly, while a high level of societal bilingualism is assumed for the contact with Scandinavian (Townend 2002: 60, 189; 2006: 70), contact with French is characterised by higher individual bilingualism (Ingham 2012: 5). Additionally, the Old English–Scandinavian contact arguably involved two adequately mutually intelligible languages in contact (Townend 2002: 183f.). This enabled speakers of either language to employ processes of accommodation in a so-called 'switching code' during mutually intelligible communication (Townend 2002: 60, 183ff.). We agree with Weinreich (1953: 56) that lexical borrowing is not restricted to the bilingual individuals of a bilingual society. Thus, borrowing of lexical material and identification of interlingual correspondences between Old Norse and Old English were available to monolingual speakers of English in this situation (Townend 2002: 60, 203). The higher degree of individual bilingualism characterising the contact with French and its implications for the integration of loan words by bilingual individuals has been discussed in Shaw (2022: 53). Following these assumptions, we concur with Wohlgemuth (2009: 30) in the proposition that the contact situations investigated in the present work allow for a comparison of loan verb integration outcomes in his typology, despite the difference of their status of societal versus individual bilingualism.

Lastly, regarding the socioeconomic dynamics between linguistic groups, the contrast in prestige and power between speakers of French and English is arguably more stark than that between speakers of ON and English, although both vary across the respective timespans (cf. Townend 2002, 2006; Ingham 2012, 2020).

As to the identifiability of Norse-derived and French loan words, words from Romance languages are more securely identifiable as loans in English than possible loans from Old Norse. The close genealogical relationship and resulting higher formal and lexical closeness of Old Norse and Old English make secure identification of lexical material as of Scandinavian origin more complex, especially in the large number of cognates between these languages. In this matter we defer to the detailed work of the *Gersum* project (Dance, Pons-Sanz & Schorn 2019) and Dance (2003, 2011, 2018) and adopt their classification of evidence and terminology for lexemes' etymological origin as being 'Norse-derived'.

Regarding the timing and nature of the influx of loan lexis from both contact settings, most Norse-derived lexis is first attested in writing only in ME (Hug 1987), at the same time as the French loan lexis entered the English language (peak between 1350 CE and 1420 CE (Dekeyser 1986)). While more recent work on the Norse element in Old English (OE) texts (Pons-Sanz 2007, 2013; Dance 2003) does reveal earlier records of Norse-derived lexis in English, the overall picture of much of Norse-derived lexis being first attested in ME still prevails (cf. Proffitt 2000–; Dance, Pons-Sanz & Schorn 2019). As Durkin (2014: 178ff.) notes, this reflects a gap in the record rather than actual borrowing of Old Norse lexemes after the end of direct contact. As it is not

reconstructable when the majority of the Norse-derived words first attested in ME would have entered spoken OE, this limits the value of the date of first written attestation as an assessment of these words' existence in the English language (Durkin 2014: 189). What is certain, however, is that the temporal distance to the period of direct contact with ME differs greatly for ON and French.

4 Research questions and hypotheses

This study seeks to gauge the impact of the etymology (ON and French) of the loan verb on its finiteness in usage, which can show to what extent the verb is morphosyntactically integrated into ME. Additionally, we take a short-term diachronic perspective towards the data and investigate the possible effects of temporal distance to the period of direct contact. Concerning these two objectives we set out two research questions:

> RQ1: Do accommodation biases shown by loan verbs from different model languages differ in strength depending on the typological closeness of the languages in contact?

> RQ2: Do accommodation biases decrease over time relative to the temporal distance to the period of direct linguistic contact?

Concerning the first research question, we subscribe to the view that linguistic closeness facilitates borrowing of more complex categories (e.g. Meillet 1921; Moravcsik 1975; Winford 2003: 51ff; cf. Johanson 2002). Accommodation biases are, therefore, hypothesised to be less strong for loan verbs from typologically closer model languages than for those from typologically less closely related model languages throughout ME. More specifically, accommodation biases are expected to be stronger for French verbs than for Norse-derived verbs. For the second research question, we hypothesise that accommodation biases weaken over time, with increased temporal distance to the period of direct contact (cf. De Smet & Shaw 2024: 5). Thus, the number of loan verbs used non-finitely is expected to be higher in earlier ME texts than in later ME texts.

5 Data and methodology

5.1 Data and operationalisation

To address these questions, a corpus study on Norse-derived and French loan verbs entering ME was conducted, comparing them to a baseline of native English verbs. Their overall usage as well as the nature and course of their morphosyntactic integration were compared. In this study, accommodation biases served as a measure for the degree of integration of loan verbs, meaning that stronger accommodation biases imply less complete morphosyntactic integration. We operationalised accommodation biases as the difference in relation of non-finite and finite uses of the verbs between foreign etymology verbs (French and ON) and English verbs.

5.2 Data extraction

Data were extracted from *The Penn–Helsinki Parsed Corpus of Middle English* (PPCME2) (Kroch, Taylor & Santorini 2000–) and *A Parsed Linguistic Atlas of Early Middle English* (PLAEME) (Truswell *et al.* 2018). The PPCME2 corpus is mostly based on the Middle English section of the diachronic part of *The Helsinki Corpus of English Texts* (Rissanen *et al.* 1991). It encompasses 56 text samples, totalling around 1.2 million words. It is subdivided into four time periods: M1 (1150–1250 CE), M2 (1250–1350 CE), M3 (1350–1420 CE) and M4 (1420–1500 CE), following the Helsinki Corpus classification.[8] The PLAEME corpus includes 68 text samples from the *Linguistic Atlas of Early Middle English* (Laing 2023–) which total roughly 173,000 words.[9] Both corpora include syntactic annotations (cf. Truswell *et al.* 2019) following the *Penn Parsed Corpora of Historical English*.[10] Together, these parsed diachronic corpora span the time between 1150 CE and 1500 CE and include prose of different genres as well as some poetry. However, there is approximately two-thirds less data for the M2 subperiod in the PPCME2 corpus than for the other subperiods in this corpus (Percillier & Trips 2020: 7172f.). This is why the PLAEME data were used as a supplement to make the data more balanced diachronically (Truswell *et al.* 2019).

This combination also leads to a more balanced representation of dialect areas (cf. Truswell *et al.* 2019: 6), as the PPCME2 contains more texts from the east and west Midlands overall and texts from the M2 subperiod only represent the southeast of England, while the smaller PLAEME corpus contains relatively more northern and southern texts (cf. Percillier & Trips 2020: 7173). As Scandinavian contact spread from the northeast in late OE and French contact spread from the south starting in 1066 CE (see section 3), the contact situations under investigation were most intense in different dialects at different times. Therefore, we controlled for the varying intensity and topological spread of linguistic contact in different regions, operationalised as four broad dialect areas, namely Northern, East Midlands, West Midlands and Southern. Herein, the latter three are operationalised as in the *Penn Parsed Corpora of Historical English* and Southern combines the Southern and Kentish dialect classifications for the PPCME2 data.[11] As the dialect text metadata for the PLAEME data is more fine-grained, its broad localisation subcategories South East, South Central, South West and Essex and London were collapsed into one Southern category while the categories North West Midlands and South West Midlands were collapsed into a general West Midlands category to make them congruent with the PPCME2 dialect groups.[12] The

---

[8] See www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4/

[9] See www.amc.lel.ed.ac.uk/amc-projects-hub/project/p-laeme-a-parsed-linguistic-atlas-of-early-middle-english/

[10] See the annotation manual for the *Penn Parsed Corpora of Historical English* at www.ling.upenn.edu/ppche/ppche-release-2016/annotation/ for detailed information.

[11] For more information on the dialect classification of texts, see www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4/

[12] Detailed information on the PLAEME text localisation may be found at https://github.com/rtruswell/PLAEME_current/blob/master/PLAEME_texts.csv

PLAEME dialect groupings for Northern and East Midlands largely correspond to those of the PPCME2. This way, patterns of loan verb accommodation biases for verbs from either contact situation can be compared between high- and low-intensity contact areas for each contact respectively across both corpora. However, these four dialect areas are not represented equally in the combined data overall, with Northern accounting for 9.77 per cent of the combined data, East Midlands for 45.61 per cent, West Midlands for 26.23 per cent and Southern for 18.39 per cent of the data, and neither are they diachronically balanced, as table 1 shows (cf. also Percillier & Trips 2020: 7173, figure 3).

Note that 29.21 per cent of text from the aforementioned corpora is based on French or Latin originals with varying degrees of literality. Following Shaw (2022), we did not exclude texts on the basis of the language of the original text. From a diachronic perspective, the ME corpora represent an ongoing contact situation with French, while contact with ON had subsided by the end of the Old English period (see section 3). This directly reflects the factor of temporal distance to the period of contact, which possibly affects accommodation biases. This makes the ME data a fitting basis for this comparative analysis of loan verb accommodation.

We queried the dataset for all occurrences of lexical verbs in three etymological groups, namely English, French, and Norse-derived using *CorpusSearch* (Randall 2010). To fulfil this aim, we used versions of the PPCME2 and PLAEME enriched with verb lemmatisations from the *BASICS* project (cf. Percillier & Trips 2020).[13] Etymological origins of verbs for these three groups were operationalised as follows: French verbs were queried following the *BASICS* etymology annotations for French origin verbs. The queried list of Norse-derived verbs was based on the *Gersum* project database (Dance, Pons-Sanz & Schorn 2019) as well as the *Oxford English Dictionary* (Proffitt 2000–.) and *Middle English Dictionary* (Lewis *et al.* 1952–2001). We restricted the set of verbs to lemmata with strong phonological and morphological evidence supporting ON influence for which no cognates are attested in OE (*Gersum* category A1–A3, Dance, Pons-Sanz & Schorn 2019, e.g. *casten*), or where they are, they are neither formally nor functionally equivalent (*Gersum* category A1*–A3*, Dance, Pons-Sanz & Schorn 2019, e.g. *raise* vs *rear*). A number of verbs which were not classified in the *Gersum* database were added to the set of Norse-derived verb lemmata under investigation. These verbs were listed in the *Oxford English Dictionary* (Proffitt 2000–), *Middle English Dictionary* (Lewis *et al.* 1952–2001) or other current research on Norse-derived lexis in ME (Pons-Sanz 2007, 2013; Dance 2003, 2012) as being of early Scandinavian origin based on sufficient formal evidence. Of these we only included verbs listing no or contrasting native West Germanic cognates to match the conditions of the set of verbs extracted from the *Gersum* database (e.g. *liten* 'to dye'). By extension, the set of English verbs serving as a baseline contained all verbs annotated as 'non-French' in the *BASICS* annotations, also excluding non-contrasting

---

[13] The project *Borrowing of Argument Structure in Contact Situations* (*BASICS*) (2015–21) investigated how the borrowing of French lexical verbs into medieval English effected grammatical changes in the recipient language. For more information, see https://tinyurl.com/dfgbasics

Table 1. *Percentual distribution of text words per dialect area and time period in the combined PPCME2 and PLAEME data (n = 1,365,624)*

	M1	M2	M3	M4	Total
Northern	0.00	6.60	2.76	0.42	9.77
East Midlands	11.02	5.98	17.11	11.49	45.61
West Midlands	9.48	1.64	10.13	4.98	26.23
Southern	0.31	7.81	7.24	3.03	18.39
Total	20.82	22.03	37.23	19.92	100.00

close cognates between English and ON. This way, we reduced the overlap of the etymological verb sets between English and Norse-derived verbs in the extensive domain of cognates between these languages. We eliminated overlap between the etymological sets by excluding all instances ambiguously lemmatised (Percillier 2016: 210; Percillier & Trips 2020) between French and non-French lexemes (e.g. orthographic type *comyn* lemmatised as either *comen* 'to come' from OE *cuman* or as *communen* 'to share, commune' from OF *com(m)uniier*) and all verbs ambiguously lemmatised between non-contrasting close Old Norse and Old English cognates or other formally close lemmata (e.g. orthographic type *lythe* lemmatised as either *lithen* 'to sail, travel' from OE *līþan* or as *lithen* 'to alleviate' from OE *līþigian* or as *lithen* 'to listen' from ON *hlýða*). This way, we excluded lemmata that are of mixed Old Norse and Old English influence that could easily be integrated by directly mapping them onto the inflectional paradigm of the native cognate by way of identification between lexemes. Such copies would likely not show accommodation biases of the nature investigated in Shaw & De Smet (2022) and would not serve to answer our research questions. Including these in our data would have conflated the effects of accommodation processes at work in mutually intelligible communication between high cognate languages and the long-term structural accommodation of loan verbs without identifiable cognates in a language contact situation (see section 3). Our query also excluded *be* and *have*, modal verbs and gerunds, the former for their status as auxiliary verbs and the latter for their status as nominalisations. Fixed expressions like *according to* (ME: *accorden* < OF) and *that is to say* (ME: *seien*) were manually excluded as they are no longer actively generated structures, but lexicalised, and thus do not require active inflection in usage (Shaw 2022: 78).

5.3 Data analysis

We automatically annotated all retrieved instances of verbs concerning their verb form, lemma, etymological origin and finiteness of the morphosyntactic realisation, drawing on the extracted corpus data and annotations. This resulted in a total of 124,308 attestations. For our operationalised diagnostic of finiteness, we distinguished between non-finite (infinitive, present participle, past participle, passive participle) and finite

(inflected present, past, imperative). Additionally, we extracted text metadata concerning dialect (i.e. Southern, Northern, East Midlands, West Midlands; see above for categorisation) and Helsinki time period (i.e. M1, M2, M3, M4; see above for categorisation). For these annotations we followed the PPCME2 classifications (Kroch, Taylor & Santorini 2000–) to retain comparability across the two corpora.

On this dataset we ran basic quantitative analyses, relating the variables of etymology and finiteness of morphosyntactic realisation generally and across the variable values of time period and dialect area. Chi-square test was used to obtain p-values for the differences in proportion of finite and non-finite forms of each of the two foreign etymology sets, comparing them to the English baseline. Yates' correction was used as a measure to prevent overestimation of statistical significance of the data.[14] For subset analyses like lemma and frequency effects (see section 7.1.1), Fisher's exact test was applied, as this type of test is typically used for smaller sample sizes than Chi-square test (Levshina 2015: 214). The imbalanced nature of the data across the four variables did not allow for valid application of regression analysis. Therefore, we only conducted pairwise comparisons in this study to test the probability of differences being significant.

6 Findings

The total number of analysed attestations for all three etymological sets is represented in table 2. The absolute number of instances in the data is by far the highest for native English verbs (103,778), followed by French loan verbs (18,676) and only a comparably small amount of Norse-derived loan verbs (1,854).

The distribution of finite and non-finite forms for all three etymologies will be visualised in figure 1, after discussing three exploratory examples, depicting finite and non-finite forms in verbs of Norse, French and English origin.

For Norse-derived verbs (5) and French loans (6), we find both non-finite (a) and finite (b) usages throughout the ME dataset, just like we do for native English verbs (7). In example (5a), for instance, the Norse-derived verb *casten* ('to cast') is used in a non-finite form, namely as an infinitive, in this case *cast* ('cast'). Example (6a), too, illustrates the use of a non-finite form, but this time of a French loan verb, namely *receiven* ('to receive'). It is used in its past participle form, *receyved* ('received'). An example of an English form used non-finitely is *fyten* ('to fight') in (7), which is the infinitival form, hence *fyten* ('fight'). The examples thus show that both Norse-derived verbs and French loan verbs can be used non-finitely, just like English verbs. However, loan verbs of both origins can also be used finitely. In (5b), *eggen* ('to egg, incite') is used in the third-person singular of the past form, namely *eggede* ('egged'), and in (6b), *tormenten* ('to torture') is used in the third-person plural of the past form, namely

---

[14] It should be noted that the need for such a measure is disputed in the literature (e.g. Adler 1951).

Table 2. *Absolute numbers of verb instances in the PPCME2 and PLAEME data across three etymological subsets and finiteness of morphological form (n = 124,308)*

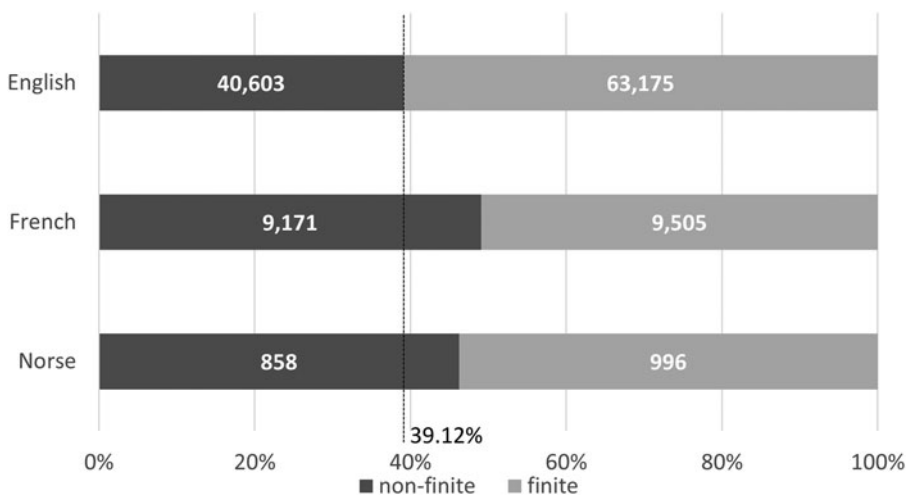|	Non-finite	Finite	Total
Norse	858	996	1,854
French	9,171	9,505	18,676
English	40,603	63,175	103,778
Total	50,632	73,676	124,308

Figure 1. General distribution of verbs of Norse-derived (p < 0.0001, Chi-square test) and French origin (p < 0.0001, Chi-square test) compared to the baseline of English verbs in finite and non-finite forms in combined PPCME2 and PLAEME data (n = 124,308)

tormentede ('tortured'). An example of native English verbs is given in (7), where *maken* ('to make') is used finitely as *makest* ('make'), in the second-singular person of the present, and *fighten* ('to fight') is used non-finitely as *fyten* in a *to*-infinitive.

(5) (a) and þou shal-t   **cast** [ON: kasta] hem   in-to dampnacioun
          and thou shall.PRS-2SG cast.INF   them into damnation
          'And you shall cast them into damnation.' (*The Earliest Complete English Prose Psalter*, CMEARLPS,170.7481)

     (b) He **egge-de** [ON: eggja] him þat he scholde sone. þe giwes ore louerd take.
          he egg-PST.3SG   him that he should   soon   the Jews our Lord   take
          'He egged him on to take our Lord to the Jews soon.' (*Life of Christ*, LAUD108ALIFE.969)

(6) (a) Another defaute   is this: that   men   doon
         another sin   is this that   men   do
         deedly synne after that they hand   **receyv-ed** [OF: recevoir] baptesme.
         deadly sins   after that they have.AUX.PRS.3PL receive-PST.PTCP   baptism
         'Another sin is this one: that men do deadly sins after they have received baptism.'
         (*The Parson's Tale*, CMCTPARS,289.C1.29)

     (b) Hi   **tormente-de** [OF: tormenter] him   strong   and   harde.
         they   torture-PST.3PL   him   strong   and   hard
         'They tortured him strongly and hard.' (*South English Legendary*, CORP145SELT.1131)

(7) þu   **make-st**   hym boþe fat &   stronge to **fyt-en** þe   ageyn
    you make-PRS.2SG   him both fat and   strong   to fight-INF   you   against
    'You make him both fat and strong to fight against you.' (*Sayings of St Bernard*, ADDE6AT.63)

The above examples show that verbs of Norse, French and English descent can be used seemingly easily in both finite and non-finite forms. However, the proportion of non-finite versus finite usage differs for the three etymological sets. Figure 1, in which the vertical dashed line (39.12 per cent) corresponds with the baseline of non-finite usage of English verbs, shows that both French-origin (48.87 per cent) and Norse-derived (46.28 per cent) verbs have significantly higher proportions of non-finite usage when compared to the usage of native English verbs (French $p < 0.0001$; ON $p < 0.0001$). Those proportions are based on the absolute frequencies, shown in white in figure 1. Note, however, that the datasets for verbs of each etymology differ vastly in absolute number of verb tokens, which has to be taken into account when looking at the findings.

Thus, the analysis shows that significant accommodation biases exist for verbs of both foreign etymologies. Whereas Shaw & De Smet (2022) had already revealed this finding for verbs of French origin[15] from a synchronic perspective on a smaller basis of ME data, it is the first time that the existence of accommodation bias towards non-finiteness is verified for Norse-derived verbs.

In examples (8)–(9), French-origin *disheriten* ('to disinherit'), *chalengen* ('to challenge') and *conqueren* ('to conquer') as well as Norse-derived *reisen* ('to raise') are used as *to*-infinitives, hence non-finitely. This is the type of construction in which they are, based on the finding above, statistically more likely to occur. English-origin verbs, in contrast, are more common in finite forms, such as *kom* ('came') in the third-person singular of the past, as was illustrated in example (4) (see section 2).

---

[15] The findings in Shaw & De Smet (2022) drew on two sources of data: (i) a full-text analysis of the PPCME2 versions (Kroch, Taylor & Santorini 2000–) of *The Parson's Tale* and *Mandeville's Travels*, with a total of 3,881 verb tokens, and (ii) the third subperiod of the Middle English section of *the Helsinki Corpus of English Texts* (Rissanen *et al.* 1991), with a total of 4,894 verb tokens.

(8) þei are more besy for-to **disherite** [OF: desheriter] here neyghbores more þan
they are more busy to disinherit.INF their neighbours more than
for-to **chalenge** [OF: chalengier] or to **conquere** [OF: conquer-re] here right heritage
to challenge.INF or to conquer.INF their rightful heritage
'They are more devoted to dispossess their neighbours than to challenge or to conquer their
rightful heritage.' (*Mandeville's Travels*, CMMANDEV,3.29)

(9) And he rod onto þe west partyes to **reyse** [ON: reisa] puple ageyn þe qween
and he rode to the western parts to raise.INF people against the Queen
'And he rode to the western parts to raise people against the Queen.' (*Capgrave's Chronicle*,
CMCAPCHR,152.3563)

These examples illustrate the trends found in figure 1. Apart from the finding that
French-origin and Norse-derived verbs are biased towards non-finite constructions as
compared to English-origin verbs, the analysis also reveals differences between verbs
of French and Norse descent: the non-finite bias is significantly stronger for French
loan verbs than for Norse-derived verbs ($p = 0.0215$, Chi-square test). This may
confirm our hypotheses (see section 4) that the non-finite bias is stronger (i) when the
two languages are typologically less close and (ii) when the temporal distance to the
period of direct contact is smaller in a synchronic comparison. However, at this point
we cannot yet confirm the two hypotheses separately as we have not yet distinguished
between them.

6.1 Disentangling typological and temporal distance effects

To disentangle the effects of typological distance and temporal distance to the time of
direct contact, we take a diachronic perspective on the ME data. Figure 2 shows the
proportion of non-finite usage for verbs of all three etymologies split up by Helsinki
subperiods, from M1 to M4.

Each subperiod of ME shows different trends concerning the proportions of
non-finite versus finite usage in the three etymological sets. Diachronically, the
non-finite bias for Norse-derived verbs steadily decreases throughout ME. In M1
Norse-derived verbs show a significant non-finite bias ($p = 0.0003$, Chi-square test)
but this is no longer significant ($p = 0.6330$, Chi-square test) by the M4 period.
This points to Norse-derived verbs not yet being well integrated at the end of the
direct contact situation between Old Norse and English. Biases for French verbs,
however, persist throughout ME and do not show a clear trend of decrease when
the temporal distance to the start of the period of direct contact increases. This
finding of persistent biases for French verbs even at the end of the direct linguistic
contact situation is parallel with the significant bias attested for Norse-derived
verbs in M1. What is more, the non-finite bias for French loans initially increases
throughout earlier ME (e.g. from M1 to M2). This may coincide with an increase
in texts translated from French and the peak of newly attested French loans, which
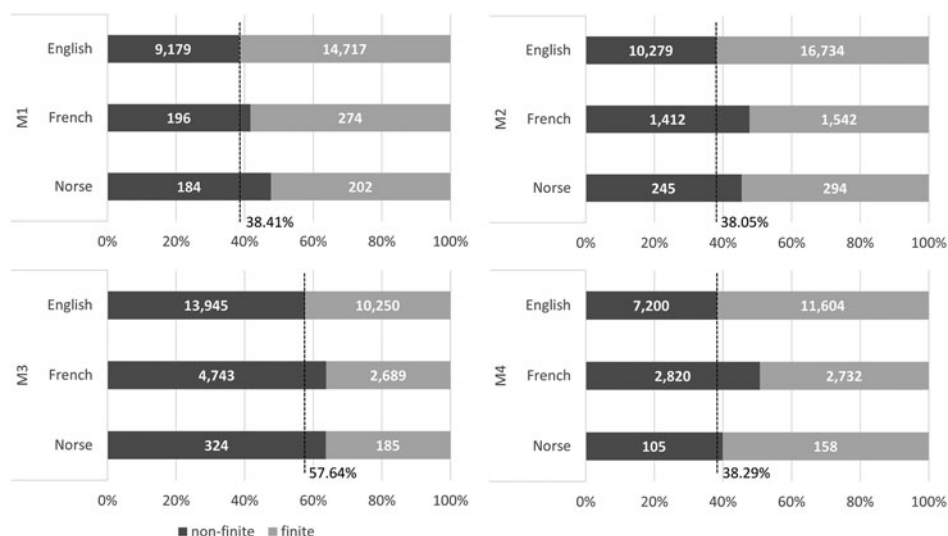are reflected in the data.

Figure 2. Diachronic distribution of verbs of Norse-derived and French origin compared to the baseline of English verbs in finite and non-finite forms in combined data from PPCME2 and PLAEME (n = 124,308)

6.2 Dialect areas

The starting locations and speeds of dispersion of the respective linguistic contact settings differ among the dialectal areas. This means that the areas were affected differently by language contact. For example, whereas French found its way into medieval England through the southern dialects (cf. Rothwell 1983: 259–60), the Old Norse language entered the country through the northern dialects (cf. Pons-Sanz 2013: 6f.). Dialectal distribution of non-finite biases for the French and Norse-derived etymological sets may reveal more about the diachronic development of accommodation biases, as dialect areas relate to areas of longest and most intense contact. From this, we hypothesise that biases would be least strong in areas where contact originated or was most pervasive. A comparison of biases across different dialects (figure 3) will reveal any diatopic trends.

In figure 3, the dashed black line and the associated percentage given at its lower end show the baseline of non-finite usage of native English verbs for the data from each dialect area. Like in the analyses above, this is the comparandum to which the proportion of non-finite usage of loan verbs is compared for each dialect.

As figure 3 shows, our hypothesis holds for French loans to some degree, as the non-finite bias is least strong in Southern texts (at 43.73 per cent non-finite usage compared to 37.04 per cent for English verbs), where French initially entered English (see section 3). However, next lowest are the biases in Northern texts (56.45 per cent compared to 47.07 per cent for English verbs) and East Midlands texts (50.40 per cent compared to 40.19 per cent for English verbs), while West Midlands texts show the
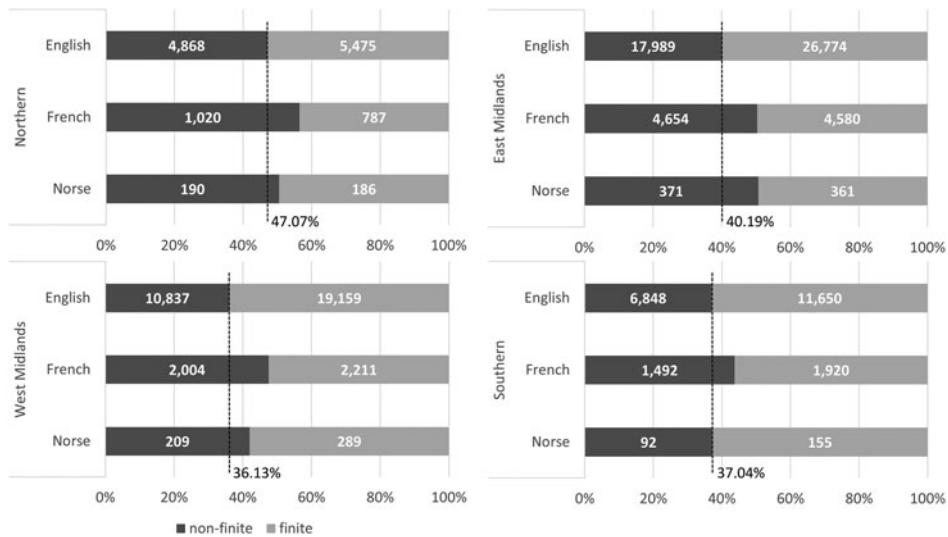
Figure 3. Finiteness distribution per dialect area for verbs of Norse-derived and French origin compared to the baseline of English verbs in combined data from PPCME2 and PLAEME
(n = 124,308)

strongest non-finite bias for French loans (47.54 per cent non-finite usage) as compared to English verbs (36.13 per cent). For Norse-derived verbs, a similar trend presents itself in the data, albeit with contrasting implications. Biases are weakest in Southern texts (at 37.25 per cent non-finite usage compared to 37.04 per cent for English verbs), followed in strength by biases in Northern texts (at 50.53 per cent non-finite usage compared to 47.07 per cent for English verbs) and West Midland texts (at 41.97 per cent non-finite usage compared to 36.13 per cent for English verbs), with East Midland texts showing the strongest non-finite bias for Norse-derived verbs with 50.68 per cent non-finite usage (compared to 40.19 per cent for English verbs), as figure 3 shows.

East Midlands dialect texts make up the largest share of data overall in the corpora used (see section 3.2, table 1) and originate from the Danelaw area, where Scandinavian influence was most intense and long-lived. Therefore, this finding is somewhat unexpected, as integration of loan verbs is hypothesised to be more advanced in high-contact areas. However, more than half of texts from the M1 subperiod, during which we would expect the highest biases for Norse verbs diachronically, are from the East Midlands dialect (52.94 per cent of M1 texts). Therefore, we might expect this higher relative bias. Moreover, the Northern and Southern dialect texts mostly stem from later Middle English (M3 and M4; see section 3.2, table 1) and there are no Northern texts from the M1 subperiod at all. Hence, the low accommodation biases for Norse-derived verbs in the Northern and Southern dialects may be a reflection of the diachronic distribution of texts rather than dialect alone.

7 Discussion and conclusion

7.1 Discussion

The above findings reveal significant accommodation biases for verbs entering into ME from both French and ON, but there exists a significant difference between the biases for French and Norse-derived verbs, as the biases are significantly weaker for ON than for French (see figure 1, p = 0.0215, Chi-square test). The strength of accommodation biases may, therefore, be directly affected by the typological closeness of replica and model language. However, as discussed above, this effect cannot be easily disentangled from the difference in temporal distance to the direct contact situation. Whereas French is at a peak point in its contact with English during the ME period, direct contact between ON and English subsides by the end of the OE period. In order to corroborate the effect of linguistic closeness of the languages in contact on the strength of accommodation biases at the smallest possible temporal distance to contact, accounting for the strength of the biases of those Norse-derived loan verbs attested earlier in OE data is a desideratum. Given the limitations of the extant OE data in accounting for the influx of Norse-derived lexis (see section 3), this merits an even stricter operationalisation of the etymological verb sets and dialectal distribution of attestations for future investigations.

As for the diachronic development of the biases, the data have shown that accommodation biases for French loan verbs are rather persistent throughout ME. For Norse-derived verbs, accommodation biases are persistent at first as well, but then decrease over time. This could be attributed to direct contact with Scandinavian already having ended by the early ME period, whereas contact with French had not. The comparison would benefit from including later diachronic data for French (e.g. Early Modern English) to assess whether biases for French weaken diachronically at a similar rate to those for Norse-derived verbs. The small number of existing data for Norse-derived non-cognate verbs which are already attested in OE should also be investigated to account for a decreased temporal distance for the Old English-Scandinavian contact situation to further enable comparison with French loan verbs in early ME.

Accommodation biases are also found to be regionally dependent, since biases for French loan verbs are stronger in areas with less intense contact than in areas with more intense contact (i.e. Southern). The data for Norse-derived verbs do not represent a clear picture across dialect areas, but the dialect representation of ME time periods is rather unequal (see sections 3 and 6.2). The low number of biases for Norse-derived verbs in texts from the East Midlands may be explained by 62.71 per cent of this text data being from later ME (M3 and M4). While the data contain no Northern texts for the M1 period, 67.51 per cent of the Northern data are from M2, representing earlier ME. This brings circumstantial evidence to our hypothesis that biases will be lower in high contact areas, even at a shorter temporal distance. The lack of early Northern texts may explain the low biases reported for either etymology in texts from this dialect area, as foreign lexis had become accommodated before occurring in the data. The weaker

biases of Norse-derived verbs in Southern texts may be due to only 1.71 per cent of Southern data being from the M1 period. Again, this allows for the possibility of verb accommodation being well under way before attestation in the data, even in this area of later and less intense contact with ON.

Additionally, the data reveal that the proportion of non-finite usage in native English verbs changes throughout the ME period (see figure 2): non-finite forms are overall more common in late ME, such as *make* ('to make') and *lawh* ('to laugh') in (10), than in early ME, which relied more on finite forms, such as *libbeþ* ('lives')**,** *healdeþ* ('holds')**,** *iualþ* ('befalls'), *leueþ* ('lives') and *sterfþ* ('dies') in (11).

(10) This son was so sobir of chere þat þere mit-e   no myrth
      this son was so sober of cheer that there might.PST.SBJV-SG no mirth
      make   him **lawh**
      make.INF him laugh.INF
      'This son was so calm in demeanour that no amusement could make him laugh.'
      (*Capgrave's Chronicle*, CMCAPCHR,57.753)

(11) Þo   þet **libb-eþ**   be fisike: hy   **heald-eþ**   þe
      those that live-PRS.3PL by physic they hold-PRS.3PL the
      mesure of ypocras   þet is lite an strait.
      measure of Hippocrates that is light and straight
      and hit **iual-þ**   ofte. þet þe ilke / þet be fisike **leu-eþ**:   be
      and it befall-PRS.3PL often that he who   by physic live-PRS.3PL by
      fizike **sterf-þ**
      physic die-PRS.3PL
      'Those that live by physic observe the measure of Hippocrates, which is little and narrow; and it often befalls that he that by physic lives by physic dies.'
      (*Ayenbite of Inwyt*, CMAYENBI,54.969–70)

The usage of periphrastic verbal structures such as *do*-support and modal verbs (see also *mite* in example (10)) increases drastically as of late ME (Görlach 2003: 97; Green 2017). Such structures typically rely on non-finite forms, as can be seen from example (12), where *blame* is supported by *do*, namely in *doth blame*.

(12) Do   he what-somever he wyll, no man do-th   **blame** [OF: bla(s)mer] hym.
      does he whatsoever   he will no man do.PRS-3SG blame.INF   him
      'He does whatever he wants to, no man blames him.'
      (*In Die Innocencium*, CMINNOCE,5.64)

Since these innovative structures heavily relied on non-finite structures, it is not unexpected that non-finite forms become increasingly common (see discussion in Shaw 2022: 160).

7.1.1 Lemma and frequency effects

Note that the general findings on accommodation biases in French and Norse-derived verbs should be interpreted in the light of some lemma and frequency effects. An individual lemma effect was identified in the Norse-derived verb set (n = 55 lemmata), and more specifically in the high-frequency lemmata. With a proportion of 43.30 per cent non-finite usages, the 5 most frequent lexemes in the dataset diverge considerably from the non-finiteness proportions for the Norse-derived verb set as a whole (46.28 per cent, see figure 1). These lexemes are *casten* (376 attestations), *foryeten* (174 attestations), *geten* (446 attestations), *geren* (108 attestations) and *forleten* (180 attestations).[16] Exactly because of their high frequency, these lemmata skew the findings for this variable, since they make up 69.27 per cent of all Norse-derived tokens in the data, and they show a significantly (p = 0.0001, Fisher's exact test) lower proportion of non-finite usage (43.30 per cent) than the other tokens (incl. low-frequency tokens) of Norse-derived verbs do at 52.98 per cent. The low rate of non-finite usage for the five high-frequency lemmata brings down the general proportion of non-finites in the Norse-derived verb set as well. Despite this significant lemma effect, the proportions of non-finite forms for the five most frequent Norse-derived lemmata is still significantly higher than the proportions of non-finite forms for the English baseline (p < 0.0025, Chi-square with Yates' correction). From this one may infer that increased usage frequency of Norse-derived verbs seems to aid the weakening of accommodation biases but does not cancel them out altogether.

Another effect in the data is the tendency of low-frequency lemmata to be used non-finitely. This finding corroborates the interaction effect found in Shaw & De Smet (2022: 11), where lemma frequency and French origin interact, meaning that the non-finite bias in French loan verbs is even stronger in low-frequency items than in high-frequency items. As suggested by De Smet & Shaw (2024: 7–8), low-frequency items are subject to stronger biases than high-frequency items since language users try to decrease the processing cost of low-frequency items.

The non-finite bias for French occurred regardless of lemma frequency (i.e. even in high-frequency items), but showed significant increase in low-frequency lemmata. This is in contrast to the non-finite bias for Norse-derived verbs, which was not significantly stronger in low-frequency items. Only when compared to high-frequency Norse-derived lemmata, which show significantly lower bias than the overall verb set, is the same trend corroborated. In summary, high-frequency French loans still show a significant bias towards non-finite forms (Shaw 2022), as do high-frequency Norse-derived loans when compared to the English baseline, albeit to a lesser extent (p = 0.0025, Chi-square test). This may suggest that high-frequency Norse-derived verbs are still somewhat easier to integrate into ME than high-frequency French loan verbs, and as a low-frequency verb is harder to integrate than a high-frequency verb, it is more likely to be biased towards a

---

[16] The next most frequent lemma, *reisen*, had 58 attestations.

non-finite form. An example of a low-frequency Norse-derived verb used non-finitely is given in (13), where the Norse-derived verb *skerrenn* ('to scare') occurs in the bare infinitive.

(13) ȝiff he seoþ þe mann forrdredd, He wil-e   himm
       if   he sees the man   frightened he will.PRS.SBJV-3SG him
       skerr-enn [ON: skirra]   mare […]
       scare-INF   more […]
       'If he [the devil] sees a man frightened, he will scare him more […].'
       (*Ormulum*, CMORM,I,132.1120)

7.1.2 Limitations

This case study is unavoidably subject to a number of limitations. First, we have not carried out regression analyses, which were conducted in Shaw & De Smet (2022), because the imbalanced nature of the data across dialects and periods in time discourages the usage of regression analysis as a statistical technique. Second, the dataset includes a number of translated texts from French and Latin originals with varying degrees of literality, and we have not controlled for the possibility of interference effects. However, since the dataset also includes non-translated texts, possible effects may already have been balanced out.

7.2 Conclusion

This study has investigated the effects of typological closeness of languages in contact as well as the temporal distance to the period of contact on constraints in loan word accommodation. Through a quantitative corpus study, the presence and strength of loan word accommodation biases in French and Norse-derived loan verbs in ME were systematically compared.

As hypothesised, typological closeness of languages in contact is inversely proportional to the strength of the accommodation biases in ME. This may strengthen the argument that linguistic closeness facilitates the borrowing of more complex categories (e.g. Meillet 1921; Moravcsik 1975; Winford 2003: 51ff; cf. Johanson 2002). Additionally, this study has confirmed the finding by De Smet & Shaw (2024: 5) that accommodation biases can weaken over time, namely in Norse-derived verbs for which the temporal distance to direct contact is longer than for the French verbs.

At a general level, this study has contributed to filling the research gap on constraints on loan word accommodation and on the morphosyntactic integration of loan words. For Norse-derived verbs in English specifically, this study has provided insight into loan word accommodation, which adds to general research on loan verbs (e.g. Wohlgemuth 2009). As with French verbs (cf. Shaw & De Smet 2022), the integration of Norse-derived verbs into ME is constrained by some factors, such as typological closeness, time distance to the period of contact, and the contact area under investigation. Investigating Norse-derived verbs has also shed light on the nature of

loan word accommodation biases across different contact situations where English is the replica language.

Additional research is needed to properly distinguish between the effects of temporal distance and typological closeness. Furthermore, the findings on typological closeness would benefit from further research into different model and replica language pairings. The question remains as to whether typological closeness facilitates the ease and speed of the morphosyntactic integration of loan verbs independently of time.

Authors' addresses:

Anglistik IV
University of Mannheim
B 6, 30–32, Room 027
68159 Mannheim
Germany
wiebke.juliane.elter@uni-mannheim.de

Blijde Inkomststraat 21 bus 3301
3000 Leuven
Belgium
marliekeshaw@gmail.com

References

Adler, Franz. 1951. Yates' correction and the statisticians. *Journal of the American Statistical Association* 46(256), 490–501. https://doi.org/10.1080/01621459.1951.10500804.

Bowern, Claire. 2013. Relatedness as a factor in language contact. *Journal of Language Contact* 6(2), 411–32. https://doi.org/10.1163/19552629-00602010.

Dance, Richard. 2003. *Words derived from Old Norse in Early Middle English: Studies in the vocabulary of the South-West Midland texts*. Tempe, AZ: Arizona Center for Medieval and Renaissance Studies.

Dance, Richard. 2011. 'Tomarʒan hit is awane': Words derived from Old Norse in four Lambeth Homilies. In Jacek Fisiak & Magdalena Bator (eds.), *Foreign influences on medieval English*, 77–127. Frankfurt am Main: Peter Lang.

Dance, Richard. 2012. 'Tor for to telle': Words derived from Old Norse in *Sir Gawain and the Green Knight*: An etymological survey. In Judith Jefferson & Ad Putter (eds.), *Multilingualism in medieval Britain (c. 1066–1520): Sources and analysis*, 41–58. Turnhout: Brepols.

Dance, Richard. 2018. Words derived from Old Norse in *Sir Gawain and the Green Knight*: An etymological survey. *Transactions of the Philological Society* 116(S2), 1–600. https://doi.org/10.1111/1467-968X.12148_02.

Dance, Richard, Sara Pons-Sanz & Brittany Schorn. 2019. *The Gersum Project: The Scandinavian influence on English vocabulary* (Cambridge, Cardiff and Sheffield). www.gersum.org (accessed 1 February 2022 – 20 November 2023).

Dekeyser, Xavier. 1986. Romance loans in Middle English: A re-assessment. In Dieter Kastovsky & Aleksander Szwedek (eds.), *Linguistics across historical and geographical boundaries: In honour of Jacek Fisiak*, vol. 1, 253–65. Berlin and New York: Mouton de Gruyter.

De Smet, Hendrik. 2014. De integratie van Engelse leenwerkwoorden in het Nederlands. In Freek Van de Velde, Hans Smessaert, Frank Van Eynde & Sara Verbrugge (eds.), *Patroon en argument: Een dubbelfeestbundel bij het emeritaat van William Van Belle en Joop van der Horst*, 75–87. Leuven: Leuven University Press.

De Smet, Hendrik & Marlieke Shaw. 2024. Missing link: Code-switches, borrowings, and accommodation biases. *Linguistic Vanguard*. https://doi.org/10.1515/lingvan-2023-0088

Durkin, Philip. 2014. *Borrowed words*: *A history of loanwords in English*. Oxford and New York: Oxford University Press.

Einhorn, Elsabe. 1975. *Old French: A concise handbook*. Cambridge: Cambridge University Press.

Finkenstaedt, Thomas & Dieter Wolff. 1973. *Ordered profusion. Studies in dictionaries and the English lexicon*. Heidelberg: Carl Winter.

Görlach, Manfred. 2003. Syntax. In Manfred Görlach (ed.), *Introduction to Early Modern English*, 95–135. Cambridge: Cambridge University Press.

Grant, Anthony P. 2009. Loanwords in British English. In Martin Haspelmath & Uri Tadmor (eds.), *Loanwords in the world's languages: A comparative handbook*, 360–83. Berlin and New York: Mouton de Gruyter.

Green, Clarence. 2017. *Patterns and development in the English clause system: A corpus-based grammatical overview*. Singapore: Springer.

Harris, Alice & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Haspelmath, Martin & Uri Tadmor (eds.). 2009. *Loanwords in the world's languages: A comparative handbook*. Berlin and New York: Mouton de Gruyter.

Hug, Sibylle. 1987. *Scandinavian loanwords and their equivalents in Middle English*. Bern, Frankfurt am Main and New York: Peter Lang.

Ingham, Richard. 2012. *The transmission of Anglo-Norman: Language history and language acquisition*. Amsterdam: John Benjamins.

Ingham, Richard. 2020. How contact with French drove patient-lability in English. *Transactions of the Philological Society* 118(3), 447–67. https://doi.org/10.1111/1467-968X.12199

Johanson, Lars. 2002. Contact-induced change in a code-copying framework. In Mari C. Jones & Edith Esch (eds.), *Language change: The interplay of internal, external and extra-linguistic factors*, 285–313. Berlin and New York: Mouton de Gruyter.

Kroch, Anthony, Ann Taylor & Beatrice Santorini. 2000–. *The Penn–Helsinki Parsed Corpus of Middle English* (PPCME2). Department of Linguistics, University of Pennsylvania. www.ling.upenn.edu/ppche/ppche-release-2016/PPCME2-RELEASE-4 (accessed 1 February 2022 – 1 July 2022). Currently distributed by the Linguistic Data Consortium as part of the Penn Parsed Corpora of Historical English.

Laing, Margaret. 2013–. *A Linguistic Atlas of Early Middle English, 1150–1325*, version 3.2. Edinburgh: University of Edinburgh. www.lel.ed.ac.uk/ihd/laeme2/laeme2.html (accessed 20 November 2023).

Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Lewis, Robert E. *et al*. (eds.). 1952–2001. *Middle English dictionary*. Online edition in *Middle English Compendium*, eds. Frances McSparran, Nigel Kerr, James Moske, Christina Powell, John Price-Wilkin & Paul Schaffner. Ann Arbor, MI: University of Michigan Press, 2000–18. http://quod.lib.umich.edu/m/middle-english-dictionary/ (accessed 1 February 2022 – 6 January 2023).

Matras, Yaron. 2020. *Language contact*, 2nd edn. Cambridge: Cambridge University Press.

Meillet, Antoine. 1921. *Linguistique historique et linguistique générale* (Collection Linguistique), vol. VIII. Paris: Édouard Champion.

Moravcsik, J. M. Edith. 1975. *Understanding language: A study of theories of language in linguistics and in philosophy*. The Hague: Mouton de Gruyter.

Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Percillier, Michael. 2016. Verb lemmatization and semantic verb classes in a Middle English corpus. *13th Conference on Natural Language Processing (KONVENS 2016)*, 209–14. www.linguistics.rub.de/konvens16/pub/26_konvensproc.pdf

Percillier, Michael & Carola Trips. 2020. Lemmatising verbs in Middle English corpora: The benefit of enriching the *Penn–Helsinki Parsed Corpus of Middle English 2* (PPCME2), the *Parsed Corpus of Middle English Poetry* (PCMEP), and *A Parsed Linguistic Atlas of Early Middle English* (PLAEME). *12th Language Resources and Evaluation*, 7170–8. Marseilles: European Language Resources Association. www.aclweb.org/anthology/2020.lrec-1.886

Pons-Sanz, Sara M. 2007. *Norse-derived vocabulary in late Old English texts: Wulfstan's works, a case study*. Odense: University Press of Southern Denmark.

Pons-Sanz, Sara M. 2013. *The lexical effects of Anglo-Scandinavian linguistic contact on Old English*, vol. 1. Turnhout: Brepols.

Poplack, Shana, David Sankoff & Christopher Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26(1), 47–104. https://doi.org/10.1515/ling.1988.26.1.47.

Proffitt, Michael (ed.). 2000–. *OED Online. Oxford English dictionary*, 3rd edn. www.oed.com (accessed 1 February 2022 – 6 January 2023).

Randall, Beth. 2010. *CorpusSearch* (2.003.00). http://corpussearch.sourceforge.net (accessed 1 February 2022 – 1 August 2022).

Rissanen, Matti, Merja Kytö, Leena Kahlas-Tarkka, Matti Kilpiö, Saara Nevanlinna, Irma Taavitsainen, Terttu Nevalainen & Helena Raumolin-Brunberg. 1991. *The Helsinki Corpus of English Texts*. https://varieng.helsinki.fi/CoRD/corpora/HelsinkiCorpus/ (accessed 1 February 2022).

Rothwell, William. 1983. Language and government in medieval England. *Zeitschrift für französische Sprache und Literatur* 93(9), 258–70.

Rothwell, William. 1996. The Anglo-French element in the vulgar register of Late Middle English. *Neuphilologische Mitteilungen* 97(4), 423–36.

Shaw, Marlieke. 2022. English phrases, French verbs: Causes and consequences of loan word accommodation biases. PhD dissertation, KU Leuven. https://researchportal.be/en/publication/english-phrases-french-verbs-causes-and-consequences-loan-word-accommodation-biases

Shaw, Marlieke & Hendrik De Smet. 2022. Loan word accommodation biases: Markedness and finiteness. *Transactions of the Philological Society* 120(2), 201–17. https://doi.org/10.1111/1467-968X.12233

Sijs, Nicoline van der. 2005. *Groot leenwoordenboek*. Utrecht: Van Dale Lexicografie.

Thomason, Sarah Grey. 2001. *Language contact: An introduction*. Edinburgh: Edinburgh University Press.

Thomason, Sarah Grey & Terrence Kaufman. 1991. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Townend, Matthew O. 2002. *Language and history in Viking Age England: Linguistic relations between speakers of Old Norse and Old English*. Turnhout: Brepols.

Townend, Matthew O. 2006. Contacts and conflicts: Latin, Norse, and French. In Lynda Mugglestone (ed.), *The Oxford history of English*, 61–82. Oxford: Oxford University Press.

Truswell, Robert, Rhona Alcorn, James Donaldson & Joel Wallenberg. 2018. *A parsed linguistic atlas of Early Middle English*. University of Edinburgh. https://datashare.ed.ac.uk/handle/10283/3032 (accessed 1 February 2022 – 1 July 2022).

Truswell, Robert, Rhona Alcorn, James Donaldson & Joel Wallenberg. 2019. A parsed linguistic atlas of Early Middle English. In Rhona Alcorn, Joanna Kopaczyk, Bettelou Los & Benjamin Molineaux (eds.), *Historical dialectology in the digital age*, 19–38. Edinburgh: Edinburgh University Press.

Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*, 7th edn. The Hague and Paris: Mouton de Gruyter.

Winford, Donald. 2003. *An introduction to contact linguistics*. Oxford: Blackwell.

Wohlgemuth, Jan. 2009. *A typology of verbal borrowings*. Berlin and New York: Mouton de Gruyter.