Language Variation and Change (2023), **35**, 199–221
doi:10.1017/S0954394523000108

CAMBRIDGE UNIVERSITY PRESS

Where did *wer* go? Lexical variation and change in third-person male adult noun referents in Old and Middle English

James M. Stratton

University of British Columbia, Canada
Email: james.stratton@ubc.ca

Abstract

The present study uses variationist quantitative methods to examine the evolution of the semantic field of third-person male adult noun referents from Old English to Middle English, covering a time depth of approximately six hundred years. Results show a shift from the favored variant *wer* in Old English to *man* in Middle English, with the diachronic change in frequency following a prototypical s-shaped distribution. Although the replacement seems to take centuries to be complete, lexical frequency and written transmission are proposed as influential explanatory factors, and a homonymic clash is suggested to have accelerated the process of replacement in Middle English. Text type and text origin contribute to variation, with alliteration significantly influencing lexical choices in Old English verse texts. When combined with findings from recent synchronic work, this study highlights a heterogeneously structured semantic domain, which has undergone lexical replacement and change over time, providing some evidence for the applicability of s-shaped patterns for lexical change.

Keywords: Old English; Middle English; lexical replacement; lexical variation; s-curve

To refer to a male adult, speakers of English have several lexical items to choose from, such as *man*, *guy*, *chap*, *dude*, *fella*, *bloke*, *gent*, *gentleman*, and *geezer*. Data from the Spoken British National Corpus (BNC; Love, Dembry, Hardie, Brezina, & McEnery, 2017) indicate that these variants are widespread and can be found in a range of referentially comparable contexts, as in (1). Because these variants have entered the English language at different intervals in time, this domain exhibits lexical layering, which, like other aspects of variation, may be influenced by linguistic and external conditioning factors.

(1)
a. he is quite a big **dude** [BNC2014, S0603]
b. he's a big **bloke** [BNC2014, S0238]
c. erm, he's a big **guy** [BNC2014, S0417]
d. Schwarzenegger […] mm you know that big **chap**? [BNC2014, S0369]
e. there was a big **man** eating [BNC2014, S0654]
f. he's a lovely **fella**, ain't he? [BNC2014, S0278]

While this semantic field has been explored in Present Day English (Franco & Tagliamonte, 2021; Tagliamonte, 2022) and a list of variants pertaining to this semantic field has been compiled for Old and Middle English (Elsweiler, 2011; Grygiel, 2006; Kleparski, 2003, 2005; Stenroos, 2002), how this domain has evolved over time, especially in the earlier stages of the English language, remains unclear.

One of the reported hallmarks of linguistic change is the presence of an s-shaped distribution where incoming forms are adopted at a rate of slow to fast to slow (Bailey, 1973:77; Weinreich, Labov, & Herzog, 1968:113-114). A comparison of the current system of variants (e.g., *man*, *guy*) with variants used in earlier stages of the language (e.g., *wer* 'man') clearly shows lexical change within this semantic field. However, whether lexical change follows a traditional s-curve trajectory remains to be investigated. Although lexical replacement is expected to follow an s-curve distribution (Blythe & Croft, 2012:278-279; Chambers, 2002:361), to date s-curve patterns have been modeled predominantly on phonetic (e.g., Labov, 1994), morphological (e.g., Nevalainen, 2015), and discourse-pragmatic features (e.g., Tagliamonte & Smith, 2021). While s-curve patterns for lexical change can be found in the literature (e.g., Chambers, 1995), they are typically based on apparent time as opposed to real time data. However, in work on short-term high density lexical change, Grieve, Nini, and Guo (2017) found s-shaped patterns for several "emerging words" such as *baeless* and *fleek*, suggesting that lexical change also follows a prototypical s-curve trajectory. Nevertheless, research on long-term s-curve patterns for lexical change is lacking. Given that the semantic field of third-person male adult noun referents is rife with variation, examining how this system has evolved holds promise for insights into the factors governing lexical variation and change in real time. Recent work on Ontario English shows that this lexical domain is in flux, with factors such as gender and age influencing speakers' choices (Franco & Tagliamonte, 2021; Tagliamonte, 2022). Against the backdrop of this work, the present study examines the system of third-person male adult noun referents in earlier stages of the English language, specifically in Old and Middle English.

Two research questions are addressed. First, what was the distribution of third-person male adult noun referents in Old and Middle English? A quantitative comparison of this domain over time can document changes within this semantic field while also providing a platform to test patterns of lexical change. Second, based on the extant metadata, is there any evidence to suggest that the use of third-person male adult noun referents was conditioned, constrained, or influenced by any attested factors of variation? To answer these questions, the *Helsinki Corpus of English Texts* (Kytö, 1996; Rissanen, Kytö, Kahlas-Tarkka, Kilpiö, Nevanlinna, Taavitsainen, Nevalainen, & Raumolin-Brunberg, 1991) was used as the principal

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

source of linguistic data, as it contains texts from Old English and Middle English, as well as metadata for potentially influential conditioning factors, such as TEXT TYPE and TEXT ORIGIN. Although larger corpora are available, given the size of the system of third-person male adult noun referents, with over ten thousand tokens of *man* attested in the *Helsinki Corpus of English Texts* alone (Rauer, 2017:142-143), to ensure confidence in the circumscription of the variable context, larger corpora were avoided.

Background

Lexical Variation

In recent decades, variationist quantitative methods have been used to examine a range of discourse-pragmatic phenomena, such as intensifiers (e.g., Stratton, 2020, 2022a; Stratton & Sundquist, 2022; Tagliamonte, 2008), quotatives (e.g., Tagliamonte & D'Arcy, 2004), general extenders (e.g., Cheshire, 2007), and evidentiality markers (e.g., Tagliamonte & Smith, 2021). Although geography is often reported as the predominant explanatory factor for lexical variation, lexis is highly structured along the axes of social and stylistic variation. In recent work, several lexical sets have been explored, such as *dinner* versus *tea* (Jankowski & Tagliamonte, 2019), words of profanity (Tagliamonte & Jankowski, 2019), adjectives of strangeness (Tagliamonte & Brooke, 2014), and adjectives of positive evaluation (Stratton, 2022b; Tagliamonte & Pabst, 2020).

Based on the premise that lexical choices are influenced by similar conditioning factors that operate on grammatical and phonological variation, Tagliamonte and colleagues probed the sociolinguistic underpinnings of the system of third-person male adult noun referents in Ontario English (Franco & Tagliamonte, 2021; Tagliamonte, 2022). The authors found that *guy* was making traction over competing variants such as *man*, a change led predominantly by young men. While *guy* was used more frequently by men than women, other factors such as socioeconomic status were also found to play a role. Following the "Uniformitarian Principle" (Lyell, 1830-1833; cf., Labov, 1972:275), one might expect similar conditioning factors to have affected this system diachronically. Although the lack of similar types of sociolinguistic metadata in the earlier stages of the English language rules out a quantitative analysis of the effect of social factors on this system in Old and Middle English, this scholarship serves as a point of departure for the present analysis.

History of third-person male referents

Variants for describing third-person male adult noun referents have a long history of lexical replacement in the English language. According to *A Thesaurus of Old English* (2017) and *The Historical Thesaurus of English* (2023), there are at least twenty-five attested lexical items that denote "man" in Old English, including, but not limited to *ceorl*, *carlmon*, *esne*, *freca*, *folcagende*, *folcwer*, *guma*, *gumrinc*, *hæle*[*þ*], *hyse*, *leod*, *mæcg*, *man*, *scealc*, *wer*, *woruldman*, *wiga*, *wæpnedmann*, and *wæpenmann*. While some of these variants (e.g., *gumrinc*) are reportedly more frequent in Old English verse (*A Thesaurus of Old English*, "13.02.10.01 A man, warrior"), others (e.g., *wer*, *guma*, *man*) are found in both prose and verse (Kleparski, 2003:49; Stenroos,

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

2002:382-383). *Wer*, *guma*, and *man* are variants that can refer to a male adult in Old English, but the word form *man* has additional functions: it can be used gender-inclusively to refer to both male and female referents (Curzan, 2003:64-65), it can be used to refer to human beings (Rauer, 2017:139-140), and it can be used as an indefinite pronoun, corresponding in translation to "one" (Raumolin-Brunberg & Kahlas-Tarkka, 1997).[1] Examples of *wer*, *guma*, and *man* referring to male referents are provided in (2).[2]

(2) a. *Ond on ðone ylcan dæg Crist gereorde fif ðusenda* ***wera*** *of fif*
and on that same day Christ fed five thousand man-GEN.PL of five
hlafum ond of twam fixum, eac wifum ond cildum
bread-DAT.PL and of two fish-DAT.PL, also woman-DAT.PL and child-DAT.PL
þara wæs ungerim
which.GEN.PL was uncountable
'And on that same day, Christ fed five thousand **men**, with five loaves of bread and two fish. In addition, he also fed women and children, of which there were many'
[Old English Martyrology, 950-1050]

b. *ðonne onwæcneð eft wineleas* ***guma***
then awakens again friendless man
'Then the **man** without any friends woke up'
[The Wanderer, 950-1050]

c. *on þære fyrde wæron þe ferdon fram Egipte, sixhund þusend*
in the army were REL traveled from Egypt six-hundred thousand
manna *butan wifum 7 cildum*
man-GEN.PL except woman-DAT.PL and child-DAT.PL
'In that army, there were 600,000 **men** who travelled from Egypt, and that number does not include women and children'
[Ælfric's Letter to Sigeweard, 1050–1150][3]

By Middle English, only half of the Germanic words for "man" (e.g., *beorn*, *cerl*, *freca*, *guma*, *hearra*, *leod*, *man*, *rinc*, *scealc*, *secg*, *wæpenmann*, *wer*, *wiga*) are reported to have remained in use (Stenroos, 2002:385). Contact with Anglo-Norman led to the emergence of new variants through lexical borrowing, such as *sire* 'man' (Kleparski, 2005:48) and *sergeant* 'servant/serving individual' (Kleparski, 2003:51). Modern English *gentleman* emerged during Middle English by compounding the French loanword *gentil* 'noble' with Germanic *mon* 'man' (literally 'nobleman'). Meanwhile, other variants underwent semantic shifts, such as Old English *æþeling* 'prince/nobleman' which became Middle English *hathel* 'man' (*Middle English Dictionary* [MED] 2021; *hathel*, n.). However, like in Old English, it is possible that many variants were affected by text type, with some variants occurring more frequently in verse than in prose due to metrical and alliterative demands. According to Eduard Sievers's *Altgermanische Metrik* (1893), lines in Old English verse, called *Langzeile*, consist of two short lines (*Kurzzeile*), and at least one lexical item in each short line alliterates. This alliterative tradition was transmitted into

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Middle English to a lesser extent, with rhyme emerging as a new feature of verse. Examples in (3) from the Middle English text *Sir Gawain and the Green Knight* illustrate the rich variation present within the semantic field, seemingly due to the alliterative demands. According to the *Oxford English Dictionary* (OED, 2021), few of the Middle English variants survived into Early Modern English. For instance, *tulk* (ON *tulk-r* 'interpreter') and *renk* (ON *rekk-r* 'warrior,' OS *rink*) were rarely used after the mid-sixteenth century, and *gome* (OE *guma*), *freke* (OE *freca*), *berne* (OE *beorn* 'man of valor,' ON *bjǫrn* 'bear'), *schalk* (OE *scealc* 'servant/man,' ON *skalk-r* 'slave/servant'), and *lede* (OE *leoda* 'people,' German *Leute* 'people') are not attested after the seventeenth century.[4]

(3) a. *and talk wyth þat ilk* ***tulk*** *þe tale þat me lyste*
and talk with that same man which tale that me desired
'and talk with the same **man** about whatever tale is pleasing to me'

b. *forþy goude Sir Gawayn let þe* ***gome*** *one*
therefore good Sir Gawain let the man alone
'so, good Sir Gawain, leave the **man** alone'

c. *for he is a* ***mon*** *methles and mercy non vses*
for he is a man measureless and mercy none use
'because he is a violent **man** and not merciful'

d. *and ȝe ar a* ***lede*** *vpon lyue þat I wel louy*
and you are a man upon life that I well love
'and you are a mortal **man** that I love very much'

e. *Þe* ***burne*** *þat rod hym by bede his mayster abide*
the warrior that rode him by commanded his master abide
'The **man** who rode past him, commanded his master to wait'

f. *here ar no* ***renkes*** *vs to rydde, rele as vus likez*
here are no men us to ride, reel as us pleases
'There are no **men** here to stop us from fighting, as it pleases us'

g. *Þat oþer* ***schalk*** *wyth a schunt þe schene wythhaldez*
the other man with a jerk the shiny withholds
'The other **man** withdrew the blade with a sudden jerk'

h. *wat, is þis Arþureȝ hous, quoþ þe* ***haþel*** *þenne*
what is this Arthur's house said the man then
'Is this King Arthur's house, said the man?'

The linguistic variable

In traditional variationist work, the linguistic variable is defined as "alternate ways of saying 'the same' thing" (Labov, 1972:188). Although this definition was originally applied to phonological variation, over time this concept was extended to the study of grammatical and lexical variation (Terkourafi, 2011). In early work on distributional semantics, Firth (1957:11) highlighted the importance of context when determining referential meaning, and in modern studies of variation, forms do not need to have the

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

exact same denotation, but overlapping uses or a shared history is often a prerequisite to be treated as variants of the same thing. So long as the analyst identifies and includes the contexts in which the referential meaning is equivalent, and removes instances in which they are not, a two-step process known as the Principle of Accountability (Labov, 1969:737-738), and circumscription of the variable context (Labov, 1969:729), the linguistic variable can be used to study variation outside of phonology.

Although there are semantic nuances between Middle English nouns such as *segge*, *hathel*, *freke*, *wyʒe*, and *mon*, at the discourse level, they can have the same referential meaning, that is, in some contexts they can be used to refer to the same male adult. The examples in (4) from *Sir Gawain and the Green Knight* illustrate that these forms could be used interchangeably, as they appear in the same contexts to refer to the same green knight (i.e., *quoþ the …* 'said the…'). Since these nouns occurred in the same contexts to refer to the same referent, these variants were, at least at the level of discourse and in specific contexts, referentially equivalent.

(4) a. *hit is sothe, quoþ the* ***segge***
it is sooth said the warrior
'it is true, said the **man**'

b. *is þis Arþureʒ hous, quoþ the* ***hathel***
is this Arthur's house said the nobleman
'is this Arthur's house, said the **man**'

c. *now iwyss, quoþ the* ***wyʒe***
now I know said the wise one
'now on my word, said the **man'**

d. *yet firre, quoþ the* ***freke***
yet further said the warrior
'yet further, said the **man'**

e. *madame, quoþ the myry* ***mon***
madam said the merry man
'my lady, said the merry **man'**

As Tagliamonte and Brooke (2014:11-12) point out, using semantic fields as a foundation for circumscribing the variable context is not new. For instance, Sankoff, Thibault, and Bérubé (1978) analyzed the semantic field of verbs which mean "to dwell," which led to the notion of "weak complementarity": the idea that linguistic variables can be identified through distributional properties and distribution across a speech community (Sankoff & Thibault, 1981:207). Recent studies have followed in this tradition when analyzing lexical variation (e.g., Stratton, 2022b; Tagliamonte & Brooke, 2014; Tagliamonte & Pabst, 2020). The concept of a semantic field, however, predates the variationist tradition and has its roots in structuralist semantics (Trier, 1931), with the important distinction between semasiology and onomasiology (Geeraerts, 2010). Semasiology, a concept that emerged in pre-structuralist work, "considers the isolated word and the way its meanings are manifest" (Baldinger, 1980:278) and is therefore concerned with polysemy, that is, the

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

various meanings a given word form can have (Geeraerts, 2010:84). In contrast, onomasiology "looks at the designations of a particular concept" (Baldinger, 1980:278), which can be conceived as studying varying levels of synonymy (Geeraerts, 2010:84). Studying the variants used to denote third-person male adult noun referents can therefore be viewed as onomasiology, which, unlike semasiology, starts with the concept and examines how it can be expressed. In a variationist framework, variants within a semantic field, lexical field, or onomasiological set can be studied as a linguistic variable. Therefore, the present study uses the notion of a semantic field, following "weak complementarity," to study the system of third-person male adult noun referents in Old and Middle English.

Methodology

Data

To examine the semantic field of third-person male adult noun referents in Old English, the *Helsinki Corpus of English Texts* (Kytö, 1996; Rissanen et al., 1991) was used, which contains 413,250 words, divided into four subperiods: O1 (2,190 words), O2 (92,050 words), O3 (251,630 words), and O4 (67,380 words). While many texts from the same corpus were also included for the analysis of variants in Middle English, because the corpus of Middle English is substantially larger than the Old English counterpart, to ensure confidence in the circumscription of the variable context only a sample of the texts from the Middle English part of the corpus was used. However, because there are fewer verse texts in the Middle English corpus, to ensure that TEXT TYPE could be included as a factor in the analysis, additional verse texts were added from Sisam (1928). The Middle English dataset in the present study therefore had three subperiods: M1 (48,336 words), M2 (30,554 words), and M3 (50,069 words), with M1 and M2 representing Early Middle English, and M3 representing Late Middle English.[5] The texts included for the Middle English analysis were as follows: M1: *Ormulum*, *Hali Meidhad*, *Peterborough Chronicle*, *Layamon's Brut*; M2: *Dame Sirith*, *Man in the Moon*, *Havelok*, *The Thrush and the Nightingale*, *Sir Orfeo*, *Ayenbite of Inwyt*; M3: *The General Prologue to the Canterbury Tales*, *The Wife of Bath's Prologue*, *The Dancers of Colbek*, *Sir Gawain and the Green Knight*, *The Pearl*, *Henry V: Letters to a Bishop*, *The New Testament: Wycliffe*, Chaucer's *Astrolabe*, *The Cloud of Unknowing*, *John Travisa: Polychronicon*.[6]

Circumscribing the variable context

A list of third-person male adult noun referents for Old and Middle English was compiled through previous literature (Elsweiler, 2011; Grygiel, 2006; Kleparski, 2003, 2005; Stenroos, 2002), dictionaries (*Bosworth-Toller* [2014]; *Middle English Dictionary* [2021], *Oxford English Dictionary* [2021]), and thesauruses (*The Historical Thesaurus of English* [2023], *A Thesaurus of Old English* [2017]). Then, search queries were run to find these variants in the data. Since word forms are not lemmatized in the Helsinki corpus, a list of spelling variants and inflectional forms was compiled with the aid of the *Dictionary of Old English* (Cameron, Amos, & Healey, 2018) and was subsequently searched for in the corpus data.

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Tokens were then downloaded and manually inspected for the removal of any nonequivalent instances.

Since the variable context was circumscribed to third-person male adult noun referents, variants such as OE *man* 'man' were only included in the analysis when they unambiguously referred to a male adult, as in (5). The presence of names, as in (5a), as well as the sociohistorical context, helped determine the gender of the referent. For instance, in (5b) it is evident that the referent *Priam* is male because he is a *preost* 'priest,' a role traditionally confined to men. Similarly, in (5c) the referent is King Arthur, referred to as a *god mon* 'good man,' who is presumably male. In (5d), the *six men* are biologically male as we are told they were castrated (their *stanes* 'testicles' were removed).[7] In contrast, examples in (6) were not included in the envelope of variation. While differentiating the gender-specific, gender-inclusive, and indefinite use of *man* is no simple task (Curzan, 2003:135; Rauer, 2017), indicators such as a preceding negative particle (e.g., *no mon here vnmanerly þe mysboden habbez* 'no one here has treated you in an unmannerly fashion') or indefinite adjectives (e.g., *forþam nat nænig man* 'therefore, nobody knows') helped disambiguate possible readings. Special attention was taken to ensure that the anachronistically homophonous and semantically nonequivalent form *mān* 'crime,' identified by a macron in editorial textual editions, was excluded. A number of functionally nonequivalent uses, such as the indefinite use of *man*, as in (6a), vocatives of address, as in (6b), as well as instances where the referential meaning was different, such as 'husband/boyfriend' in (6c), were excluded.[8] Instances in which *wer* meant 'wergild,' that is, a compensation tariff, were also excluded on grounds of being semantically nonequivalent.[9] *Eorl* 'earl' was not included when used as a term of address or rank (e.g., *Harold eorl* 'Earl Harold,' *Godwine eorl* 'Earl Godwin').

(5) a. *ða wæs Apollonius gehaten sum iung* ***man*** *se wæs swiðe welig*
then was Apollonius called some young man who was very wealthy
and snotor
and wise
'there was a young **man** called Apollonius, who was very wealthy and wise'
[Apollonius of Tyre, 950-1050]

b. *þa cleopede þe king Piram, ænne preost mæren*
then called the King Priam, a priest famous
he wes swiðe wis ***mon*** *and witful on bokken*
he was very wise man and witful in books
'Then the king summoned Priam, a famous priest, he was a very wise **man** and well read'
[Layamon, 1150-1250]

c. *Arður wes an weorlde wis king and riche god* ***mon*** *and griðful*
Arthur was one world wise King and rich good man and peaceable
'In this world, Arthur was a wise and powerful King, a good **man** and amicable'
[Layamon, 1150-1250]

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

d. *six* ***men*** *spilde here ægon 7 of here stanes*
six men deprived their eyes and of their stones
'six **men** had their testicles and their eyes removed'
[Peterborough Chronicle, 1150-1250, written under the entry for 1125AD]

(6) a. *hwæðer ænig* ***mon*** *be norðan þæm lande westenne bude*
whether any man by north the-DAT.SG land waste lived
'whether **anyone** lived to the north of the wasteland'
[Ohthere & Wulfstan, 950-1050]

b. *ne sorga, snotor* ***guma****, selra bið æghwæm þæt he his freond wrece*
no worry, wise man, better is for each that he his friend avenge
'do not worry, wise **man**, it is better for everyone that he avenges his friend'
[Beowulf, 950-1050]

c. *gif* ***ceorl*** *7 his wif bearn hæbben gemæne*
if churl and his woman child have together
'if a **husband** and his wife have a child together'
[Laws of Ine, 850-950]

For texts translated from Latin, comparisons between the Old English variants and the Latin counterparts aided in identifying the gender of the referents. For instance, in (7) it is evident that the people the *idesa* 'women' have not slept with are male, not only from context, but also because *vir* 'man' was found in the Latin text (*habeo duas filias, quae necdum cognoverunt virum* 'I have two daughters who are yet to have known/slept with men').[10] Old English *wer* is cognate with Latin *vir*, but *beorn* was likely used because it alliterates with *gebedscipe* (cf., *ge-* prefixes are unstressed in Old English verse).

(7) ne can þara *idesa* owðer gieta
not can the-GEN.PL woman-GEN.PL either yet
þurh gebedscipe ***beorna*** *neawest*
through intercourse men-GEN.PL proximity
'Neither of these women have slept with a **man** before'
[Genesis, 950-1050]

After circumscribing the variable context, each token was coded according to the available metadata: TEXT TYPE, TEXT ORIGIN, and TIME. The factor TEXT TYPE had two levels (prose, verse), TEXT ORIGIN had two levels (translated, not translated), and TIME had four levels for Old English (O1, O2, O3, O4) and three levels for Middle English (M1, M2, M3). While some metadata for DIALECT was available, DIALECT was not included as a factor for three reasons. First, not all texts contained such metadata. Second, there is some disagreement regarding the dialect in which specific manuscripts are written. Third, for Old English, there is a bias toward West Saxon texts, which substantially outweigh Northumbrian, Kentish, and Mercian texts. To test whether alliteration had a significant effect on lexical choices in verse, each variant in verse texts was coded binomially for the presence or absence of neighboring words with which the variant could alliterate. For the multivariate analyses, binary mixed effects logistic

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Table 1. Distribution of third-person male adult noun referents in Old English

Variants	*n*	%
wer	266	42.2
man	86	13.6
guma	85	13.5
secg	31	4.9
beorn	29	4.6
hæle[*þ*]	18	2.9
rinc	18	2.9
freca	9	1.4
wæpned	9	1.4
ceorl	8	1.3
wæpman	6	1.0
Other	66	10.4
Total	631	100

regressions were developed in *Rbrul* (Johnson, 2009), with TEXT ID run as a random intercept. In all models, the most frequent variant of the period (i.e., *wer* in Old English, and *man* in Middle English) was run against all other variants within that period, coded binomially.

Results

Old English distributional analysis

For Old English, a total of 631 tokens were included in the envelope of variation. Of the twenty-five attested variants (*beorn*, *carlman*, *cempa*, *ceorl*, *cniht*, *duguð*, *eorl*, *freca*, *guma*, *hæle*[*þ*], *hildedeor*, *hyse*, *leod*, *magu*, *man*, *rinc*, *scealc*, *secg*, *sundbuend*, *þegn*, *wæpman*, *wæpned*, *wer*, *wiga*, *wigmen*), *wer* was most frequent, which made up 42.2% of the semantic field. The variants *man* and *guma* competed for second place, each occupying approximately 13.5% of the system. The overall distribution of variants is reported in Table 1, with some examples of use in (8).

(8) a. *Iob wæs gehaten sum heah Godes þegen*
Job was called some high God-GEN servant
on þam lande Chus swiðe geleafull ***wer***
on the land Chus very faithful man
'Job was the name of a high servant of God in the land of Chus, who was a very faithful **man**'
[Ælfric's Letter to Sigeweard, 950-1050][11]

b. *on þis ilcan tyme forðferde Ælfsine abbot of Burh*
on this same time passed Alfsine abbot of Burh

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

7 man ceas þa Arnwi munec
and man chose then Earnwig monk
to abbod forþan þe he wæs swiðe god ***man*** *7 swiðe bilehwit*
to abbot because that he was very good man and very sincere
'During this time, Elfsinus of Peterborough died and they chose Anry, a monk, to be their abbot because he was a very good and very benevolent **man**'
[Chronicle MS E Early, 1050-1150][12]

c. *Hie þa æt burhgeate* ***beorn*** *gemitton sylfne sittan sunu*
they then at city-gate-DAT.SG man met.PL self sit son
'They then met the **man**, the son of Haran himself, sitting at the city-gates'
[Genesis, 950-1050]

To explore differences in use across prose and verse, the variants were cross-tabulated by TEXT TYPE (see Table 2). Different variants were favored by different text types. In prose texts, *wer* was the number one variant, at 63.6%, compared with 19.7% in verse texts. In contrast, *guma* was the number one variant in verse texts, at 24.8%, compared with 2.2% in prose. The type-token ratio (TTR), a common measure of lexical density, indicates that a wider range of variants was found in verse ($n = 23$ types, 314 tokens → TTR = .073) than in prose ($n = 12$ types, 317 tokens → TTR = .037). Of the 314 tokens found in verse, 76% ($n = 239$) alliterated with words in proximity. A chi-square test found that alliteration had a significant effect ($p < .001$) on the lexical choices within the semantic domain of variants found in Old English verse. The need

Table 2. Distribution of Old English variants by text type

Variants	Prose		Verse	
	n	%	*n*	%
wer	204	63.6	62	19.7
man	72	22.4	14	4.4
guma	7	2.2	78	25
secg	2	.6	29	9.2
beorn	0	0	29	9.2
ceorl	2	.6	6	1.9
freca	0	0	9	2.9
rinc	0	0	18	5.7
hæle[þ]	0	0	18	5.7
wæpned	8	2.5	1	.3
wæpman	6	1.9	0	0
Other	16	6.2	50	16
Total	317	100	314	100

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

for alliteration may explain the wider range of variants found in verse compared to prose.

As for the effect of provenance (Table 3), nontranslated texts contained a wider a range of variants (types = 25, tokens = 409 → TTR = .06) than translated texts (types = 13, tokens = 222 → TTR = .058), but the type-token ratio was almost identical. In translated texts, *wer* made up 68.3% of the semantic field. In contrast, although *wer* was also the most widely used variant in nontranslated texts, it made up a smaller share of the system (28.1%).

Old English multivariate analysis

To examine the statistical significance and relative strength of the factors operating on this semantic field, a binary mixed effects logistic regression was computed in *Rbrul* (Johnson, 2009). TIME, TEXT ORIGIN, and TEXT TYPE were run as fixed effects, with all possible interactions. The output is summarized in Table 4. Factor weights (FW) indicate the probability of the application value (i.e., *wer*) to occur in the listed context. Factor weights closer to 1 indicate a favoring effect whereas factor weights closer to zero indicate a disfavoring effect. Although TEXT was originally coded with four levels (O1, O2, O3, O4), due to the limited data available for O1 (2,190 words), for which only two tokens were included, O1 and O2 were collapsed into one level.

All three factors were found to significantly affect the probability of *wer* to occur in the Old English texts, with a significant interaction between TIME and TEXT ORIGIN. *Wer* occurred more frequently at the beginning of the Old English period (O2) than at the end (O4), in prose than verse, and in translated texts than non-translated texts.

Table 3. Distribution of Old English variants by origin

Variants	Non-translated		Translated	
	n	%	*n*	%
wer	115	28.1	151	68.3
man	50	12.2	36	15.8
guma	72	17.6	13	5.9
secg	30	7.3	1	.5
beorn	25	6.1	4	1.8
ceorl	7	1.7	1	.5
freca	9	2.2	0	0
rinc	15	3.7	3	1.4
hæle[*þ*]	17	4.1	0	0
wæpned	6	1.5	3	1.4
wæpman	5	1.2	1	.5
Other	58	14.3	9	3.9
Total	409	100	222	100

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Table 4. Logistic regression of the factors influencing the use of *wer* versus all other Old English variants

	n	%	FW
FIXED EFFECTS			
TIME**			
O3	418	32.3	.70
O2	140	69.3	.50
O4	73	46.6	.34
Range			*36*
TEXT TYPE***			
prose	317	64.4	.81
verse	314	19.7	.20
Range			*61*
TEXT ORIGIN**			
translated	222	68.0	.69
non-translated	409	28.1	.30
Range			*39*
RANDOM EFFECTS			
TEXT ID	*SD* = 2.12		
	Total *N* = 631, Input = .453, Texts = 72, *p < .05, **p < .01, ***p < .001		

The range for the factor groups, calculated by subtracting the lowest factor weight from the highest, indicates that, of the three factors, TEXT TYPE had the strongest effect on the absence or occurrence of *wer* in Old English. A random forest (Hothorn, Hornik, Strobl, & Zeileis, 2015) was run to confirm the hierarchical ordering of the constraints: TEXT TYPE ranked first, followed by TEXT ORIGIN, and then TIME. Although the higher frequency of *wer* in O4 (46.6%) than in O3 (32.3%) suggests the decrease in *wer* was not linear, when only prose texts are considered it is evident that *wer* continued to decrease throughout Old English. To confirm this, two follow-up models were run using data from only prose texts: *wer* occurred at a significantly greater frequency in O2 than O3 and significantly more frequently in O3 than O4, illustrating a significant downward trajectory over time.

Middle English distributional analysis

For Middle English, 246 tokens were included in the envelope of variation, with twenty-seven attested variants (*bachelor*, *baroun*, *beorn*, *burne*, *carlman*, *cherl*, *duʒeðe*, *erl*, *freke*, *gome*, *hathel*, *kempe*, *knape*, *kniʒt*, *ladde*, *lede*, *man*, *rahʒe*, *renk*, *schalk*, *segge*, *swein*, *þein*, *tulk*, *wepmann*, *wer*, *wyʒe*). The number one variant was *man*, with 57.3% (see Table 5). Examples of use are provided in (9). In contrast with Old English, the only attestation of *wer* referring to a male individual came

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Table 5. Distribution of third-person male adult noun referents in Middle English

Variants	*n*	%
man	141	57.3
kniȝt	33	13.4
gome	10	4.1
burne	8	3.3
shalk	6	2.4
beorn	5	2.0
freke	4	1.6
segge	4	1.6
cherl	3	1.2
hathel	3	1.2
wepmann	3	1.2
Other	26	10.7
Total	246	100

from the Early Middle English text *Ormulum* (9b). The low frequency of *wer* is consistent with evidence from the OED of its reported demise by Late Middle English (OED, *were*, n.1).[13] With the exception of (9b), when *wer* occurred in the Middle English data, it referred to a specific type of man, namely a married man, a use which was later usurped by the lexical item 'husband.'[14]

(9) a. *Þiss gode* ***mann*** *þiss gode prest þatt we nu mælenn offe*
this god man this good priest that we now speak of
wass alls I seȝȝde nu littlær ȝehatenn Zacaryas
was as I said now little-early called Zacharias
'This good **man**, this good priest, that we now talk of, was, as I said earlier, called Zacharias'
[Ormulum, 1150-1250]

b. *Uss birrþ heroffe witenn wel 7 seon 7 unnderrstanndenn, þatt David*
us behooves thereof know well and see and understand that David
kingess kinness men, off ***weress*** *oþþr off wifess wiþþ Aaroness kinness men*
king-GEN kins men from men or from women, with Aaron-GEN kins men
Off siþre wærenn sammnedd to streonenn streon to wurrþenn sibb,
from lately were gathered to acquire offspring to become relation
wiþþ kingess 7 wiþþ preostess
with kings and with priests
'It is necessary for us to know, see and understand, that the lineage of King David's kin, from **men** or from women, were gathered lately to have offspring in

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

order to be related to Kings and to priests'
[Ormulum, 1150-1250]

c. *Þa namen hi þa men þe hi wenden ðat ani god hefden, bathe*
then took they the men who they turned that any goods had both
be nihtes 7 be dæies, ***carlmen*** *7 wimen, 7 diden heom in prison*
by night and by day men and women and did them in prison
'Then they seized those people who had any goods, both during the night and during the day, both **men** and women, and threw them in prison'
[Peterborough Chronicle, 1150-1250]

As for the distribution by TEXT TYPE (see Table 6), *man* was the overwhelming choice in Middle English prose (92%) but occupied 41.1% of the field in verse. Of the 168 tokens of third-person male adult noun referents in Middle English verse, forty-eight (18%) alliterated, suggesting that while alliteration still influenced lexical choices in Middle English verse, it played less of a role than in Old English verse. The diminished role of alliteration in Middle English, however, is indicative of a larger change in verse style, as 34% of the Middle English variants that did not alliterate in verse, instead rhymed (e.g., *kniȝt - riȝt*, *man - þan*). In texts such as *Sir Orfeo* and the *Dancers of Colbek*, rhyme is the emphasis, not alliteration.

Middle English multivariate analysis

A binary mixed effects logistic regression was run on the Middle English data, using *man* (the most frequent variant) as the application value. The output is reported in Table 7. The model found TEXT TYPE to significantly influence the probability of *man* to occur over any other variant, appearing more frequently in prose than in verse. TEXT ORIGIN and TIME also significantly affected the occurrence of *man*. While, like in Old English, the model suggests that the increase in frequency of *man* was not consistent across

Table 6. Distribution of Middle English variants by text type

	Prose		Verse	
Variants	*n*	%	*n*	%
man	72	92	69	41.1
kniȝt	1	1.2	32	19.1
gome	0	0	10	6.0
burne	0	0	8	4.7
shalk	0	0	6	3.5
beorn	0	0	5	3
Other	6	7.7	38	22.6
Total	78	100	168	100

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Table 7. Logistic regression of the factors influencing the use of *man* versus all other Middle English variants

	n	%	FW
FIXED EFFECTS			
TIME**			
M3	77	55.8	.88
M1	126	53.2	.34
M2	43	72.1	.24
Range			*64*
TEXT TYPE***			
prose	78	92.3	.84
verse	168	41.1	.16
Range			*68*
TEXT ORIGIN**			
translated	37	81.1	.62
non-translated	209	53.1	.37
Range			*25*
RANDOM EFFECTS			
TEXT ID	*SD* = 2.7		*n* = 20
	Total *N* = 246, Input = .868 *p < .05, **p < .01, ***p < .001		

time, when only prose texts are considered it is evident that *man* continued to increase in frequency throughout Middle English. TEXT TYPE significantly interacted with TIME due to high frequency of genre-specific variants in verse (e.g., *knight*). These variants occur frequently, not necessarily because they were used frequently in everyday discourse, but due to the nature of verse content, wherein references to knights and chivalry are common. The range for the factor weights, along with a random forest indicates that TEXT TYPE had the strongest effect on the variation, followed by TIME and TEXT ORIGIN.

Changes from Old English to Middle English

To examine changes from Old English to Middle English, the frequency of *wer*, *guma*, and *man* was plotted across time (Figures 1 and 2). Frequency was measured by comparing the number of times a variant occurred versus the total number of referentially equivalent tokens by subperiod. Figure 1 shows the demise of *wer* from Old English to Middle English and its gradual replacement by *man*. The spike in frequency of *wer* in O4 (Old English: 1050-1150 CE) and the dip in frequency of *man* in M3 (Middle English: 1350-1420 CE) is a function of TEXT TYPE interference. When only prose texts are considered, the trend is clearer (see Figure 2). The change from *wer* to *man* follows a clearly identifiable s-curve pattern.

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

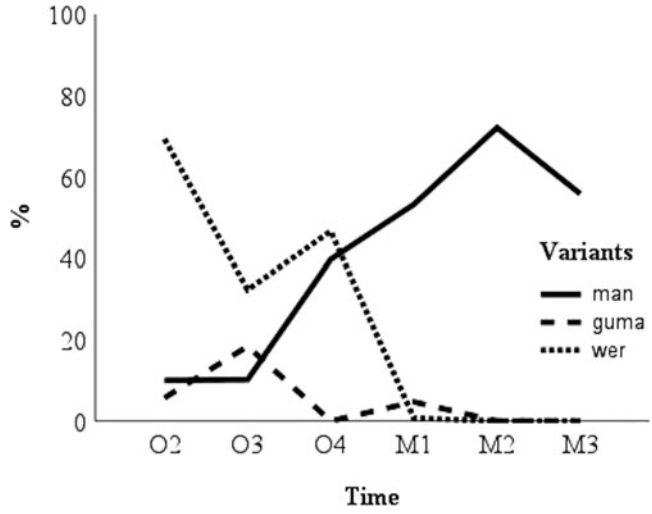

Figure 1. The frequency of *wer*, *guma*, and *man* from Old English to Middle English.

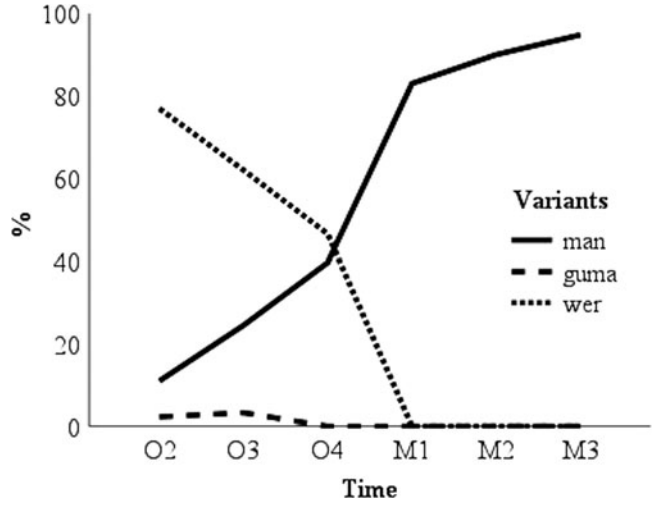

Figure 2. The frequency of *wer*, *guma*, and *man* in Old and Middle English prose texts.

Discussion

To examine changes within the set of third-person male adult noun referents in the early history of the English language, the present study examined the frequency of variants pertaining to this semantic field from Old English to Middle English. Distributional and multivariate analyses show a clear shift from *wer* in Old English to *man* in Middle English. Although the gender-inclusive use of *man* continued into Middle English, the number of instances in which *man* was used with reference to male-only individuals increased. The shift from *wer* to *man* follows an s-shaped

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

distribution, a pattern typically observed in other areas of linguistic change (Nevalainen, 2015; Tagliamonte & Smith, 2021). While language change does not have to follow an s-shaped pattern (Kauhanen, 2017; Newberry, Ahern, Clark, & Plotkin, 2017), an s-curve temporal trajectory is often viewed as a clear indication of lexical replacement (Blythe & Croft, 2012:278-279; Chambers, 2002:361). Assuming the data in the present study are representative, the low frequency of *wer* and its eventual demise in Middle English illustrates lexical change within this onomasiological set. While the actuation problem occludes the causation of this change (Weinreich et al., 1968:102), one might speculate that the Anglo-Norman loanword *werre* 'war' (MED, *werre*, n.), which shows up in twelfth century texts, had an influence on the demise of *wer*, as *wer* and *werre* could have been homophonous.[15] Although homophony of forms may be too simplistic of an explanation to account for the loss of *wer* "man," an explanation of this kind would be in line with the notion of a "homonymic clash" (Samuels, 1972:67-75), which has been proposed as a mechanism of change for several lexical items in Middle English. Since data show that *wer* was already decreasing throughout Old English, a homonymic clash could not have been the sole cause of this change, but this clash may have accelerated a change that was already underway.

Wer temporarily retreats to use as part of a related, but different, semantic field, namely "husband," which later too was replaced by the competing lexical item *husband*.[16] To the question *Where did* wer *go?*, the present study shows that *wer* was gradually replaced by the competing variant *man* from Old English to Middle English, with a clear relationship between the increase in frequency of *man* and the decrease in frequency of *wer*. The present-day compound *werewolf*, literally 'man-wolf,' is one of the few remaining breadcrumbs of this once frequently used noun and remains in the language today only as a vestige.[17]

One implication of this study is that lexical replacement is gradual, as the shift from *wer* to *man* appears to have taken place over approximately 400-500 years. However, there are two alternative explanations for the tardiness of this change. First, because frequently used lexical items are typically replaced less frequently (Pagel, Atkinson, & Meade, 2007), the high frequency of *wer* in Old English may account for why this replacement took centuries to be complete. The second factor to consider is the written transmission. Given that lexical choices are known to shift from generation to generation (Tagliamonte & Brooke, 2014; Tagliamonte & Jankowski, 2019; Tagliamonte & Pabst, 2020), and the locus of linguistic change is generally acknowledged to be in spoken as opposed to written language (Milroy, 1992:32), the Old and Middle English extant manuscripts may leave the impression that this replacement was gradual even though the change may have been accelerated in spoken language but remained in the language in formal written contexts, as is attested in the extant manuscripts.[18] After all, there are well documented register effects that condition and constrain language variation and change (Biber, 2012), which may have contributed to the longevity of this lexical replacement. The seemingly gradual nature of this change may therefore be a byproduct of the limited data that remain, that is, the notorious "bad data problem" (Labov, 1994:11), as it is inevitable that an analysis of this kind would be construed through a written lens.

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

As for the factors contributing to variation, distributional and multivariate analyses indicate that TEXT TYPE and TEXT ORIGIN significantly affected lexis. Variants such as *wer* and *man* were more frequent in prose than in verse, with Old English verse texts making use of a wider range of variants, likely due to alliterative requirements. Variants such as *shalk* (OE *scealc*) and *renk* (OE *rinc*) rarely occurred in Old English prose, suggesting that these variants were bound by stylistic tradition. Whether a text was translated from a Latin source also significantly influenced the lexical decisions in Old English, with *wer* occurring more frequently in translated texts. In contrast, in nontranslated texts a wider range of variants was employed (e.g., *freca*, *rink*), but this effect may be due to the skewed proportion of verse in nontranslated texts compared to translated ones. The higher frequency of *wer* in translated texts may also be attributed to the fact that Old English *wer* and Latin *vir* 'man' are cognates (Proto Germanic **uiraz*/*uiraR*), although counterexamples in translation choice were found, as in (7) above.

Conclusion

The semantic field of third-person male adult noun referents is a dynamic and heterogeneous one, with analyses of present-day varieties of English pointing to recent changes within this domain. The present study showed that variation within this onomasiological set is not new and has existed since the beginning of the history of the English language. *Wer* was once the most frequently used variant to refer to a male adult, but it was gradually replaced over time by *man*. This diachronic shift in lexical preference followed a prototypical s-shaped distribution, suggesting that, like other areas of linguistic change, lexis may follow similar patterns of change. While research using apparent time data (Chambers, 1995) or short periods of time (Grieve et al., 2017) point to the applicability of s-shaped trajectories for lexical change, the analysis of change within the semantic field of third-person male adult noun referents over approximately six hundred years adds an important diachronic dimension to this discussion.

Acknowledgments. This paper was the recipient of the Richard M. Hogg Prize, awarded by the International Society for the Linguistics of English (ISLE, 2021). The author would like to thank Karen Beaman, Laurel Brinton, Thijs Porck, and Mo Pareles for their comments on the paper, along with the three anonymous reviewers.

Notes

1. An example of the gender-inclusive use of *man* in Old English is the reference to Adam and Eve as *twegen men* 'two men': *on ðam sixtan dæge he gesceop eal deorcynn 7 ealle nytenu þe on feower fotum gað 7 þa twegen men Adam 7 Euan* 'on the sixth day, he created all wild animals, and all four-footed animals, and the two humans, Adam and Eve.' An example from Middle English can be found in the *Ormulum*: *Zacarie, Godess preost 7 ȝho þatt wass hiss macche wærenn rihhtwise 7 gode menn* 'Zachary, God's priest, and she who was his wife, were righteous and good people.' The gender-inclusive use of *man* still lives on in present-day substantives such as *manslaughter* (Old English *mansliht*). As for the gender-specific use, *man* can also appear in compounds such as *wæpenmann* 'weapon man,' which refers to male individuals, such that *wæpen* 'weapon' is also interpreted as meaning 'penis' (Holthausen, 1934:380). *Wæpnedmenn* are therefore male warriors; the gendered term reflecting the military gender makeup of Early Medieval England, but also individuals with male anatomy. Of the nine examples of *wæpnedmenn* in the dataset, seven occurred alongside inflectional forms of *wifman* 'woman,' illustrating, by contrast, the referential

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

gender of *wæpnedmenn* (e.g., *ge wæpmenn ge wimmenn* 'both men and women'). For information on the replacement of indefinite *-one* with *-body*, see D'Arcy, Haddican, Richards, Tagliamonte, and Taylor (2013).

2. Variation between *wer* and *guma* predates Old English. In the Gothic manuscript *Ambrosianus A*, *waír* is used as a translation of Greek ἀνήρ 'man,' but glosses show evidence that an additional scribe wrote *guma* in the margins (*du waira fullamma* → *du gumin fullamma* 'to the perfect man'), suggesting the two were synonymous at the time (see Falluomini, 2015:124).

3. Swain (2009:292) pointed out that in the Old English translation of the text *Heptateuch*, when referring to the number of Israelite males in the military, *wæpmanna* was used, which unambiguously refers to male individuals.

4. Abbreviations: OE = Old English, ON = Old Norse, OS = Old Saxon.

5. Time periods: O1 [-850], O2 [850-950], O3 [950-1050], O4 [1050-1150], M1 [1150-1250], M2 [1250-1350], M3 [1350-1420].

6. Given the size of the text (approx. 16,000 lines), only lines 9,229-12,400 from Layamon's Brut were included in the analysis. In contrast, an additional 657 words (not included in the Helsinki Corpus) from *Ayenbite of Inwyt* were added, and additional parts of the *Peterborough Chronicle* were added (starting from folio 84r).

7. The six men were presumably castrated because castration was a Medieval punishment for cis men.

8. The Old English pairing *ceorl 7 wif* suggests a marital or sexual relation, often parents of a common child, whereas *wer 7 wif* does not necessarily have this connotation. This is reflected in the verbal lexicon *ceorlian* 'to take a husband' (i.e., 'to marry') and *wifian* 'to take a wife'. Note, however, that *wer* was sometimes used to mean 'husband,' - the adjective *werleas* (literally *wer* 'man' plus the derivational suffix *-leas* 'less') meant 'unmarried' (see also *ceorleas* 'unmarried'). For instance, *ða þæt Latinus hiere wer geascade* 'when Collantinus her husband asked.' Such uses of *wer* were also not included in the envelope of variation.

9. In Early Germanic law, *wergild* (literally 'man-money') referred to a financial tariff that the perpetrator had to pay to the victim, or if dead, their family, when a crime was committed. This type of compensation was paid if the victim was killed, deliberately wounded, insulted, or dishonored.

10. Mo Pareles pointed out that the men are being offered the *idesa* as local substitutes for the (male) guests they are trying to rape.

11. Faithfulness note: in Bodleian Library MS. Bodl. 343 fol. 131v, *gehaten* is written as *í haten*.

12. This example illustrates the multifunctionality of *man* in Old English. The first occurrence of *man* is an indefinite pronoun whereas the second is a noun.

13. In the OED, n.1 notation refers to noun number 1, indicating that other nouns are attested under the same basic orthographic form. In the case of *wer*, n.1 refers to 'a male person; a man' in contrast to n.2 that refers to *weregild* (see note 9).

14. There are nineteen tokens of *wer* in the Middle English text *Hali Meidhad* that mean 'husband.' One could speculate that the retention of *wer* in reference to third-person males in the Middle English text *Ormulum* is due to dialectal influence, as the text appears to have been written in an East Midlands dialect spoken in the region that historically was part of the *Danelaw* (perhaps of importance as Old Norse had the cognate *verr* 'man').

15. According to the OED (*war*, n. 1), the word 'war' comes from the Germanic root **werz*/*wers* meaning 'discord/confusion.' Medieval Latin borrowed this word from Old High German *werra*; *w* became *gu*, leading to Spanish *guerra* 'war.' In Norman French, the *gu* became *w*, which was later borrowed into English through contact with Anglo-Norman. The earliest attestation given in the OED is suspicious as the period given is Late Old English, but the source, the Anglo-Saxon Chronicle from 1154 CE, suggests it was later, likely Early Middle English.

16. The semantic shift from 'man' to 'husband' appears to be a common trend for lexical items within the semantic field of 'male adult.' For instance, *gome* (OE *guma*) lives on today only as 'groom,' *cherl* (OE *ceorl*) had already developed the additional meaning of 'married man,' and *man* later developed the specific meaning of 'husband,' as is evident in somewhat archaic expressions such as *I now pronounce you man and wife* (Curzan, 2003:158). Similar trends can also be observed in Present Day English with variants such as *fella* and *man* (e.g., *how's your fella/man doing*?), as well as other Germanic languages (e.g., Icelandic *versæll* 'well married,' literally, 'husband-blessed').

17. The noun *world* is also a surviving remnant of *wer*, which historically was a compound consisting of *wer* 'man' and *old* 'age,' translating literally as 'age of man' (OE *weorld*, OS *uuerold*). Other etymologically

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

related words such as *werod* 'troop' (OS *uuerod* 'crowd') have died out. See, for instance, retentions in literary Icelandic, such as *ver-gjarn* 'nymphomaniac' (literally 'man-eager').

18. Note, however, that there are instances of language change that start in writing and diffuse to spoken language, as has been argued to be the case of *wh*-relative pronouns in English (e.g., Romaine, 1982:122) and *s*-genitive with inanimate possessors (e.g., Hinrichs & Szmrecsanyi, 2007:441; Jankowski, 2013:103-105).

References

Bailey, Charles-James. (1973). *Variation and linguistic theory*. Washington, DC: Center for Applied Linguistics.

Baldinger, Kurt. (1980). *Semantic theory*. Oxford: Blackwell.

Biber, Douglas. (2012). Register as a predictor of linguistic variation. *Corpus Linguistics and Linguistic Theory* 8:9–37.

Blythe, Richard, & Croft, William. (2012). S-curves and the mechanisms of propagation in language change. *Language* 88(2):269–304.

Bosworth-Toller (2014). In T. Northcote Toller, C. Sean, & O. Tichy (eds.), *An Anglo-Saxon Dictionary Online*. Faculty of Arts, Charles University. https://bosworthtoller.com/

Cameron, Angus, Amos, Ashley Crandell, & Healey, Antonette diPaolo. (2018). *Dictionary of Old English Project*. https://tapor.library.utoronto.ca/doe/

Chambers, Jack. (1995). The Canada-US border as a vanishing isogloss: The evidence of *chesterfield*. *Journal of English Linguistics* 23(1-2):155–166.

Chambers, Jack. (2002). Patterns of variation including change. In J. Chambers, P. Trudgill, & N. Schilling Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 349–372.

Cheshire, Jenny. (2007). Discourse variation, grammaticalisation and stuff like that. *Journal of Sociolinguistics* 11(2):155–193.

Curzan, Anne. (2003). *Gender shifts in the history of English*. Cambridge: Cambridge University Press.

D'Arcy, Alexandra, Haddican, Bill, Richards, Hazel, Tagliamonte, Sali, & Taylor, Ann. (2013). Asymmetrical trajectories: The past and present of *-body/-one*. *Language Variation and Change* 25 (3):287–310.

Elsweiler, Christine. (2011). *Laȝamon's Brut between Old English heroic poetry and Middle English romance: A study of the lexical fields 'hero', 'warrior' and 'knight'*. Frankfurt: Peter Lang.

Falluomini, Carla. (2015). *The Gothic version of the Gospels and Pauline Epistles*. Berlin: Walter de Gruyter.

Firth, John. (1957). *Papers in Linguistics*. London: Oxford University Press.

Franco, Karlien, & Tagliamonte, Sali A. (2021). Interesting fellow or tough old bird? 3rd person male referents in Ontario. *American Speech* 96(2):192–216.

Geeraerts, Dirk. (2010). *Theories of lexical semantics*. Oxford: Oxford University Press.

Grieve, Jack, Nini, Andrea, & Guo, Diansheng. (2017). Analyzing lexical emergence in Modern American English online. *English Language and Linguistics* 21(1):99–127.

Grygiel, Marcin. (2006). On the cyclicity of meaning alterations in English historical synonyms of man/male human being. In R. W. McConchie, O. Timofeeva, H. Tissari, & T. Säili (eds.), *Selected proceedings of the 2005 symposium on new approaches in English historical lexis* (HEL-LEX). Somerville, MA: Cascadilla Proceedings Project. 60–68.

Hinrichs, Lars, & Szmrecsanyi, Benedikt. (2007). Recent changes in the function and frequency of standard English genitive constructions: a multivariate analysis of tagged corpora. *English Language and Linguistics* 11(3):437–474.

The Historical Thesaurus of English. (2023). (2nd ed., version 5.0). University of Glasgow. https://ht.ac.uk/

Holthausen, Ferdinand. (1934). *Altenglisches etymologisches Wörterbuch*. Heidelberg: Carl Winters Universitätsbuchhandlung.

Hothorn, Torsten, Hornik, Kurt, Strobl, Carolin, & Zeileis, Achim. (2015). *Party: A laboratory for recursive partytioning*. http://cran.r-project.org/web/packages/party/index.html.

Jankowski, Bridget L. (2013). *A variationist approach to cross-register language variation and change*. Doctoral dissertation, University of Toronto.

Jankowski, Bridget, & Tagliamonte, Sali A. (2019). Supper or dinner?: Sociolinguistic variation in the meals of the day. *English World-Wide* 40(2):170–201.

https://doi.org/10.1017/S0954394523000108 Published online by Cambridge University Press

Johnson, Daniel E. (2009). Getting off the *GoldVarb* standard: Introducing *Rbrul* for mixed effects variable rule analysis. *Language and Linguistics Compass* 3:359–383.
Kauhanen, Henri. (2017). Neutral change. *Journal of Linguistics* 53(2):327–358.
Kleparski, Grzegorz. (2003). Churls, harlots and sires: The semantics of Middle English synonyms of man. *Studia Anglica Posnaniensia* 39:47–55.
Kleparski, Grzegorz. (2005). Towards the semantics of Middle English synonyms of MAN. *Studia Anglica Resoviensia* 3:88–95.
Kytö, Merja. (1996). *Manual to the diachronic part of the Helsinki Corpus of English Texts: Coding conventions and list of source texts*. (3rd ed.). Helsinki: University of Helsinki.
Labov, William. (1969). Contraction, deletion, and inherent variability of the English copula. *Language* 45 (4):715-762.
Labov, William. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
Labov, William. (1994). *Principles of linguistic change: Volume 1: Internal factors*. Cambridge: Blackwell.
Love, Robie, Dembry, Claire, Hardie, Andrew, Brezina, Vaclav, & McEnery, Tony. (2017). The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics* 22(3):319–344.
Lyell, Charles. (1830-1833). *Principles of geology, in three volumes*. London: John Murray, Albemarle-Street.
MED = *Middle English Dictionary*. (February, 2021). https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary
Milroy, James. (1992). *Language variation and change: On the historical sociolinguistics of English*. Oxford: Blackwell.
Nevalainen, Terttu. (2015). Descriptive adequacy of the s-curve model in diachronic studies of language change. In C. Sanchez-Stockhammer (ed.), *Can we predict linguistic change?* (Studies in Variation, Contacts and Change in English 16), University of Helsinki: VARIENG. https://varieng.helsinki.fi/series/volumes/16/nevalainen/index.html
Newberry, Mitchell, Ahern, Christopher, Clark, Robin & Plotkin, Joshua B. (2017). Detecting evolutionary forces in language change. *Nature* 551:223–226.
OED = Oxford English Dictionary. (February, 2021). https://www.oed.com/
Pagel, Mark, Atkinson, Quentin & Meade, Andrew. (2007). Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449:717–720.
Rauer, Christine. (2017). *Mann* and gender in Old English prose: a pilot study. *Neophilologus* 101(1):139–158.
Raumolin-Brunberg, Helena & Kahlas-Tarkka, Leena. (1997). *Indefinite pronouns with singular human reference*. In M. Rissanen., M. Kytö & K. Heikkonen (eds.), *Grammaticalization at work. Studies of long-term developments in English. Topics in English linguistics 24*. Berlin and New York: Mouton de Gruyter. 17–86.
Rissanen, Matti, Kytö, Merja, Kahlas-Tarkka, Leena, Kilpiö, Matti, Nevanlinna, Saara, Taavitsainen, Irma, Nevalainen, Terttu & Raumolin-Brunberg, Helena. (1991). *Helsinki Corpus of English Texts: Diachronic and Dialectal*. University of Helsinki. https://varieng.helsinki.fi/CoRD/corpora/HelsinkiCorpus/
Romaine, Suzanne. (1982). *Socio-historical linguistics: its status and methodology*. Cambridge: Cambridge University Press.
Samuels, Michael L. (1972). *Linguistic evolution: With special reference to English*. Cambridge: Cambridge University Press.
Sankoff, David & Thibault, Pierrette. (1981). Weak complementarity: Tense and aspect in Montreal French. In B. B. Johns & D. R. Strong (eds.), *Syntactic change*. Ann Arbor, MI: University of Michigan Press. 205–216.
Sankoff, David, Thibault, Pierrette & Bérubé, Hélène. (1978). Semantic field variability. In D. Sankoff (ed.), *Linguistic variation: Models and methods*. New York, NY: Academic Press. 23–43.
Sievers, Eduard. (1893). *Altgermanische Metrik*. Halle: Niemeyer.
Sisam, Kenneth. (Ed.). (1928). *Fourteenth century verse & prose*. Oxford: Clarendon Press.
Stenroos, Merja. (2002). Words for *MAN* in the transmission of Piers Plowman. In J. E. Díaz Vera (ed.), *A changing world of words: Studies in English historical lexicography, lexicology and semantics*. Amsterdam: Editions Rodopi. 375–409.
Stratton, James M. (2020). Adjective intensifiers in German. *Journal of Germanic Linguistics* 32(2):183–215.
Stratton, James M. (2022a). Old English intensifiers. The beginnings of the English intensifier system. *Journal of Historical Linguistics* 12(1):31–69.

Stratton, James M. (2022b). Tapping into German adjective variation: A variationist sociolinguistic approach. *Journal of Germanic Linguistics* 34(1):63–102.
Stratton, James M. & Sundquist, John D. (2022). A variationist sociolinguistic analysis of intensifiers in Oslo Norwegian. *Journal of Germanic Linguistics* 34(4):385–419.
Swain, Lary. (2009). *Ælfric of Eynsham's Letter to Sigeweard: An edition, translation and commentary*. Doctoral dissertation, University of Illinois at Chicago.
Tagliamonte, Sali A. (2008). *So different* and *pretty cool!* Recycling intensifiers in Toronto, Canada. *English Language and Linguistics* 12(2):361–394.
Tagliamonte, Sali A. (2022). Dialects as a mirror of historical trajectories: Canadian English across Ontario (North America). In M. Kytö & L. Siebers (eds.), *Early North American Englishes*. Amsterdam: John Benjamins. 231–258.
Tagliamonte, Sali A. & Brooke, Julian. (2014). A (weird) language tale: Variation and change in the adjectives of strangeness. *American Speech* 89(1):4–41.
Tagliamonte, Sali A. & D'Arcy, Alexandra. (2004). *He's like, she's like*: The quotative system in Canadian youth. *Journal of Sociolinguistics* 8(4):493–514.
Tagliamonte, Sali A. & Jankowski, Bridget L. (2019). *Golly, gosh, and oh my God!* What North American Dialects can tell us about swear words. *American Speech* 94(2):195–222.
Tagliamonte, Sali A. & Pabst, Katherina. (2020). A *cool* comparison: Adjectives of positive evaluation in Toronto, Canada and York, England. *Journal of English Linguistics* 48(1):3–30.
Tagliamonte, Sali A. & Smith, Jennifer. (2021). Obviously undergoing change: Adverbs of evidentiality across time and space. *Language Variation and Change* 33(1):81–105.
Terkourafi, Marina. (2011). The pragmatic variable: Toward a procedural interpretation. *Language in Society* 40(3):343–372.
A Thesaurus of Old English. (2017). Glasgow: University of Glasgow. http://oldenglishthesaurus.arts.gla.ac.uk/
Trier, Jost. (1931). *Der deutsche Wortschatz im Sinnbezirk des Verstandes. Die Geschichte eines sprachlichen Feldes I. Von den Anfängen bis zum Beginn des 13. Jhdts*. Heidelberg: Winter.
Weinreich, Uriel, Labov, William & Herzog, Martin. (1968). *Empirical foundations for a theory of language change*. Austin: University of Texas Press.

Cite this article: Stratton JM (2023). Where did *wer* go? Lexical variation and change in third-person male adult noun referents in Old and Middle English. *Language Variation and Change* **35**, 199–221. https://doi.org/10.1017/S0954394523000108