
FUNCTIONAL DIFFERENTIATION AND GRAMMATICAL COMPETITION IN THE ENGLISH JESPERSEN CYCLE

PHILLIP WALLAGE

Northumbria University

ABSTRACT Wallage (2008) argues for a model of the Middle English Jespersen Cycle in which each of its diachronic stages are functionally equivalent competitors in the sense proposed by Kroch (1989). However, recent work on the Jespersen Cycle in various Romance languages by Schwenter (2006), Hansen (2009) and Hansen & Visconti (2009) has argued that the forms in competition during the Jespersen Cycle are not simply diachronic stages, but perform different pragmatic or discourse functions. Hansen (2009) and Hansen & Visconti (2009) suggest that functional change may therefore underpin the Jespersen Cycle in these languages. Hence this paper explores the interface between pragmatic or functional change, and change in the syntax of sentential negation.

Analysis of data from the PPPCMEz (Kroch & Taylor 2000) show that *ne* (stage one) and *ne...not* (stage two) are similarly functionally differentiated during the ME Jespersen Cycle: *ne...not* is favoured in propositions that are discourse-old (given, or recoverable from the preceding discourse), whereas *ne* is favoured in propositions that are discourse-new. Frequency data appear to show the loss of these constraints over time. However, I argue that these frequency data are not conclusive evidence for a shift in the functions of *ne* or *ne...not*. Indeed, the results of a regression analysis indicate that these discourse constraints remain constant throughout Middle English, in spite of the overall spread of *ne...not* as the Jespersen Cycle progresses. Therefore, I conclude the spread of *ne...not* is independent of these particular discourse constraints on its use, rather than the result of changes in, or loss of, these constraints.

1 INTRODUCTION

This paper examines some of the constraints on variation and change in sentential negation strategies in early English from the Middle English period (1150-1500AD). The changes we see in this period of English exemplify the Jespersen Cycle (Jespersen 1917), and are well known from a number of languages. Examples (1) to (3) summarise the stages of the cycle that we find in Middle English.

- (1) STAGE 1:
we ne migen þat don
 We NEG can that do
 ‘We cannot do that’ (TRINIT,103.1369)
- (2) STAGE 2:
I ne may nat denye it
 I NEG may not deny it
 ‘I may not denye it’ (BOETH,435.C1.262)
- (3) STAGE 3:
Thou sall nought do so
 You shall not do so
 ‘You shall not do so’ (ROLLTR,43.880)

In common with Wallage (2008), a central focus of this paper will be to examine whether we can analyse developments within the Middle English Jespersen Cycle in terms of Kroch’s (1989) model of competition between mutually exclusive morphosyntactic options. However, while Wallage’s (2008) primary concern was a formal syntactic account of the transition from stage one (*ne*) to stage two (*ne...not*), this paper will focus more on the relationship between Wallage’s (2008) syntactic account of this change and the pragmatic functions of *ne* at stage one and *ne...not* at stage two. Hence this paper will examine the extent to which these two forms are pragmatically differentiated during the period of greatest variation and competition between them (the 12th-14th centuries). It will investigate the contribution that change in the pragmatic functions of these forms makes to the replacement of *ne* by *ne...not* during this period. The role of pragmatic factors at later stages of the Jespersen Cycle, in the transition from stage two *ne...not* to stage three *not*, requires detailed research and discussion of its own; hence, pending its further examination, the distribution of stage three *not* will not be discussed in detail in this paper.

Theories of morphosyntactic change such as Kroch (1989) assume competition between functionally equivalent syntactic options. Wallage (2008) presents an account of the Middle English Jespersen Cycle in these terms, where each formal stage (1)-(3) performs the same function. However, this assumption potentially obscures an interplay of form and function within the cycle. Recent work on other languages, principally French and Italian (Hansen 2009, Hansen & Visconti 2009), suggests that succeeding stages of the Jespersen Cycle tend to be employed in different functions. This suggests that processes of functional change (generalisation and specialisation in functions) may be involved in the Jespersen Cycle, and indeed may play a pivotal role in motivating the transition between stages.

Therefore, this paper has two objectives. First, I will establish whether stages one and two of the Middle English Jespersen Cycle are functionally differentiated, if not in absolute terms, in relative terms of particular forms being more frequently used in certain functions. Then I will examine how variation in functions and variation in forms interact in the replacement of *ne* by *ne...not*. Through this analysis we will see how particular types of functional change may be implicated in the spread of *ne...not* at stage two of Jespersen Cycle. The structure of the paper is as follows. In section 2, I outline some of the functions negation may have within a discourse, and observe how different negative forms associate with these different functions crosslinguistically. In section 3, I argue that stage one (*ne*) and stage two forms (*ne...not*) tend to be employed for different discourse functions in early Middle English. Section 4 examines whether changes in the functions associated with *ne...not* accompany the spread of this form – that is, whether the preference to associate *ne* and *ne...not* with certain functions changes over time as the overall frequencies of *ne* and *ne...not* increase during Middle English. In conclusion, section 5 will outline how we can accommodate discourse-functional constraints within a model of the Jespersen Cycle as morphosyntactic competition between forms at successive stages.

2 DISCOURSE FUNCTIONS OF NEGATION

Much recent work on negation, for example Horn (2001), Schwegler (1988), Israel (2001), Detges & Waltereit (2002), Schwenter (2006), Hansen (2009) and Hansen & Visconti (2009), has argued that negation performs a number of pragmatic or discourse functions. Schwegler (1988), Israel (2001) and Detges & Waltereit (2002) make a distinction between emphatic and neutral negation. For Israel (2001), emphatic negation is highly informative relative to normal default sentential negation. Thus emphatic negation is a hyponym of neutral negation, and performs a subset of the functions available to neutral negation. (Detges & Waltereit 2002: 183) claim that emphatic negation of a proposition *p* is ‘used whenever speakers want to act against some strong counter-expectation on the part of their listeners’, in other words, in contexts where the negation acts against the listener’s assumption that the affirmative proposition *p* is true.

Thus, the function of emphatic negation makes crucial reference to the preceding discourse, presuppositions or inferences arising from it, the listener’s prior beliefs relating to or arising from it, and the extralinguistic context in which it takes place. However, Schwenter (2006) observes that this distinction between emphatic and neutral negation does not fully explain the distribution of so-called ‘emphatic’ negative markers. While they can work against the assumption that a proposition is true, so-called ‘emphatic’ negative markers are not restricted to these contexts. Schwenter observes that in Italian and Catalan, for example, a formal distinction is made between single negative markers (*non* and *no*) and bipartite negative markers (*non...mica* and *no...pas*) in terms of the information status of the negated proposition

– that is, whether the proposition that is negated is already present as part of the discourse or not. This distinction makes no reference to the beliefs or expectations of the hearer or reader. As Dryer (1996) observes, it is possible for a proposition to be discourse-old without being part of the set of beliefs held to be true by either conversational participant. A discourse-old proposition is not necessarily one assumed or presupposed by the listener.

Discourse-old propositions are those which are already present to the attention to of the reader or hearer of the discourse, and in which the underlying proposition is recoverable from the preceding discourse, either because it has already been stated explicitly, as in (4), or because it arises through an implicature of some kind (scalar implicature, presupposition, conversational implicature), as in (5).

- (4) He said he went. *In fact he did not go*
- (5) a. I'm not hungry
b. I don't want anything to eat

Schwenter argues that Catalan *no...pas* or Italian *non...mica* tend to negate propositions that are discourse-old, that is, already part the discourse context and present to the attention of the participants within the discourse. Thus, in Catalan 'there must be a salient proposition, evoked either linguistically ... or situationally for *pas* to be felicitous' (Schwenter 2006: 333). Is a similar formal distinction made between ME negative markers in accordance with differences in their function or status in the discourse?

Hansen (2009) and Hansen & Visconti (2009) note a similar functional distinction between medieval French *ne* (stage one) and *ne...pas* (stage two). However, their work raises further questions about pragmatic change. They argue that this functional distinction weakens over time, as *ne...pas* becomes functionally unmarked through functional extension into discourse-new propositions. Their findings might lead us to wonder whether ME stage two *ne...not* also becomes pragmatically unmarked or neutral over time, and whether this change might explain the increasing overall frequency of *ne...not* in ME. Using diachronic data from ME, we can examine first whether stage one *ne* and stage two *ne...not* mark different discourse functions, and second, whether any such distinction is lost as *ne...not* becomes more frequent.

3 DISCOURSE FUNCTIONS OF NEGATIVE MARKERS IN MIDDLE ENGLISH

3.1 Methodological issues and procedures

We find the first three stages of a Jespersen Cycle in Middle English (1150-1500): the verbal proclitic *ne* at stage one, stage two in which the verbal proclitic comes to be supplemented by the postverbal negative *not*, and stage three in which *ne* is lost and *not* becomes the sole negative marker in the clause. This section will focus on the

first two stages and examine whether *ne* (stage one) and *ne...not* (stage two) are associated with any of the discourse functions listed here as (6).

- (6)
 - a. Denial of antecedent proposition: the negative proposition denies an earlier proposition which was explicitly stated in the discourse
 - b. Cancellation of an inference: the negative proposition cancels an implicature arising out of the preceding discourse
 - c. Repetition of an antecedent proposition: the negative proposition repeats an earlier proposition which was explicitly stated in the discourse
 - d. Statement of an inference: the negative proposition explicitly states a proposition which is implied by the preceding discourse
 - e. Negation of a proposition that is new to the discourse: the proposition is not identified by an antecedent proposition in the earlier discourse and not inferentially linked to the preceding discourse.

All of the negative forms *ne* and *ne...not* and *not* appear in each of these discourse functions in Middle English, as examples (7) to (11) from the PPCME2 corpus (Kroch & Taylor 2000) illustrate. The (a) examples involve stage one *ne*, the (b) examples stage two *ne...not* and the (c) examples stage three *not* in each discourse function.

- (7) Negation of earlier proposition:

- a. *þenne þe prest þe menezed rihtliche teðien. þenne
when the priest them admonishes rightfully tithen then
cumed þe werse to sume mannes heorte and minzed hine
comes the worst to some men's hearts and advises him
þat he swo **ne** do
that he so NEG do*

'When the priest admonishes them to give their tithes aright, then comes the devil to a man's heart and advises him that he do not so.'
(CMTRINIT,215.3014)

- b. *Alle ðo men ðe swinkeð on ðessere swinkfulle world, alle
All the men that labour in this toilsome world, all
he swinkeð for sumere hope ðe hie habbeð, ðe hem
they labour for some hope that they have, that them
oft eaten ande beswinkð ... Ac ðo ðe swinkeð for
often at end deceives ... But those that labour for
ðessere eadi hope, hie **ne** bieð **naht** becaht
this blessed hope, they NEG are not deceived*

‘All the men who labour in this toilsome world, they all labour for some hope they have which often deceives them in the end...But those who labour for this blessed hope, they are not deceived.’

(CMVICES_{1,33.385})

- c. *For it peyneth hem evere as though they sholde dye anon*
 For it pains them ever as though they should die now
but certes, they shal not dye
 but truly they shall not die

‘For it pains them always as if they should die now, but truly they shall not die’ (CMCTPARS_{292.C2.194})

(8) Repetition of an earlier negative proposition¹:

- a. *Ariseð þanne ge hauen seten, ac we ne mugeþ þat*
 Arise when you have sat but we NEG may that
done wiðuten his elpe. Seie we þanne to him
 do without his help. Say we then to him
Domine tu cognouisti sessionem meam et resurrectionem meam –
Domine tu cognouisti sessionem meam et resurrectionem meam –
louerd þu wost hu ich habbe seten. and þat ich ne
 lord you know how I have sat and that I NEG
mai wið-uten þin elp risen.
 may without your help rise.

‘Arise when you have sat, but we are not able to do that without His help. Let us say then unto him – Domine tu cognouisti sessionem meam et resurrectionem meam – Lord, you know how I have sat and that I am unable without your help to rise.’ (CMTRINIT_{103.1372})

- b. *zef þu ne cnawest þe seolf ... zef þu ne cnawest naut*
 If you NEG know yourself ... If you NEG know NEG
þe seolf
 yourself

‘If you not know yourself ... If you not know not yourself.
 (CMANCRIW_{II.80.941 & 948})’

- c. *And secoundely, he that is irous and wroth, he ne*
 And secondly he that is angry and full-of-wrath he NEG
may nat wel deme and he that may nat wel deme may
 may not well judge and he that may not well judge may
nat wel conseille
 not well advise

¹ As a reviewer notes, this function may prove problematic, since the form of the second negative proposition may be influenced by, or parallel, the form of the antecedent negative proposition.

‘And secondly, he that is angry and full of wrath, he may not judge well, and he who may not judge well may not give good advice.’
(CMCTMELI,222.C1.195)

(9) Statement of a negative inference:

- a. *Drihten seið ec on his godspelle þet þa beoð godes*
Lord says also in his gospel that they are God’s
bern þe beoð isibsumme ac sake ne sturiað.
children that are peaceable but strife NEG raise-up
‘The Lord says also in his gospel that they are God’s children who are peaceable and raise not up strife.’ (CMLAMBX1,113.1092)
- b. *Ich nam noht giet sad of mine sines and forþi*
I not-am not yet sated of my sins and therefore
ne mai ich hie noht forlete.
NEG can I them not renounce
‘I am not yet sated of my sins and therefore I cannot renounce them’
(CMTRINIT,75.1028)
- c. *þei dreven Brut out of þe lande & wolde not suffre*
they drove Brut out of the land and would not allow
hym among hem.
him among them
‘they drove Brut out of the land and would not allow him among them’ (CMBRUT3,6.131)

(10) Cancellation of an inference:²

- a. *And þah þes patriaches alse abel and noe and*
And though the patriarchs as Abel and Noah and
abraham and ysaac gode men weren þurh þet ho
Abraham and Isaac good men were through that they
weren itende of þan halie gast and al þos godnesse
were enlightened of the holy ghost yet all this goodness
hom ne mihte werien þet ho ne wenden
them NEG could prevent that they redundant NEG go
alle in to helle.
all in to hell.

² In Middle English, *ne* is sometimes used redundantly, without contributing negative semantics, to the complement clause of verbs like *werien* ‘prevent’. See [van der Wurff \(1999\)](#) and [Wallage \(2008\)](#) for discussion.

‘And though the patriarchs as Abel and Noah, Abraham and Isaac, were good men, being enlightened of the Holy Ghost, yet all this goodness could not preserve them from going into hell..’

(CMLAMB1,81.153)

- b. *and þe lage hadde þo alle þe mihtes þe haueð nu*
and the law had then all the virtues that has now
fulluht for ðat clensede þe man of sinne: swa doð nu
baptism for that cleansed the man of sin: as does now
fulluht ac it ne openede hem noht þe blisse of
baptism but it NEG opened them not the bliss of
heuene also fulcneng doð us.
heaven as baptism does us.

‘And that rite had then all the virtues which baptism now has, for that cleansed man of sin even as baptism now does, but it opened not to them the bliss of heaven as baptism does to us.’ (CMTRINIT,87.1165)

- c. *Als es o þaim at saie als þe gold þu may se, þat*
Also is of them that say as the gold you may see, that
semis gode and es noht:
seems good and is not:

‘Also is of them that say as the gold you may see, that seems good and is not’ (CMBENRUL,4.105)

(11) Discourse-new proposition:

- a. *and here wuneð on wanrede and þoleð his unwilled,*
and here dwells in distress and suffers his discomfort,
hwile druie and hwile wete hwile chele
sometimes dry and sometimes wet sometimes cold
wile hete hwile hunger wile
sometimes hot sometimes hunger sometimes
þurst...hwile unhele hwile sorinesse and wile
thirst...sometimes sickness sometimes soreness and sometimes
werinesse and hwile wurmene cheu and fele oðre
weariness and sometimes worm’s biting and many others
þe ich telle ne mai and ne mai wiðuten helpe
that I tell NEG can and NEG may without help
him seluen þerwið werien
himself therewith protect

‘and here he dwells in distress and endures discomfort, sometimes dry and sometimes wet, sometimes cold and sometimes hot, sometimes hunger, sometimes thirst, sometimes sickness sometimes soreness,

sometimes weariness and sometimes the biting of worms, and many others that I cannot tell of, and may not protect himself against them without help' (CMTRINIT,123.1648)

- b. *ðis sade sanctus Iacobus, ðe hali apostel: 'Swiðe michel*
 This said Saint James the holy apostle: 'So much
helpð þas rihtwismannes bede', and ðar of he seið
 avails the righteous's prayer', and there of he told
ðelliche forbisne: 'Hit was on mann, dedlich alswa we bieð,
 such example 'It was a man mortal as we are,
*and he besohte at gode þat **naht ne** scolde reinin, for*
 and he besought of God that not NEG should rain for
ðe folc to kastin. Godd him ihierde – þat was Helyas
 the people to chasten. God him heard – that was Elias
– and wiðheld alle reines þrie hier and six moneþes...'
 – and withheld all rains three years and six months...'
 'This said St. James, the holy apostle: 'The prayer of a righteous man
 avails much', and he gives such an example of it: 'There was a man,
 mortal as we are, and he besought God that it should not rain, to
 chasten the people. God heard him – that was Elias – and withheld all
 the rains three years and six months...' (CMVICES1, 143,1787)
- c. *And moorover, whan oure Lord hadde creat Adam, oure*
 And moreover, when our Lord had created Adam, our
*forme fader, he seyde in this wise: It is **nat** good to*
 first father he said in this manner: It is not good to
been a man alloone ...
 be a man alone ...
 'And moreover, when our Lord had created Adam, our first father, he
 said this: 'It is not good to be a man alone...' (CMCTMELI,221.C2.165)

Since *ne*, *ne...not* and *not* can all appear in all five discourse functions, any correlation between particular forms and particular functions will only manifest itself under a quantitative analysis, such that the frequency of *ne* or *ne...not* is higher or lower in a particular context than in other contexts. To facilitate this quantitative analysis a large number of clauses (n=1922) are taken from the PPCME2 and each one categorised by discourse function.

The PPCME2 is a 1.2 million word corpus of English from the period 1150-1500. It is tagged for grammatical parts of speech and parsed at the clause level to facilitate the search for syntactic structures, and organised into four subperiods (1150-1250, 1250-1350, 1350-1420 and 1420-1500) to facilitate investigation of morphosyntactic change. The two earlier periods are longer than the later ones. This reflects differences in the amount of textual data available in each period. There are fewer texts

in earlier periods, hence the periods are longer in order to include a larger sample of different texts in these periods. Given the extant medieval English texts, the corpus comprises mostly religious texts, histories and literary texts, with a predominance of religious texts in the earlier centuries. Spoken registers are not represented in the corpus. The corpus includes texts translated from Latin or French originals, as well as texts originally written in English.

All the negative clauses from the two earlier subperiods 1150-1250 (n=674) and 1250-1350 (n=724) are included in the initial frequency counts for *ne*, *ne...not* and *not* in each discourse function. However, only 25% of the data from the period 1350-1420 are examined, due to the large number of negative clauses in this subperiod of the corpus (n=2254). The sampling procedure takes the first 25% of main clauses and subordinate clauses³ from each text to provide a sample which represents the source PPCME2 corpus as closely as possible in terms of the balance of texts and clause types. This results in a sample size of 564 clauses. The period 1420-1500 is excluded from the analysis because there is insufficient variability in the form of negation at this period. As the ME changes draw to completion, over 98% (n=1843/1878) of negative clauses in this period are negated by stage three *not*, making both stage one *ne* and stage two *ne...not* too infrequent to draw reliable conclusions about their distribution from the frequency counts for this period.

Another issue for the analysis is how to identify the discourse function of each proposition reliably. Identity relations between propositions in the discourse are easily identified in the texts. However, identifying implicatures is less straightforward. As texts cannot be read in the social and cultural contexts in which they were written, the socio-cultural common ground between writer and reader that informs interpretation of the discourse is missing. Therefore we can only examine the relationships between propositions within the texts themselves. These relationships are of five main kinds: identity, entailment, presupposition, scalar implicature and conversational implicature. Present-day English translations of the corpus texts allow us to identify the implicatures of propositions within the translated text. Where the translations are somewhat more free than literal, a word for word gloss of the original text is used. Propositions whose discourse function remained unclear or potentially ambiguous after reference to gloss and translation are few in number. These are simply excluded from the analysis.

3.2 Discourse functions of Middle English negatives

The examples given earlier as (7)-(11) show that *ne* and *ne...not* occur in all five discourse functions – in this sense *ne* and *ne...not* are functionally equivalent. However, Table 1 reveals that discourse function does constrain the use of each form in terms of its frequency of occurrence within each discourse function or context.

³ As Wallage (2008) provides data to show that the distribution of *ne*, *ne...not* and *not* are different in main clauses and subordinate clauses.

It reports the frequency of stage one *ne*, stage two *ne...not*, and stage three *not* in each of the five discourse functions given in examples (7) to (11), subdividing these functions into two types: those that introduce information new to the discourse (discourse-new) and those in which the content of the proposition is recoverable from the preceding discourse (discourse-old). Counterfactual and irrealis clauses, like the one in (12), are counted separately because Hansen (2009: 244) observes that such clauses favour stage one negation in medieval French more than other clauses. This observation is borne out for English by the ME data in Table 1.

- (12) *For soothly oure sweete Lord Jhesu Crist hath spared us ...,*
For truly our sweet Lord Jesus Christ has spared us ...,
that if he ne hadde pitee of mannes soule, a sory song
that if he NEG had pity of mans soul, a sorry song
we myghten alle synge.
we might all sing

‘For truly, our sweet Lord Jesus Christ has spared us ... that if he didn’t have
pity on man’s soul, a sorry song we might all sing’
(CMCTPARS,296.C2a.344)

	1150-1250			1250-1350			1350-1420			
	ne	ne...not	Total	ne	ne...not	Total	ne	ne...not	Total	
Denial of antecedent p	27.2% (n=12)	72.8% (n=32)	44	10.7% (n=3)	85.7% (n=24)	28	— (n=0)	2.2 (n=1)	97.8% (n=44)	45
Repetition of antecedent p	37.5% (n=3)	62.5% (n=5)	8	7.1% (n=1)	92.9% (n=13)	14	— (n=0)	15.4% (n=2)	84.6% (n=13)	15
Cancellation of inference	12.8% (n=5)	84.6% (n=33)	39	— (n=0)	89.7% (n=35)	39	— (n=0)	10.8% (n=4)	89.2% (n=33)	37
Assertion of inference	17.1% (n=28)	82.3% (n=135)	164	1.9% (n=5)	88.7% (n=228)	257	— (n=0)	15.5% (n=22)	84.5% (n=120)	142
TOTAL discourse-old	14.9% (n=38)	80.4% (n=205)	245	4.1% (n=9)	87.5% (n=300)	338	— (n=0)	15.4% (n=29)	84.6% (n=215)	244
Discourse-new	85.2% (n=335)	14.5% (n=57)	393	39.0% (n=135)	50.9% (n=176)	346	0.7% (n=2)	9.2% (n=27)	90.1% (n=265)	294
Counterfactual	92.3% (n=24)	7.7% (n=2)	26	68.8% (n=11)	31.2% (n=5)	16	33.3% (n=4)	25.0% (n=3)	41.7% (n=5)	12
TOTAL	397	264	664	155	481	700	6	59	480	545

Table 1 The distribution of clauses with and without *not* in each discourse function

The data in Table 1 show that the discourse status of the proposition has an effect on the distribution of stage one *ne* and stage two *ne...not*, but seems to have little effect on the distribution of stage three *not*. The most obvious distinction between contexts for *ne* and those for *ne...not* is between clauses that are discourse-new and clauses that are discourse-old. This is the case in both the periods 1150-1250 and 1250-1350, although the differences between the frequencies of *ne...not* in discourse-old and discourse-new contexts lessen over time as *ne...not* becomes more frequent. In both periods the distinction between discourse-old and discourse-new contexts has a statistically significant effect on the distribution of *ne...not* versus *ne*, where a p-value of <.05 indicates statistical significance.⁴ By the time we reach the period 1350-1420, stage one *ne* is very marginal and largely restricted to two contexts, discourse-new propositions and counterfactuals. In this period *ne* is too infrequent (n=6) to perform chi-square tests on its distribution across the different pragmatic contexts.

By contrast, although the frequencies of stage three *not* are low in Table 1, they are approximately equal across discourse-new and discourse-old propositions. Chi-square tests report no statistically significant differences in the distribution of stage three *not* versus other negative forms in discourse-new and discourse-old environments.⁵ These figures suggest that stage three *not* is not favoured in any of these five discourse functions at any period of ME, while stage two *ne...not* is favoured in discourse-old propositions. It follows that the transition from stage two *ne...not* to stage three *not* does not simply involve loss of the morpheme *ne*, but also a change in the discourse-functional constraints on *not* when it begins to appear independently of *ne*. The data in Table 1 show that *ne...not* and *not* are distinct in their pragmatic functions. Taking these data at face value, they appear to show that stage three *not* is not pragmatically marked for any of the five functions examined in Table 1. If this is right, more detailed research is required to provide a more precise characterisation of the role of pragmatic change or loss of pragmatic marking in the transition from stage two to stage three of the Jespersen Cycle. However, the remainder of this paper will focus particularly on the role of pragmatic factors in the transition from stage one *ne* to stage two *ne...not*.

In terms of diachronic developments, the changing frequencies of *ne* and *ne...not* in Table 1 indicate an extension in the functions of *ne...not* from marking discourse-new propositions to marking both discourse-new and discourse-old propositions, whereas *ne* seems to undergo a narrowing of function such that it goes

4 For the period 1150-1250, the results of a chi-square test on the distribution of *ne* versus *not* in the two contexts are chi-square (1df) = 281.8, p<.01. For the period 1250-1350, a similar chi-square test reports chi-square (1df) = 200.7, p<.01.

5 For the period 1250-1350 the figures for stage three *not* are: discourse-old n=29/343, discourse-new n=35/346; chi-square (1df) = 0.564, p = 0.45. For the period 1350-1420 the figures for stage three *not* are: discourse-old n=29/244, discourse-new = 27/294; chi-square (1df) = 0.989, p = 0.31. I assume statistical significance with a p-value of <.05, hence the differences in the distribution of stage three *not* across discourse-old and discourse-new contexts are not significant in either period.

from marking negation mostly in discourse-new propositions in the 12-13th centuries to marking negation most frequently in counterfactuals in the 14th century. Tables 2 and 3 show more clearly how the discourse-functions associated with stage one *ne* and stage two *ne...not* change over time. They subcategorise all occurrences of *ne* and *ne...not* according to discourse function.

Function of stage one <i>ne</i>	1150-1250	1250-1350	1350-1420
Discourse-old	11.6% (n=48)	16.0% (n=12)	— (n=0)
Discourse-new	82.3% (n=335)	74.6% (n=56)	33.3% (n=2)
Counterfactual	5.9% (n=24)	9.3% (n=7)	66.6% (n=4)
TOTAL	407	75	6

Table 2 Discourse functions associated with stage one *ne*

Function of stage two <i>ne...not</i>	1150-1250	1250-1350	1350-1420
Discourse-old	77.7% (n=205)	62.4% (n=300)	49.2% (n=29)
Discourse-new	21.6% (n=57)	36.6% (n=176)	45.8% (n=27)
Counterfactual	0.7% (n=2)	1.0% (n=5)	5.0% (n=3)
TOTAL	264	481	59

Table 3 Discourse functions associated with stage two *ne...not*

Tables 2 and 3⁶ suggest that change in the discourse functions of negative markers is an integral part of the Jespersen Cycle. As *ne...not* extends from discourse-old propositions to discourse-new ones, it becomes pragmatically unmarked, while

⁶ The use of *ne...not* in discourse-old contexts antedates the Middle English period. Looking at examples of non-argument ‘not’, that is *noht* or *nawiht* which is not assigned a theta-role, in Old English data from the York Corpus of Old English Prose (Taylor, Warner, Pintzuk & Beths 2003) dating from the period 850-1150, 82.2% (n=60/68) of them appear in discourse-old propositions. Further work remains to be done on the pragmatics of Old English negative markers, particularly to ascertain if Old English *ne...na* performs the same function as *ne...not* in Old and Middle English (with thanks to an anonymous reviewer for this suggestion).

ne becomes increasingly marked, becoming largely restricted to counterfactuals. Hansen (2009) and Hansen & Visconti (2009) argue on the basis of medieval French data that the use of French stage two *ne...pas* extends from discourse-old propositions into discourse-new propositions during the Jespersen Cycle at the expense of stage one *ne*. They assign this change a causal role in the Jespersen Cycle, arguing that through this change *ne...pas* becomes grammaticalised as the pragmatically unmarked or neutral way of marking sentential negation.

The data presented in this section seem to point to the conclusion that stage one *ne* and stage two *ne...not*, while equivalent in basic grammatical function (i.e. both markers of sentential scope negation), are highly differentiated in their discourse functions or discourse contexts in early ME. Furthermore, the discourse-functions and contexts for both negative markers seem to change over time. Is it possible to accommodate these apparent functional changes within a model of the Jespersen Cycle which assumes morphosyntactic competition between functionally equivalent forms? Section 4 explores this question in more detail. I will argue that despite these apparent changes in the discourse functions of *ne* and *ne...not* within the frequency data, the discourse functional constraints on *ne* and *ne...not* remain constant during the 13th and 14th centuries, in a way which fits within a model of the change as morphosyntactic competition.

4 THE DISCOURSE FUNCTIONS OF ME NEGATIVE MARKERS AND A MODEL OF THE JESPERSEN CYCLE AS MORPHOSYNTACTIC COMPETITION

4.1 *Issues for a model of the Jespersen Cycle as morphosyntactic competition*

The findings of the previous section appear to be problematic if we adopt a model of the Jespersen Cycle as competition between two morphosyntactic options, such as the one advocated by Wallage (2008). Wallage (2008) argues for competition between two lexemes: ne_1 at stage one of the Jespersen Cycle, which is negative in both semantics and form; and ne_2 at stage two, which is negative in form only and therefore must occur in concord with another negative such as *not* in order for the clause to be interpreted as negative semantically. The model of competition between ne_1 and ne_2 , as set out in Wallage (2008), does not take into account the discourse functions of the two negative markers. Instead, it assumes (a) that ne_1 (at stage one) and ne_2 (at stage two) are equally likely to occur in all discourse functions, and (b) that the discourse functions associated with ne_1 and ne_2 remain the same throughout the ME period.

However, the data presented in section 3.2 suggest that (a) *ne* and *not* are not equally likely to occur in all discourse contexts, and (b) that the Jespersen Cycle is characterised by ongoing change in the functions of both *ne* and *ne...not*. The increasing frequency of *ne...not* in discourse-new propositions suggests that the discourse-functional constraints operating on *ne...not* may weaken as *ne...not* becomes more frequent, so that they all but disappear by the 14th century, as the

frequencies of *ne...not* in discourse-old and discourse-new propositions converge (see Table 1, section 3.2).

These data potentially have considerable implications for our understanding of the ME Jespersen Cycle. The frequency data in section 3.2 are compatible with an account in which *ne...not* spreads through functional extension. Under such an account, the discourse-functional constraints on *ne...not* weaken over time so that the overall frequency of *ne...not* increases because it spreads from one discourse context to another. Hence it is crucial to establish whether the discourse functional constraints on *ne...not* do actually weaken as the form becomes more frequent overall. However, frequency data alone do not demonstrate whether the discourse functions associated with *ne* and *ne...not* remain constant or change over time.

Kroch (1989) observes that the progress of a morphosyntactic change over time follows a logistic curve. Since the distributions of *ne* and *ne...not* are not static but changing over time, any attempt to examine the effect of discourse function on the distribution of these two forms needs to estimate the constraining effect of these factors in a way that takes into account the fact that the distribution of the forms changes over time in a non-linear way. Thus we need to ask two inter-related questions. First, does the competition between *ne* and *ne...not* progress at the same rate in all contexts – that is, does competition between *ne* and *ne...not* result in a single logistic curve with the same slope in all discourse functions, or does it proceed at a faster rate in some contexts than in others? Second, does the effect of discourse function on the distribution of *ne...not* remain constant or does it change over time? Do *ne* and *ne...not* more strongly correlate with different discourse functions at earlier points in the change than at later points, or does the strength of these correlations remain the same in probabilistic terms at all points throughout the change, whatever the overall frequencies of *ne* and *ne...not* happen to be? If the rates of change are the same in all functions and the effect of contextual factors constant, this forms a strong argument for regarding the change from *ne* to *ne...not* as a single change across all discourse functions. The change is thereby independent of the discourse-functional constraints on *ne* and *ne...not*. However, if *ne...not* spreads by extension across each discourse function, we would predict that the effect of contextual factors changes over time. Any correlation between negative markers and discourse function would weaken over time until it ceases to be significant.

It is not possible to answer these questions using frequency data alone. The frequency data presented in Table 1 (section 3.2) appear to show the frequencies of *ne...not* in discourse-old and discourse-new contexts converging over time, and appear to show the frequency of *ne...not* increasing much more in discourse-new contexts than discourse-old ones during the 13th and 14th centuries. However, this does not necessarily mean that there are different changes in discourse-old and discourse-new contexts, that follow different logistic curves and progress at different rates. In fact, the apparently greater increase in *ne...not* in discourse-new

contexts during the 13th and 14th centuries could result even if the introductions of *ne...not* proceeds at the same rate in discourse-old contexts and discourse-new contexts and thus follows the same logistic curve in both contexts, as Figure 1 illustrates.

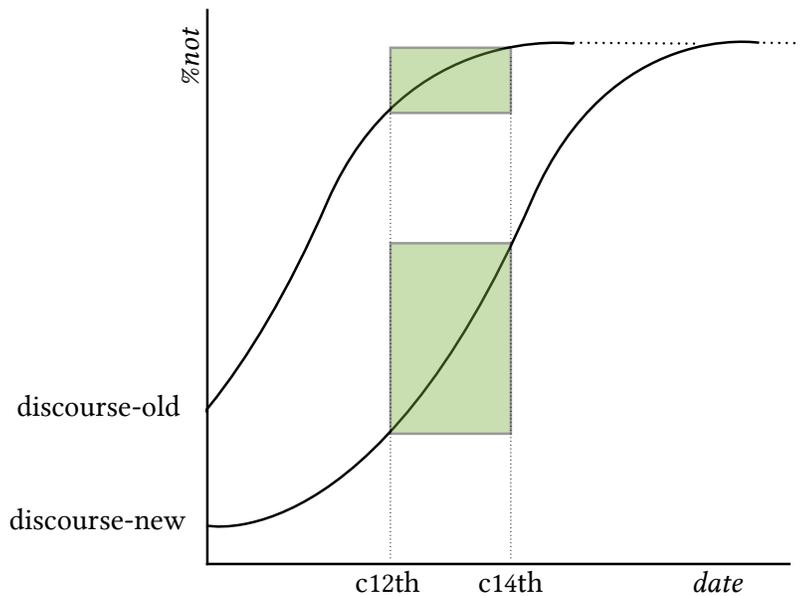


Figure 1 The increased frequency of *ne...not* modelled using the same logistic curve in discourse-old and discourse-new contexts.

Figure 1 shows a pair of parallel logistic curves in which the rate of change is the same. The only difference between them is that the frequency of *ne...not* starts out higher in one than in the other (observe the different points at which the two curves intersect the Y-axis). The upper curve only shows an 11% increase in the innovative form during the 13th century, whereas the lower curve shows a 55% increase in the innovative form during the same period, simply because the frequency of the innovative form starts out much lower on the second curve. The frequencies of *ne...not* represented by the two curves seem to converge over time as the change progresses beyond the 13th century, even though the two curves have the same slope, indicating that *ne...not* is introduced at the same rate in both contexts.

Furthermore, if both curves change at the same rate, it follows that the relationship between the curves, as manifest in the initial differences between the two, must remain the same at all points on those curves. This manifests itself, not in terms of the frequencies of *ne...not* at successive points on those two curves, but

probabilistically in terms of how much more likely discourse-old contexts are to involve *ne...not* than discourse-new ones are at each particular point in time.

Competition between *ne* and *ne...not* will produce a logistic curve in each of our five discourse contexts. Therefore, in order to see whether the change from *ne* to *ne...not* progresses at the same rate in each context, we need to ascertain whether these curves are parallel – that is, whether they have the same slope, and whether the relationship between the curves is the same at successive points of the change. If this is the case, then the overall increase in *ne...not* will be independent of these discourse contexts, and these discourse-contextual constraints will not themselves change over time. Kroch proposes a method to establish whether or not this is the case:

A constant rate of change across contexts is mathematically equivalent to fixity of contextual effects, in direction and size, across time periods. Thus, if a study reports a series of multivariate analyses for different time periods, and the contextual effects are constant across these analyses, the rate of change of each context measured separately would necessarily be the same. This equivalence holds because, in statistical terms, the constant rate hypothesis is the claim is that the overall rate of use of a form is independent of the contextual effects on its use. (Kroch 1989: 204)

What we need then is a series of multivariate logistic regression analyses at different time periods. These convert frequency data into a format from which we can estimate the contextual effects of discourse function on the distribution of *ne...not* and compare them across different points of the change. We estimate these effects in probabilistic terms: the factor weight for *ne...not* in each context is an estimation of the difference between the likelihood of finding it in that particular context and finding it in the whole dataset. In other words, the factor weight indicates how much more or less likely *ne...not* is to appear in each individual discourse function than in any other function. Therefore the factor weights for each context can be compared across successive periods irrespective of differences in the overall frequency of *ne...not* over time.

A further issue for a logistic regression analysis, not yet mentioned, concerns the interaction of discourse function with other factors pertinent to the distribution of *ne*, *ne...not* and *not*. Wallage (2008) observes a difference between the frequencies of *ne*, *ne...not* and *not* in main clause declaratives and subordinate clauses, particularly a higher frequency of *ne* in conditional *if* clauses (13) and clauses which appear within the scope of negation (14), as shown in Table 4.

- (13) *And also þenk what-maner þou art bisi and turmented*
 And also think what-manner you are fearful and tormented

aboute mony þinges 3if þou hem ne haue
about many things if you them NEG have

‘And also think in what manner you are fearful and tormented about many things if you do not have them’ (CMEDVERN,242.98)

- (14) *‘it may nat be,’ seith he, “that where greet fyr hath long
“it may not be,” says he, “that where great fire has long
tyme endured that ther ne dwelleth som vapour of
time endured that there NEG dwells some vapour of
warmness.”*
warmness.”

‘ “it may not be,” he said, “that where a great fire has been for a long time that it isn’t warm there” ’ (CMCTMELI,223.C2.269)

	1150-1250			1150-1250			1150-1250		
	ne	ne...not	not	ne	ne...not	not	ne	ne...not	not
Main	36.8% (n=84)	62.3% (n=142)	0.9% (n=2)	11.0% (n=45)	78.0% (n=320)	11.0% (n=45)	— (n=0)	9.1% (n=23)	90.9% (n=230)
Subordinate	70.6% (n=264)	29.1% (n=109)	0.3% (n=1)	33.6% (n=81)	60.6% (n=146)	5.8% (n=14)	0.4% (n=1)	12.0% (n=30)	87.6% (n=219)
If clause	76.1% (n=32)	23.9% (n=10)	— (n=0)	52.4% (n=11)	33.3% (n=7)	14.3% (n=3)	10.5% (n=2)	5.3% (n=1)	84.2% (n=16)
In scope of nega- tion	90.0% (n=27)	10.0% (n=3)	— (n=0)	67.7% (n=21)	25.8% (n=8)	6.5% (n=2)	13.0% (n=3)	21.7% (n=5)	62.2% (n=15)
TOTAL	407	264	3	158	481	64	6	59	480

Table 4 The distribution of clauses with *ne*, *ne...not* and *not* in each grammatical context

Although Wallage (2008) argued that the effect of clause type is consistent (in probabilistic terms) from the 13th century until stage one *ne* is lost in the 15th century, this may not be the case once we consider the interaction of clause type and discourse function over time. Indeed, the clause type effect observed by Wallage (2008) may result from the different discourse functions associated with main and subordinate clauses, if subordinate clauses tend to introduce discourse-new propositions and main clauses discourse-old propositions.

4.2 *The effect of discourse function within a logistic regression model*

In order to estimate the probabilistic effect of discourse function on the distributions of *ne* and *ne...not*, two separate multivariate logistic regression analyses are carried out for the periods 1150-1250 and 1250-1350. The resulting factor weights from each of the two periods can then be compared to ascertain if these discourse functional constraints have the same probabilistic effect on the distribution of *ne* and *ne...not* in both periods, or whether the effect of these constraints changes over time.

Unfortunately, in the data for the period 1350-1420 there are no examples of *ne* in discourse-old contexts. Thus discourse-old contexts are not a variable context for *ne*. Hence it is not possible to perform regression analysis to compare the distribution of *ne* and *ne...not* in discourse-new and discourse-old contexts during the period 1350-1420.

Regression analyses are therefore carried out on data from the periods 1150-1250 and 1250-1350. In order to identify any interaction between them, both clause type and discourse function are included as independent variables. The dependent variable is the form of negative marker: stage one *ne* or stage two *ne...not*. The factor weights represent the likelihood that *ne...not* will appear in each context. The input probability represents the likelihood that *ne...not* will occur in all the data for a particular period. Then the analysis produces a factor weight for each context, on a scale of .01 to .99. A context with a factor weight of .01 is one in which *ne...not* is least likely to occur. Conversely, a context with a factor weight of .99 is one where *ne...not* is most likely to occur. Contexts where *ne...not* is favoured, that is more likely to appear than average for the period, have a weight of greater than .5. Those contexts where it is less likely to occur than average for the period have a weight of less than .5. The analysis reports a *p*-value which is a measure of how significant the differences between the contexts are within the analysis. As is typical, I assume the threshold for statistical significance is $p \leq .05$.

The quantitative data on which the regression analysis is based are the data in Tables 1 and 4 in section 3.2. The results of the two logistic regression analyses are presented in Table 5. There are some gaps in Table 5 for which factor weights cannot be estimated. There are two reasons why it is not possible to estimate a factor weight. First, there may be no variation between *ne* and *ne...not* in a particular context. In the period 1250-1350, propositions which cancel an inference are not a variable context. The form *ne* is completely absent and *ne...not* is used in all examples in this

context, hence regression analysis cannot be carried out. Second, some of the gaps in Table 5 occur when there are too few data in a particular context for multivariate analysis to yield reliable factor weights – that is, fewer than 5 instances either of *ne* or of *ne...not*. In the period 1150-1250, there are only 3 instances of *ne* in propositions which repeat an antecedent proposition, and only 1 instance in the period 1250-1350. Hence the context “Repetition of an antecedent proposition” is excluded from the analysis for both periods. Similarly, because there are only 3 instances of *ne* used to deny an antecedent proposition in the period 1250-1350, the context “Denial of an antecedent proposition” is excluded from the regression analysis for this period. Hence, Table 5 shows the likelihood of *ne...not* in two discourse contexts for which there are sufficient data to permit regression analysis, one discourse-new, the other discourse-old.

	1150-1250	1250-1350
Input probability	.359	.870
Clause type		
Main	.686	.669
Subordinate	.401	.310
If-clause	.371	.216
(Range)	(315)	(453)
Discourse function		
Denial of antecedent p	.839	
Repetition of antecedent p		
Statement of inference	.899	.880
Cancellation of inference	.900	
Discourse-new	.234	.196
(Range)	(666)	(684)
P ≤	.0001	.0001

Table 5 Logistic regression analysis of the competition between stage one *ne* and stage two *ne...not*

Table 5 shows that despite a very large increase in the likelihood of *ne...not* in the dataset overall (the input probability), the distinctions between both the different clause types and the different discourse functions remain generally consistent across the two periods in probabilistic terms, at least for those contexts which have sufficient data from which to estimate a factor weight. In particular, the likelihood of finding *ne...not* in discourse-new contexts remains remarkably consistent across

the two periods, irrespective of the increasing overall frequency of *ne...not*. The very low p-values indicate that discourse function continues to be a very highly significant factor in both periods in spite of the increasing overall frequency of *ne...not* in the period 1250-1350.

5 CONCLUSION

The data and findings presented in this paper point to two main conclusions. First, although equivalent in their basic syntactic function as markers of sentential scope negation, stage one *ne* and stage two *ne...not* are highly functionally differentiated throughout Middle English in terms of their use across different discourse contexts. These constraints on their use mean that *ne* and *ne...not* are not simply functionally equivalent diachronic stages within the Jespersen Cycle, but are specialised for different discourse functions. This provides an explanation for the persistence of variation between stage one *ne* and stage two *ne...not* over more than three centuries in the ME period. Stage one *ne* continues to be favoured in a particular set of discourse functions even in the 14th century.

Second, although the changing frequencies of *ne*, *ne...not* and *not* seem to suggest a functional shift of the kind proposed by Hansen (2009) and Hansen & Visconti (2009) for medieval French *ne...pas*, when we model discourse-status as a constraint on the distribution of ME *ne* and *ne...not* within a regression analysis, its effect remains constant in probabilistic terms throughout the 13th and 14th centuries – at least in the two contexts for which regression analysis can be performed reliably in both periods – despite the increasing overall frequency of *ne...not* at that time.

Although it is possible that pragmatic change may be implicated in the Jespersen Cycle in a way that this analysis does not consider, the regression data presented here are consistent with the conclusion that the pragmatic or discourse-functional distinctions examined in this paper have a consistent effect on the distributions of *ne* and *ne...not* throughout early ME. It follows that *ne...not* replaces *ne* at the same rate in all propositions, irrespective of their discourse status. So *ne...not* remains favoured in discourse-old propositions, and *ne* favoured in discourse-new ones, throughout the 12th-14th centuries, until the form *ne* is itself lost. Stage two *ne...not* does not become pragmatically unmarked, or neutral, as it becomes more frequent. It remains pragmatically marked even in 14th century English, despite its increasing overall frequency. Thus, an appeal to loss of these particular pragmatic constraints does not explain the increasing overall frequency of *ne...not* in early ME. The constraints are not lost as the frequency of *ne...not* increases. Instead, these data indicate that the discourse-old/discourse-new distinction is independent of the increasing overall frequency of *ne...not*.

REFERENCES

- van der Auwera, Johan. 2009. The Jespersen Cycles. In E. van Gelderen (ed.), *Cyclical change*, 35–71. Amsterdam: John Benjamins.
- Detges, Ulrich & Richard WALTEREIT. 2002. Grammaticalisation vs. reanalysis: a semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21(2). 151–195.
- Dryer, Matthew S. 1996. Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics* 26(4). 475–523.
- Hansen, Maj-Britt Mosegaard. 2009. The grammaticalisation of negative reinforcers in Old and Middle French. In M.-B. Mosegaard Hansen & J. Visconti (eds.), *Current Trends in Diachronic Semantics and Pragmatics*, 227–251. Bingley: Emerald Publishing.
- Hansen, Maj-Britt Mosegaard & Jacqueline Visconti. 2009. On the diachrony of “reinforced” negation in French and Italian. In C. Rossari, C. Ricci & A. Spiridon (eds.), *Grammaticalisation and Pragmatics: facts, approaches, theoretical issues*, 137–171. Bingley: Emerald Publishing.
- Horn, Laurence R. 2001. *A natural history of negation*. Stanford: CSLI.
- Israel, Michael. 2001. Minimisers, maximisers and the rhetoric of scalar reasoning. *Journal of Semantics* 18. 279–331.
- Jespersen, Otto. 1917. *Negation in English and Other Languages*. Kobenhavn: Enjar Munksgaard.
- Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1. 199–248.
- Kroch, Anthony S. & Ann Taylor. 2000. *The Penn-Helsinki Parsed Corpus of Middle English*. Second Edition. Philadelphia: University of Pennsylvania.
- Schwegler, Armin. 1988. Word-order changes in predicate negation strategies in Romance languages. *Diachronica* 5(1). 21–58.
- Schwenter, Scott. 2006. Fine tuning Jespersen’s Cycle. In L.R. Horn, B.J. Birner & G.L. Ward (eds.), *Drawing the boundaries of meaning: neo-Gricean studies in pragmatics and semantics in honor of Laurence R. Horn*, 327–344. Amsterdam: John Benjamins.
- Taylor, Ann, Anthony Warner, Susan Pintzuk & Frank Beths. 2003. The York-Toronto-Helsinki Parsed Corpus of Old English Prose. York: Department of Language and Linguistic Science, University of York. Distributed through the Oxford Text Archive.
- Wallage, Phillip W. 2008. Jespersen’s Cycle in Middle English: parametric variation and grammatical competition. *Lingua* 118. 643–674.
- van der Wurff, W. 1999. On Expletive Negation with Adversative Predicates in the History of English. In I. T.-B. van Ostade, G. Tottie & W. van der Wurff (eds.), *Negation in the History of English*, 295–329. Berlin: Mouton de Gruyter.

Functional differentiation and grammatical competition in the English Jespersen Cycle

Phillip Wallage
Faculty of Arts, Design and Social Sciences
Northumbria University
Newcastle upon Tyne
NE1 8ST, United Kingdom
phillip.wallage@northumbria.ac.uk