13

Persistence as a diagnostic of grammatical status: The case of Middle English negation

AARON ECAY AND MEREDITH TAMMINGA

13.1 Introduction

The field of diachronic syntax encompasses the analyses of both historical grammatical structures and the processes by which they change. Processes of change are in many cases easily measured by changes in surface patterns, but without access to native speakers and their grammaticality judgements, theoretical analysis of the syntactic structures of past language stages is at a disadvantage. This analysis is impeded by our inability to collect new data probing crucial minimal pairs or rarely-occurring contexts, as well as by our own lack of intuitions about the variety. As diachronic syntacticians, then, we are pressed to be resourceful and creative in our analytic approaches.

One method by which researchers have made headway on the problem of analysing underlying structural unity (or disunity) across surface strings in historical data is by application of the Constant Rate Hypothesis (CRH; Kroch 1989*b*). In this approach, quantitative data on rates of change across different syntactic contexts is used to infer grammatical analyses. Although the CRH has been applied successfully to a range of cases, it is not the only logically possible solution to the methodological dilemma at hand. In this chapter we propose an independent source of quantitative evidence about historical grammatical analyses: the observation of repetitiveness in variant choice, or *persistence*. Persistence in corpus data is generally understood to reflect the psycholinguistic mechanism of structural priming, whereby reuse of a recently-processed syntactic structure is facilitated. The reasoning underlying this type of evidence is quite distinct from the application of the CRH, thus allowing a single corpus to generate multiple strains of quantitative evidence. We present a case study, namely the change in negation in Middle English, for which CRH evidence has been extensively investigated without yielding a definitive conclusion. Our argument is that persistence evidence favours one extant analysis over its competitor, thereby demonstrating the utility of persistence as a diagnostic tool that can be used in tandem with other approaches.

Micro-change and Macro-change in Diachronic Syntax. First edition. Eric Mathieu and Robert Truswell (eds)
This chapter © Aaron Ecay and Meredith Tamminga 2017. First published 2017 by Oxford University Press.

The chapter is organized as follows. In Section 13.2, we discuss the literature on corpus persistence and describe the assumptions necessary to link written historical data to psycholinguistic processes. In Section 13.3, we introduce the syntactic change that forms our case study, the replacement of the sentence negator *ne* with *not* by way of a period where both co-occur. We outline two competing accounts for the structures at play during the period of co-occurence; one account due to Frisch (1997) and the other to Wallage (2008). Section 13.4 first makes explicit the predictions that these two accounts make about the expected persistence behaviour of the negators, then tests these predictions against quantitative data from the PPCME2 (Kroch and Taylor 2000). Finally, in Section 13.5 we discuss the implications of our results and the potential for more widespread application of the methodology we propose here.

13.2 Background: Persistence

In quantitative studies of language variation, it is generally assumed that each token is an independent observation of a random variable. While this is a useful simplifying assumption, it is widely known not to hold as strictly as a researcher might hope (for an early discussion, see Sankoff and Laberge 1978; Paolillo 2011 offers a more general critique of independence assumptions in quantitative linguistics). *Persistence* is the tendency for a recently-used variant of a linguistic variable to be used again—each token is to some extent dependent on the token or tokens before it. Such repetitiveness in variant choice is of interest in its own right, but in this chapter we focus on its use as a tool for understanding the relationships between grammatical strings. By taking persistence as a dependent variable that is sensitive to the structural identity of sequential constructions in written corpora, it is possible to determine whether those constructions are related or not. This section briefly outlines the assumptions necessary to make such a claim and the previous research supporting those assumptions.

Early demonstrations of the persistence effect come from spoken corpora. The earliest such study, to our knowledge, is Sankoff and Laberge (1978), which documents a persistence effect on three variables in Montreal French: the pronominal alternations between *on* and *tu*/*vous* for general indefinite human reference, *on* and *ils* for exclusive indefinite reference, and *nous* and *on* for first person plural. The authors show that in sequences of tokens, speakers switch from one option to the other only about one-third as often as the null hypothesis of total statistical independence would predict. Another influential early demonstration of such an effect comes from Poplack (1980, 1984), two studies of the factors conditioning the variable deletion of inflectional /s/ and /n/ in Puerto Rican Spanish. In these papers, Poplack shows that preceding /s/-deletion within the noun phrase favours further deletion, while preceding retentions disfavour deletion. In her terms, 'One marker leads to more, but zeros lead to zeros' (Poplack 1980: 378). Poplack and Tagliamonte, in their study of verbal /s/ in nonstandard English, similarly demonstrate that 'once a zero is used . . . another zero is most likely' (1989: 70).

One of the earliest corpus studies to identify a persistence effect for a syntactic variable is Weiner and Labov (1983). This study of the constraints governing selection of the active or passive voice, which the authors treat as functional alternants, finds that while the external factors of style, sex, class, ethnicity, and age have small-to-insignificant effects on the selection of the active or the passive, the internal effects of information structure (given vs. new) and parallel surface structure (persistence) are large and significant. They conclude that 'the ordering of surface syntax across clauses is the predominant linguistic influence on this choice' (Weiner and Labov 1983: 52). This conclusion holds even when excluding from the analysis clause pairs with coreferential subjects. The same variable is revisited in more detail by Estival (1985). More recent corpus-based studies of syntactic persistence include Gries (2005), in which the variables under consideration are the dative alternation and particle placement, and Szmrecsanyi (2006) on future marking, particle placement, complementation, comparatives, and genitives in English dialects. The Gries study is particularly informative because he finds comparable results using written and spoken data sources, thus bolstering the viability of persistence studies in the context of historical syntactic research.

The corpus persistence literature is widely understood to be intimately linked with the experimental structural priming literature, in which speakers who have heard a particular syntactic construction are more likely to use that construction in a picture description task. Pickering and Branigan (1999: 136) define structural priming as 'the phenomenon whereby the act of processing an utterance with a particular form facilitates processing a subsequent utterance with the same or a related form'. The pioneering paper in this domain, Bock (1986), was motivated by the corpus results from Estival (1985) and Weiner and Labov (1983) described above. In Bock's original methodology, which heavily influenced subsequent studies, participants are exposed to sentences and then asked to describe pictures. The syntactic structure the participant chooses for the description (passive or active, double object or prepositional object) is the experimental target and is dependent on the structures in sentences that the participant is exposed to as primes. Pickering and Ferreira (2008) identify sentence recall, written sentence completion, and spoken sentence completion as additional methodologies that have subsequently been added to the structural priming toolbox. They also provide an overview of the wide range of results on the robustness and sensitivity of structural priming.

The experimental structural priming literature has made progress by exploiting priming relationships to answer questions of structural identity. The general mode of reasoning is that experimental observations of priming relationships between different syntactic structures are taken as evidence for the nature of the mental representation of the underlying linguistic objects. Branigan *et al.* (1995) lay out the argument explicitly:

> If the processing of a stimulus affects the processing of another stimulus, then the two stimuli must be related along a dimension that is relevant to the cognitive system. Under certain circumstances, we can conclude that they are represented in a related manner. If the relationship between two stimuli is syntactic, then we can use this relationship as a way of understanding what syntactic information is represented, and how that information can interact with other information.' (Branign *et al.* 1995: 490)

Pickering and Branigan also argue this point when they 'claim that syntactic priming taps into knowledge of language itself, and as such can inform linguistic theories

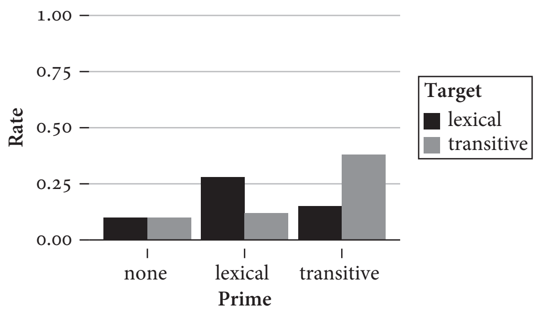

Figure 13.1. Results from Estival (1985), showing that English lexical and transformational passives prime themselves, but not each other, in conversational speech

that are concerned with accounting for knowledge of language' (1999: 140). They specifically advocate the use of priming studies to determine relationships between sentences. Taking the argument one step further, Pickering and Ferreira (2008: 429) suggest that in the face of overwhelming evidence for structural priming in many constructions, a failure to observe priming for a specific construction may indicate that the putative prime and target are *not* representationally related.

A handful of studies have taken up these types of grammatical questions. Estival (1985) asks whether the persistence effect holds across different types of passives or whether it shows a within-type restriction. The two types of passives she considers are transformational passives ('John is believed to have left') and lexical passives ('John is interested in music'—the examples are hers). She finds that each type of passive makes it more likely that a passive of its type will be used again, but that the effect does not extend to passives of the other type; this is illustrated in Figure 13.1.[1]

One early experimental demonstration that priming reflects structural identity beyond simple surface repetition comes from Bock and Loebell (1990). They show through the use of a picture description task that prime sentences with PPs introduced by *to* facilitate the choice of the prepositional structure over the double-object one in subjects' subsequent production of ditransitive sentences. This relationship holds regardless of whether the *to*-phrase in the prime sentence introduces a beneficiary ('The wealthy widow gave her Mercedes to the church') or a location ('The wealthy widow drove her Mercedes to the church'). The same priming is not observed, on the other hand, when the prime contains *to* functioning as an infinitival marker ('Susan brought a book to study').

In a more recent paper, Ferreira (2003) investigates the sensitivity of the optional complementizer *that* to priming by other functions of *that*. Using a sentence recall task, he shows that when the prime *that* heads a CP which is the object of a verb as in (1), it boosts the rate at which subjects choose to use *that* in their productions in just

[1] Estival's study was carried out at a time when the nature of these different passives was in question, but her result is in line with the modern consensus that the two types of passives are structurally distinct (see Embick 2004 for discussion).

that context (where it is optional). But when the prime is a determiner as in (2) or the head of a noun's complement as in (3), no such priming effect results.[2]

(1) The company ensured that the farm was covered for two million dollars.

(2) The company insured that farm for two million dollars.

(3) The theory that penguins built the igloos was completely false.

The core methodological premises of this chapter, then, are twofold: that persistence effects observed in written corpora reflect priming effects in the language production of the time period; and that repetition reveals grammatical sameness while lack of repetition indicates grammatical difference. In the context of historical syntax, where we have no access to native speaker judgements or experimental behaviour, there is a need for as many independent sources of evidence as possible. When different sources of evidence for a theoretical analysis converge, it considerably strengthens our confidence in that analysis. We make the case that persistence can be one of these sources of evidence using the case study of a change in sentence negation in Middle English.

13.3 Background: Negation

In the Middle English period, there is a change in the exponence of sentence negation (represented as the category Neg in our syntactic diagrams). The negator *ne*, which was inherited from Old English, is eventually lost, being replaced by *not*, originally an adverb of emphatic negation.[3] During the period when this change is ongoing, there are a large number of sentences which appear with both *ne* and *not*, as in the following example.

(4) he ne shal nouȝt decieue him (Early Prose Psalter, 161:131:11, from Frisch 1997)

The progress of this change is charted in Figure 13.2.

13.3.1 Frisch (1997)

The literature contains two different analyses of the grammatical underpinnings of this change. The first is due to Frisch (1997), who analyses the change as stemming from competition between (just) two grammars. The first grammar, inherited from

[2] A reviewer points out that many syntactic theories treat the *that*s in (1) and (3) as the same lexical item. If these theories are taken to be true (and we see no compelling reason not to do so), the lack of priming effects between sentences like (1) and (3) needs to be explained. Ferreira's data indicate that lexical and categorial identity is a necessary but not sufficient condition for syntactic priming. A contextual factor is also at play—*that* alternates with Ø in (1), but is obligatory in (3). Precisely characterizing the contextual conditions on priming effects in terms of grammatical structures is an open question for psycholinguistic research, and clearly has implications for the generality of our proposed corpus-based diagnostic. However in the context of this paper, where we compare two analyses of the same construction rather than looking at the behaviour of lexically identical material across constructions, the issue does not arise.

[3] *Not* could also be used as a DP, with approximately the same meaning as modern *nothing*. In modern standard English this use of the word is spelled *naught* or *nought*.

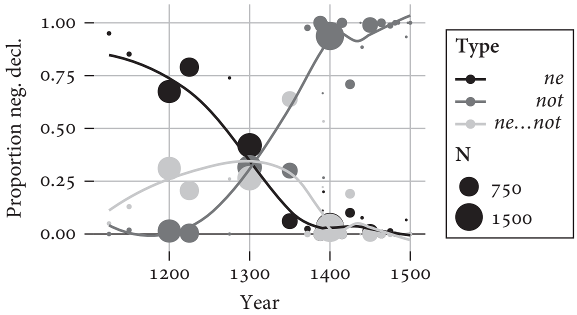

Figure 13.2. The change from *ne* to *not* in Middle English, as measured in negative declarative sentences in the PPCME2.

Note: Sentences with both *ne* and *not* are a separate category from sentences with just one or the other negator. The sizes of the circles represent the total number of negative declarative tokens in a time period (the denominator of the proportion), so triplets of points in the same year are identically sized. See Section 13.4.1 for details on how the dataset was collected.

OE, contains an entry for *ne* as the head of NegP.[4] The second, innovative, grammar contains *not* as the specifier of NegP. In both grammars, the single negator contributes the semantics of sentence negation to the derivation. Because one grammar contains an entry for a head and the other for a specifier, there is no reason why the negator from both grammars could not be merged in a single derivation (this would create a situation with some parallels to the phenomenon of code-switching, though Frisch does not invoke that term). This is precisely the situation that, on Frisch's account, causes *ne . . . not* sentences to surface.

Since both lexical entries have the semantics of negation, it might be feared that merging both would result in double negation (in the logical sense, i.e. $\neg\neg p \equiv p$). Frisch accounts for this by putting forth an Economy of Projection principle, taken from Speas (1994). Under his account, it is the presence in the derivation of NegP (the maximal projection of Neg^{0}) which contributes negative semantics. The presence of either *ne* or *not* is a sufficient condition for the projection of NegP, but neither is necessary; both together introduce only one NegP. The structures of the different options are presented in Figure 13.3.

The evidence that Frisch musters for his analysis is as follows. First, he must distinguish between the use of *not* as sentence negation and as an (emphatic) sentence adverb. This is not always possible for a given token; however, it is clear that occurrences of *not* before a tensed verb must be sentence adverbs. The adverb *never* has a similar distribution to sentence-adverb *not*, and occurs preverbally 16 per cent of the time in Frisch's measurement. Thus, Frisch infers a similar rate of preverbal position-

[4] Consistent with the Borer–Chomsky Conjecture (Baker 1998), we are using the term 'grammar' to mean a list of lexical entries for functional categories. Because *ne* and *not* are not mutually exclusive, it is possible to activate both in the structure of a single clause. This lack of mutual exclusivity differs from classical cases of grammar competition as discussed by Kroch (1989*b*). It is clear from the developments in ME that *not* replaces *ne*, but this competition is not necessitated by the structure of UG. Instead it may be driven by pragmatic concerns; see Ahern and Clark (to Appear) for discussion.

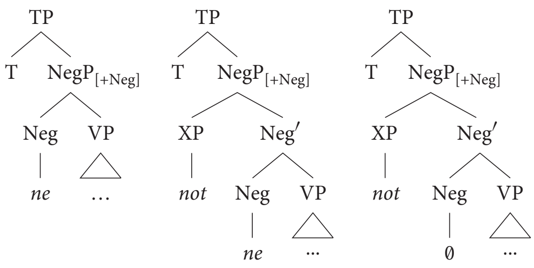

Figure 13.3. The analysis of three stages in the evolution of English negation, from Frisch (1997).

Note: The notation [+Neg] indicates where the semantics of sentence negation is introduced—in all cases, at the NegP level.

ing for *not*, and thus that the number of sentence adverb *not* occurrences in a given time period is the number of preverbal occurrences divided by 0.16 (or equivalently, multiplied by 6.25).

Frisch argues that there is no single grammatical change underlying the shift from *ne* to *not*. By the CRH (Kroch 1989*b*) the trajectories of different surface realizations of a single change will be the same. Unless exogenous complications present themselves, the natural spread of a language change through a population will proceed along an S-shaped curve in probability space, which is equivalent to a straight line in logit-transformed space. In Frisch's data, the decline of *ne* and the rise of *not* proceed with different slopes in logit space, leading him to conclude that the two phenomena are each linked to different underlying changes.

He also argues that co-occurrences of *ne* and *not* arise from the combination of two independent lexical insertions. In order to do this, he demonstrates in his data that across various time periods the rate of occurrence of *ne . . . not* is roughly equal to the product of the rate of occurrence of *ne* and the rate of occurrence of *not* (in both cases, the relevant measurement is of all the occurrences of these negators, whether alone or in combination with the other). In other words, insertion of either negator is governed by its own binary random process (effectively, a coin flip). Seeing both negators in the same sentence is the result of both coins happening to turn up heads at once, rather than a face of a special three-sided coin.

13.3.2 Wallage (2008)

The second analysis of the change comes from Wallage (2008) and is based on typologies of Jespersen's Cycle (JC; Jespersen 1917). Wallage holds that *ne*, *ne . . . not*, and *not* are each stages of an instantiation of JC in English, and that all have a separate grammatical identity. In *ne . . . not* constructions, the *ne* that appears does not have negative force, and is rather a manifestation of negative concord (or a similar process driven by formal agreement between negative features and Neg°). The various structures needed for this analysis are illustrated in Figure 13.4.

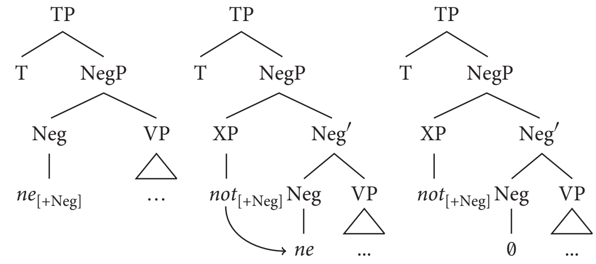

Figure 13.4. The analysis of three stages in the evolution of English negation, from Wallage (2008).

Note: The notation [+Neg] indicates where the semantics of sentence negation is introduced. An arrow indicates negative concord (or a similar agreement relationship).

Wallage provides several pieces of evidence arguing against Frisch's account. For example, Wallage shows that the distribution of *ne* appearing by itself differs between main and (several different types of) subordinate clauses. On the other hand, the distribution of *ne … not* is identical across clause types. This distinction militates against the hypothesis that there is a unified process generating *ne* when it appears with and without *not*. Wallage goes on to show that the loss of *ne* by itself in main and subordinate clauses occurs with the same trajectory (logistic regression slope), thus providing evidence that *ne* by itself shares an underlying cause in main and subordinate clauses, which differs from the cause of *ne … not*.

The second part of Wallage's argument, that the *ne* of *ne … not* constructions is generated by negative concord, is based on the comparison of different types of 'redundant *ne*' in ME. The two types of redundant *ne* are (A) that licensed by a higher negative (negating a verb of doubt) as in (5) and (B) that licensed by a verb of prohibition or denial as in (6).

(5) *ne doute the nat that alle thinges ne ben don aryght*
NEG doubt you not that all things NEG are done rightfully
'Do not doubt that all things are done rightfully.'
(Chaucer's Boethius; Wallage 2008: (25a))

(6) *Iesus hire þo for-bed þat heo attryne ne sceolde his hond ne his fet*
Jesus her though forbade that she bind NEG ought his hands nor his feet
'though Jesus forbade her to bind his hands or his feet'
(Passion 581; Wallage 2008: (24b))

Further examples of the second type can also be found in Early Modern English as in (7), whereas the first type is lost during the ME period.

(7) You may deny that you were not the meane of my Lord Hastings late imprisonment (Shakespeare, *Richard III*)

Wallage claims that the *ne* that appears jointly with *not* is another instantiation of type A redundant *ne*—that is, licensed by the presence of a Neg projection.[5] This claim, insofar as it links the behaviour of sentential negation to other relevant linguistic phenomena, strengthens Wallage's overall picture.

13.3.3 Summary

There is a fundamental disagreement between Frisch's account and Wallage's about how to analyse the change in ME negation in terms of grammatical structures. This disagreement boils down to the question of whether there are two or three atomic objects interacting in speaker's–learners' minds during the change. Each position is supported by some amount of quantitative evidence analysed according to the CRH paradigm. The corpus of ME is strictly limited in size, and Frisch's and Wallage's analyses have already incorporated most of it. The amount of further progress that CRH-based approaches can make is thus also quite limited. We propose instead to inject a new methodology into the debate. Priming is a quantitative diagnostic which, like the CRH, is attuned to the grammatical status of constructions. An appropriately-constructed examination of priming effects (or the lack thereof) in the available ME corpus will shed new light on the question of which of these two analyses is correct.

13.4 Results

13.4.1 Background

The dataset used in this investigation was derived from the Penn Parsed Corpus of Middle English, version 2 (PPCME2; Kroch and Taylor 2000).[6] All tokens of negative declarative clauses were extracted from the corpus, and matched into consecutive pairs. We placed no limit on the number of affirmative clauses intervening between members of a pair. However, the intervening occurrence of a negative clause other than a declarative (such as a question) would prevent two negative declaratives from being joined into a pair.[7] For the bulk of the study, we focus on the period from 1250 to 1350, which is the approximate middle of the change when all three surface variants are roughly equiprobable. This maximizes the opportunity to find unequal variants in paired contact with each other, and thus the greatest opportunity to observe prim-

[5] Wallage's presentation of the proposal is not without some degree of hedging: 'Hence we might try to link change in the distribution of redundant *ne* to the change which happens in sentential negation contexts. The effect of change in both contexts is the same: restriction of *ne* to contexts in which another negative [i.e. *not* in the sentential negation case—AE&MT] is present' (Wallage 2008: 667). This proposal generates a CRH prediction which is not tested.

[6] Specifically, Penn's copy of the working files as of 20 June 2012 were used, incorporating various fixes since the last release. Layamon's *Brut*, a poetic text not included in any release, was excluded.

[7] We also excluded the following constructions, which introduce tokens of *ne* or *not* which are not part of the envelope of variation: contracted *ne* as in *nis* = *ne is*; *ne* in concord with *none* or *never*; *not only* and similar constructions; probable cases of *not* as constituent negation of verbs and adverbs. The interested reader is referred to column 4 of the `queries/coding.c` file in the associated GitHub repository for specific implementational details of these exclusions.

TABLE 13.1. Counts in each prime/target position of our primary negation dataset

		Target			
		ne	*not*	both	Sum
Prime	*ne*	124	13	58	195
	not	20	156	39	215
	both	69	41	74	184
	Sum	213	210	171	594

The dataset comprises PPCME2 texts from 1250 to 1350 (inclusive), with specific characteristics as described in Section 13.4.1.

ing effects. The corpus contains 598 target–prime pairs during this time period. The counts within each prime/target condition are given in Table 13.1.

In examining the data, we looked for evidence of priming effects (or the lack thereof) by examining each prime condition separately. These are the divisions along the *x*-axis of our plots. Within each condition, we compared the rate of realization of the possible targets. If the rate of realization of each target type is equal to the others within a single prime condition, there is no effect of the prime type on the realization of the target—that is, there is no priming. On the other hand, if there is a difference in target realization proportions, this constitutes a priming effect.[8]

13.4.2 Two-atom model

The predictions of Frisch's two-atom model are shown in Figure 13.7(a). Under the independence assumptions of this model, we expect that uses of *ne* alone will facilitate following *ne* (alone or with *not*), and similarly for *not* alone. We also predict that tokens of both negators together will have the same effect as *ne* alone on following use of *ne*, and similarly for *not*. This prediction is not borne out completely, as illustrated in Figure 13.5. We expect the middle dark grey bar (measuring the priming effect of *ne . . . not* on a *ne* target) to be as tall as the middle light grey one, and the middle light grey bar (*ne . . . not* on a *not* target) to be as tall as the right-hand one. The binomial confidence intervals, represented by the error bars, should not be taken as definitive, since there are non-independence effects (such as that generated by repeatedly sampling a small population of speakers) which are not accounted for by the simple underlying model. Nonetheless, they may be taken as a rough guide to the degree to which the point estimates in the bars are uncertain based on the size of the sample subtending them. Note that in the present case, the confidence intervals in the crucial cases discussed above are entirely non-overlapping.

[8] A reviewer points out that priming effects might arise because of heterogeneities in the levels of different negation strategies on a text-by-text basis. This is indeed a concern, though it cannot explain our most important result, which is a difference in the fulfilment of the priming predictions of the two different analyses considered.

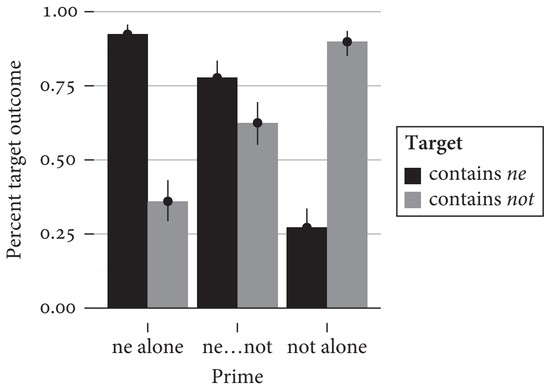

FIGURE 13.5. The predictions of the two-atom model measured in the corpus data.
Note: The error bars give 95% confidence intervals according to the binomial distribution.

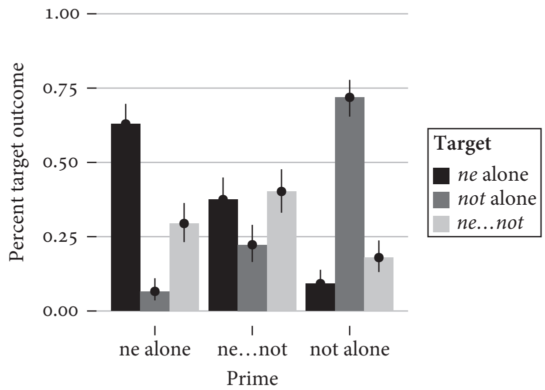

FIGURE 13.6. The predictions of the three-atom model measured in the corpus data.
Note: The error bars give 95% confidence intervals according to the binomial distribution.

13.4.3 Three-atom model

If the three-atom model is correct, then we predict that each kind of negation should facilitate itself, and not any of the other forms. This prediction too is only partially borne out. In Figure 13.6, the right-hand case, where the prime is *not*, exhibits the expected behaviour, whereby it strongly primes a following *not* but the other two types of negation have low and (most importantly) equal rates. On the other hand, in the other two cases, there is some degree of cross-priming between *ne* and *ne . . . not*, in the sense that, in the *ne* or *ne . . . not* prime cases, the bar for the opposite target from this set is higher than the bar for *not* alone.

However, a degree of cross-priming can be explained by appealing to the notion that some of the surface *ne . . . not* tokens are in fact tokens where *ne* is the negator, and *not*

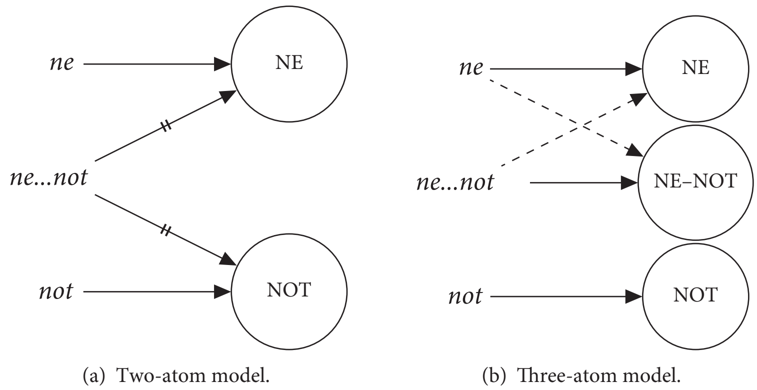

FIGURE 13.7. Priming predictions and actual results of the two models.

Note: Each diagram shows the predicted priming relationships between surface forms (left) and grammatical objects (right). Solid lines indicate priming relationships which are both predicted and observed. Dashed lines indicate relationships which are not predicted, but are nonetheless observed in the corpus. Finally, solid lines with cut marks indicate relationships which are predicted but not observed.

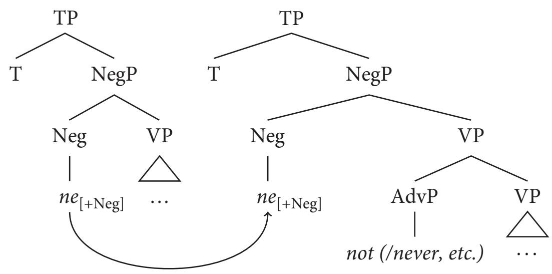

FIGURE 13.8. Structure with *ne* as sentence negation in combination with *not* or another reinforcing adverb.

Note: The arrow indicates a priming relationship between two tokens.

is a sentence adverb. Such a structure is illustrated in Figure 13.8. In these cases, the negator *ne* facilitates itself and emphatic *not* is additionally either added or subtracted. It is possible to test this fix, using the divide-by-0.16 method from Frisch (1997) to calculate the rate of *ne . . . not* tokens which contain adverbial *not*. For *ne . . . not* targets, the test is exact: we discount the number of observed *ne . . . not* tokens by 16%. For *ne . . . not* primes, we cannot assume that the distribution of adverbial *not* is consistent across target categories. However, we can set an upper bound on the discount by assuming that all adverbial *not* primes (i.e. negator *ne* primes) precede *ne* by itself. The (unknowable) true amount which must be corrected is less than or equal to the output of such a calculation.

The results of applying this analysis can be seen in Figure 13.9, which presents a picture much closer to the predicted situation. (The remaining discrepancy is the

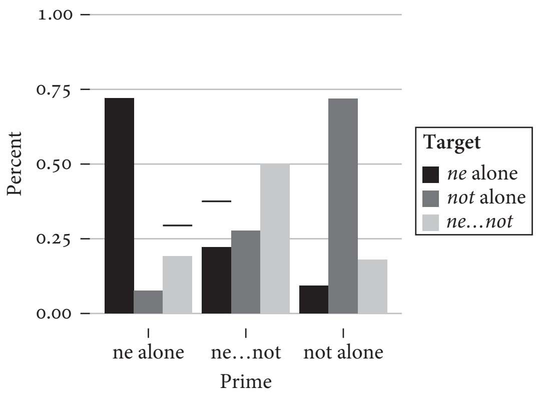

Figure 13.9. The three-way priming results, after applying Frisch's divide-by-0.16 adjustment for adverbial *not*.

Note: The black marks indicate the original height of the bars directly affected by the patch. (Because the bars in this graph are constrained to sum to 100%, the other bars' heights are affected as well.)

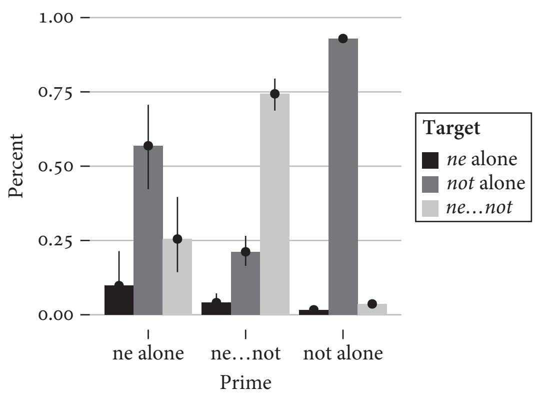

Figure 13.10. Priming effects on ME negation in the period 1350–1400.

Note: The error bars give 95% confidence intervals according to the binomial distribution. Note that invariant texts are excluded from the figure.

relatively small amount by which the height of the leftmost light grey bar exceeds the leftmost medium grey one, when they are predicted to be equal in height.)[9]

[9] The curious reader may wonder whether this same adjustment will improve the picture for the two-atom model depicted in Figure 13.5. It is in fact of equivocal benefit: in the crucial *ne … not* prime condition it raises the *not* + *ne … not* bar while lowering the *ne* + *ne … not* bar. Thus, the existence of this patch does not increase our confidence in the two-atom model. For reasons of clarity, we do not present the result here, although the code is available in the associated GitHub repository in the function `two.way.patch.graph` in the file `scripts/digs-proceedings-graphs.R`.

Another piece of evidence in favour of the three-atom model comes from the later period of the change (1350–1400; $N = 1617$), and is depicted in Figure 13.10. In this period, we see that *ne* facilitates *not* more strongly than *ne . . . not* does. That is to say, the leftmost medium grey bar is taller than the middle one. The two-atom model assumes that there is a relationship of structural identity between *ne . . . not* primes and *not* targets, of the sort that should give rise to priming. There is no such relationship between *ne* primes and *not* targets. Thus, the two atom model predicts that priming should raise the height of the middle bar above that of the left-hand one, contrary to fact. The three-atom model on the other hand has no problem with this data; it predicts (accurately) that the bar for targets identical to the prime should be higher than the same-colour bar for the other two possible primes.

13.5 Conclusion

In this chapter, we have presented corpus priming data which are inconsistent with the two-atom model and provide tenuous support for the three-atom model proposed by Wallage (2008). We note that this evidence by itself does not dispose of the question of which of the two models best describes the historical data. The total picture must also include the CRH-based quantitative evidence discussed by Frisch (1997) and Wallage (2008).

Broadening our view beyond the ME negation case study, we observe that the CRH is important because it provides a link between frequency data attested in historical corpora and the mental representations that underlie language and language change. We have here argued that persistence data constitute another, independent source of linkage between these two domains. The investigation of persistence in historical corpora (used in combination with other techniques) can contribute to a body of convergent evidence about syntactic history, and thus to conclusions more robust than those confined to any single methodological framework. Furthermore, the existence in historical corpora of patterns that can be interpreted as priming, in line with studies that uncover priming in purely synchronic textual corpora, lends confidence that the subject of historical corpus investigation has an underlying psychological reality.

Acknowledgements

We would like to thank the following people for their assistance: The compilers of the PPCME2, Beatrice Santorini, Anthony Kroch, Phillip Wallage, our fellow graduate students at Penn, the audiences of DiGS15 and PLC38, and an anonymous reviewer of this manuscript. All remaining imperfections are attributable only to the authors. Further, we point out that the data and code used in this analysis, including the LaTeX source code for this article, is available on GitHub: https://github.com/aecay/digs15-negative-priming