



## Article

Ulrike Schneider\*

# *I couldn't help but wonder: do modals and negation attract?*

<https://doi.org/10.1515/cllt-2023-0029>

Received March 17, 2023; accepted October 13, 2023; published online November 16, 2023

**Abstract:** The present paper focusses on the historical development of the relationship between the English core modals *can*, *could*, *shall*, *should*, *will*, *would*, *may*, *might* and *must* and the negator *not*. It explores whether semantic and morpho-syntactic factors, particularly the emergence of DO-support in Early Modern English, the increase in the popularity of contracted forms such as *won't* in the nineteenth century and the loss of core modals in the twentieth century, had an influence on negation rates. Large-scale empirical analyses of modal use in historical corpora of British prose fiction published between ca. 1500 and 1990 reveal that many modals—particularly high-frequency *will*, *would*, *can* and *could*—indeed attract *not*. The establishment of the contractions *n't*, *'ll* and *'d* had the strongest effect on the modal-negation system after 1500. The availability of the contracted modals *'ll* and *'d* led to a functional split whereby *will* and *would* became much more strongly associated with negation while contracted *'ll* and *'d* repel *not*-negation.

**Keywords:** modals; negation; diachronic change; contractions; Early Modern English; Late Modern English

## 1 Negation and modals

In previous analyses of diachronic change in verbal constructions (cf. Mondorf and Schneider 2016; Schneider 2021a, 2021b, 2023), a notable attraction surfaced repeatedly: If a verb is accompanied by a modal, the chance that it is also negated is often significantly higher than in non-modal uses of the same verb. The correlation appears in American and British data. It is found in Early Modern English (EModE) and Late Modern English (LModE) corpora and irrespective of whether a verb generally attracts modals or not.

---

\*Corresponding author: Ulrike Schneider, Department of English and Linguistics, University of Mainz, Philosophicum, Jakob-Welder-Weg 18, 55128 Mainz, Germany, E-mail: [ulrike.schneider@uni-mainz.de](mailto:ulrike.schneider@uni-mainz.de)

While the relationship between modals and negation has attracted some attention in past decades, questions have often centred on scope (cf. e.g. De Haan 1997; Palmer 1997; Radden 2009) and on how differences in scope can be best modelled within generative frameworks (cf. also Roberts 1985; Beukema and van der Wurff 2002; Hankamer 2011; Puskás 2018). Empirical and diachronic investigations, on the other hand, have mostly focussed on the spread and functions of contracted *n't* when combined with (specific) modals (cf. e.g. Tagliamonte and Smith 2002; Bergs 2008; Bybee 2010: 151–164; Dausgs 2021; Hejná and Walkden 2022: 77–79; López-Couso and Pérez-Guerra 2023; Nakamura 2023; Varela and Ramón 2013). Thus we still know little about the general co-occurrence patterns of modals and negation in English.

The relationship between the two factors is particularly interesting when analysed in the context of Hopper and Thompson's (1980) theory of Transitivity, which distinguishes binary grammatical transitivity from gradable semantic Transitivity. While the former hinges on the presence or absence of a direct object, the latter is defined as the "effectiveness or intensity with which the action is transferred from one participant to another" (Hopper and Thompson 1980: 252). In both (1)b. and (1)c. the effectiveness is reduced compared to non-modal and non-negated (1)a. because in (1)b. and (1)c. no action has taken place and the band thus remains unaffected. Therefore, modals and negation are considered Transitivity-lowering features.

- (1) a. John left the band.  
 b. John could have left the band. MODAL  
 c. John did not leave the band. NEGATED

Hopper and Thompson (1980: 254) argue that "the component features of Transitivity co-vary extensively and systematically". To this "Transitivity Hypothesis" they add the disclaimers that a) it does not predict *when* features will cluster, only that, when a cluster emerges, the clustered properties "will agree in being either both high or both low in value" (Hopper and Thompson 1980: 254) and b) that it pertains "to obligatory morphosyntactic markings or semantic interpretations" only (Hopper and Thompson 1980: 256). Previous studies have both criticised the vagueness of b) and produced counter-examples to a) (see, for instance, Tsunoda 1985: 392–394, Lazard 2002: 175 and Malchukov 2006: 333). Hopper and Thompson (1980: 279) themselves seem to weaken b) when they argue that Transitivity may offer "explanatory principles in [...] documented or presumed types of change", which are gradual in nature and do not abruptly bring about obligatoriness.

The present study explores whether Transitivity offers an explanatory principle for co-occurrence patterns of the English core modals (*can, could, may, might, must, shall, should, will, would*) and the negator *not/n't* (jointly referred to as NOT). Based on large-scale diachronic analyses of collections of British prose fiction from the EModE and LModE periods, it additionally address the following questions:

1. Do core modals attract NOT-negation?
2. Has the relation between the core modals and NOT changed over the past 500 years? If so, which factors influenced this change?
3. Did the emergence of contracted *n't* have an influence on the modal–negation relationship?

The remainder of this paper is structured as follows. Section 2 discusses what we currently know about the combined use of modals and negation in EModE and LModE as well as about historical changes which may have influenced their co-occurrence. Section 3 introduces the Chadwyck-Healey corpora of historical British fiction as well as the imaginative prose section of the British National Corpus (BNC), from which the data for the present analysis was drawn and details how databases of modals and NOT-negation were compiled and annotated. Section 4 first provides baseline rates of NOT-negation and of modals before comparing these with the co-occurrence rates of modals and NOT. This is followed by a statistical evaluation of the strength and significance of the attraction between different modals and NOT by means of collostructional analysis (cf. Stefanowitsch and Gries 2003; Gries 2023) and finally by a comparative analysis of the behaviour of full-form *not* versus contracted *n't*. The results are discussed in Section 5.

## 2 Background

Before we delve deeper into the matter, a brief note on the state of the modal and negation systems in EModE is in order. By late Middle English (ME), a system of auxiliary verbs had developed and, partially aided by the gradual loss of the subjunctive, “the modals commonly appeared as ‘semantic substitutes’ for verbal inflection” (Roberts 1985: 42; cf. also Beukema and van der Wurff 2002). This means that by the beginning of the EModE period, the core modals were well-established (although some of the lexical verbs from which the modals derived were still in use) and had largely also taken on their modern form except for spelling variants and inflection for the second person singular (cf. Denison 1993: 297–298).

The same can be said for the negator *not*. The dominant form of sentence negation in Old English was *ne* + finite verb, as in (2). Starting in Old English, the adverb *not* as well as related forms, like *nowiht*, *naht* or *nought*, were increasingly used together with *ne* to reinforce negation, such as in (3). This happened so often that “[b]y the end of the ME period the use of *not* was virtually obligatory if there was no other negative element than *ne*” (Denison 1993: 450; see also 470 en.3). Around the same time, speakers began omitting *ne* so that *not* remained as the sole negator in the

clause (cf. also Denison 1993: 450; Laing 2002: 298). As a result, *not*-negation had become the norm by the EModE period.

- (2)    *ic ne secge*  
       *I not say*  
       STAGE 1 IN THE “NEGATIVE CYCLE” (Jespersen 1917: 9–14)
- (3)    *I ne seye not*  
       *I not say not*  
       STAGE 2 IN THE “NEGATIVE CYCLE”
- (4)    *I say not*  
       STAGE 3 IN THE “NEGATIVE CYCLE”

Nevertheless, two major changes affected negated VPs in the EModE and LModE periods, namely the establishment of obligatory operators and the emergence of contracted forms. These will be addressed below together with the loss of core modals, which so far has not been linked to negation, but which may have had an influence on modal–negation rates. Yet, I will first touch upon the (combined) semantics of modals and negation.

## 2.1 Scope and suppletion

In the simplest terms, what is described in a negated or modal clause is marked as neither having happened nor happening at the time of speaking, i.e. both independently shift a proposition to irrealis or false (cf. e.g. Givón 1984: 321). This is why Hopper and Thompson (1980) classify both as Transitivity lowering—the action has not affected the patient. There may be a possibility for it to do so in the future—even an obligation—but the actual event happening is not part of the proposition of the clause.

Semantics get more complex where the two are combined. When a modal and a negator co-occur, the negator can either have narrow scope, which only extends over the proposition (in which case the modal scopes over the negator), as illustrated by (5) or it can have wide scope extending over the modal, as in (6) (cf. Klima 1964; De Haan 1997: 12; Palmer 1997: 137; Quirk et al. 1985: 794–796; Radden 2009: 170).

- (5)    John [must **not** leave the band].  
       ‘What John is obliged to do, is not leave the band.’  
       \*‘What John is not obliged to do, is leave the band.’  
       Narrow scope of negation: Negation of the proposition

- (6) John [[can]n't] leave the band.  
 \*‘What John is permitted to do, is not leave the band.’  
 ‘What John is not permitted to do is leave the band.’  
 Wide scope of negation: Negation of the modality

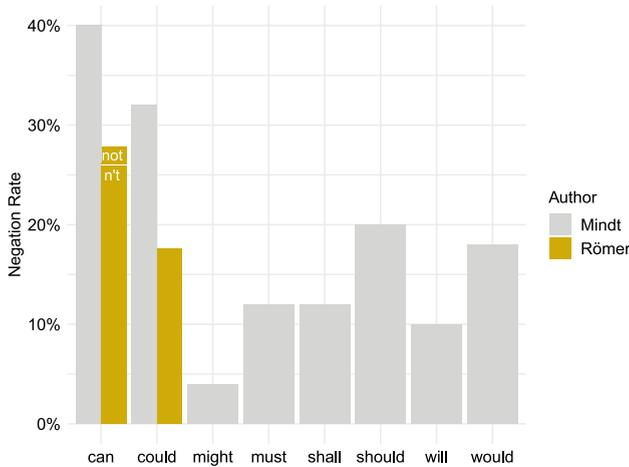
While many other languages use word-order to distinguish between these (cf. e.g. De Haan 1997: 86–109), English does not.<sup>1</sup> Instead, which reading(s) is/are permitted, depends on the modal and to some degree on modal flavour (i.e. deontic vs. epistemic modality). Thus negation is most likely to have narrow scope when combined with *should*, *might*, *must* and epistemic *may*, while it tends to have wide scope with *can*, *could* and deontic *may* (cf. Coates 1983; De Haan 1997: 60–67; Francis and Iatridou 2020: 285; Palmer 1997: 136; Quirk et al. 1985: 794–797; Radden 2009: 174). Quirk et al. (1985: 795–797) argue that the distinction between narrow and wide scope “is neutralised” in the case of *will* and *would* as both interpretations are semantically equivalent. Coates (1983: 64) assigns *should* to this ‘scope neutralised’ category, too. De Haan (1997: 60–67) lists alternative—and presumably much rarer—scope interpretations which are possible with some modals. He claims that deontic *can not* actually permits both narrow and wide scope interpretations of negation while *cannot* as well as contracted *can't* only permit a wide-scope reading of negation (De Haan 1997: 62).

This may prompt the question whether negation rate and scope are linked; i.e. whether modal–negation pairs with more flexible interpretations occur more frequently than ones with restricted scope interpretations. However, scope is unlikely to be a good predictor of negation rate, not least because there are additional complicating factors; namely that not all semantic contexts are equally frequent and that speakers often adopt suppletion strategies; i.e. “the difference between wide scope interpretation of the negation and narrow scope interpretation is expressed by means of different modal elements” (De Haan 1997: 58). If speakers, for instance, wish to express ‘not necessary’ they cannot use a necessity modal, such as *must*, due to scope restrictions. Instead, they opt for a possibility modal like *may* (cf. Palmer 1997: 135; see also Radden 2009: 170). Suppletion may explain some of the differences in use of between negated and non-negated modals (cf. e.g. Bybee 2010: 151–164; Leclercq 2022).

Negation rates determined by Mindt (1995; grey bars in Figure 1; based on a corpus consisting primarily of British fiction) suggest, on the one hand, that *would* and *should* with their potentially neutralised scope interpretations have higher negation rates than scope-restricted *might* and *must*. On the other hand, negation rates of *will* and *shall* are not elevated despite neutralised scope and *could*, with

---

<sup>1</sup> See Beukema and van der Wurff (2002: 88–90) for a discussion of examples, which show that even in ME, when English permitted both OV and VO word-orders, there was no one-to-one relation between scope and word-order.



**Figure 1:** Negation rates of different modals as determined by Mindt (1995: 176) and Römer (2004: 189).

which negation is restricted to narrow scope, has some of the highest negation rates. Moreover, more nuanced data provided by Römer (2004; yellow bars in Figure 1; based on the spoken section of the British National Corpus) shows that the frequency of *can't* far exceeds that of *cannot*, despite the fact that *cannot* is possibly less restricted in scope than the contracted form.

Finally, language change also affects scope—some older sources still list combinations of scope and modal flavour which are missing from later lists and, of course, there are differences between British and American English (cf. Tottie 1985). Finally, core modals are not only replaced by other core modals, but also by semi-modals as well as lexical expressions of modality etc. (cf. also Leech 2013). All of these factors make it unlikely that flexibility of scope has an influence on negation rates of modals.

## 2.2 do-support

Besides semantics, there are a range of syntactic changes which could potentially have influenced modal-negation rates. The first of these is the emergence of *do*-support, or operator *DO*, between 1500 and 1700. Within this timeframe, English shifted from mostly 'finite verb + *not*', as in (7), to predominantly '(finite) operator + *not* (+ lexical verb)' in negative declarative sentences, as in (8) (cf. Denison 1993: 451; Ellegård 1953; Jespersen 1917: 10–11; Strang 1970: 151; Visser 1969: §1440–1441).

(7) I say not  
STAGE 3 IN THE “NEGATIVE CYCLE” (Jespersen 1917: 9–14)

(8) I do not say.  
STAGE 4 IN THE “NEGATIVE CYCLE”

Ellegård (1953) shows that the rate of *do*-support in negative declarative sentences (in which no other operator is present) rises from 5 % in 1500 to 75 % in 1700. In negative questions, *do* spreads even faster. When no other auxiliary is present, these are almost invariably formed with *do* by 1700 (cf. Ellegård 1953).<sup>2</sup> Warner (1993: 221–222) speculates that the parallel development of the modals and operator *do* “strongly suggests that [the two] are interconnected”.

In fact, we have evidence that modals functioning as operators paved the way for operator *do*. Denison (1993: 467–468) analyses six Paston Letters, written in the late fifteenth and early sixteenth centuries, “when periphrastic *do* was in infrequent use and had by no means reached its peak, let alone been regulated” (Denison 1993: 467). He finds that the chance of any operator being present is significantly higher in negated than in affirmed sentences and concludes that this “suggests that the essence of the process now labelled as the ‘regulation of *do*’—namely the tendency for negatives [...] to contain an operator—had already begun long before *do* itself was a statistically important element in the language, possibly even before periphrastic *do* first arose” (Denison 1993: 467–468). Budts and Petré (2020) corroborate this in an analysis of Early English Books Online (EEBO), covering the timespan 1580–1700. They confirm that “periphrastic *do* became distributionally more and more similar to the modal auxiliaries (and, possibly, also the reverse)” (Budts and Petré 2020: 346).

Warner (1993: 209) suggests that this reverse influence consisted in an early effect on the emergence of the modals as a class of auxiliaries. Budts (2022) provides tentative suggestions as to what a later influence of *do* on the modals could have looked like. Her neural network analysis of late sixteenth-century uses of *do* as well as modals uncovers functional overlap between them. Particularly *will* shows “extensive overlap with *do*” (Budts 2022: 357). Her results further suggest that negated *do* and *can* in combination with perception and cognition verbs are functionally interchangeable, as “the inability (*can not*) to perceive something inevitably leads to the perception not taking place (*do not*)” (Budts 2022: 355). She concludes that, under the influence of other modals, *do* briefly expressed epistemic modality in sixteenth century affirmative declaratives and further that

---

<sup>2</sup> With some high-frequency verbs, like *say* and *know* (cf. Ellegård 1953: 198–200; Denison 1993: 451, 459), as well as with imperatives, non-operator negation survived longer (cf. Strang 1970: 152).

the loss of affirmative *DO* seems to suggest that the story of periphrastic *DO* and the modals in the 17th century is one of divergence rather than attraction. Perhaps affirmative *DO* had grown so similar to the modals, especially to *WILL*, that the two forms entered in a competition that pushed *DO* to the fringes of the paradigm, allowing it to preserve its auxiliary syntax but forcing it to let go of its modal semantics. (Budts 2022: 359)

The present study allows us to assess whether there are changes in the rate of modal-negation after 1700 which point to a divergence from or further convergence with *DO*.

## 2.3 Decline in core modals

The second syntactic reorganisation process which may have triggered changes in modal-negation rates is the decline in core modals and the concomitant rise of semi-modals such as *need to* and *want to* over the course of the twentieth century (cf. Leech 2003, 2013; Leech et al. 2009). As already pointed out in Section 2.1 above, speakers today have the option to use semi-modals in contexts where core modals would have dominated earlier. Leech (2003) shows that this leads to a shift in LModE. He analyses the frequency of modals in the Brown family of corpora and finds significant decreases in the frequency of core modals. Rates of decline range from 2.7 % for *will* to 43.7 % for *shall*. Yet, in British English *can* and *could* do not seem to be affected by this shift; they are actually on the rise.

It is however possible that this process has not affected negated and affirmative contexts in the same way. If affirmative uses of modals are lost at a different rate than negated uses, this would have an effect on negation rates. The present analysis investigates whether we see any evidence of this happening.

## 2.4 Negative contraction

The final diachronic change to consider is the emergence of negative contraction. In the sixteenth century, some modal-negation combinations take on a different form: they are contracted and the vowel in *not* is elided (cf. e.g. Jespersen 1917: 116; Denison 1993: 309).<sup>3</sup> For about 300 years, the contractions are much less frequent than the full forms.

---

<sup>3</sup> There were earlier contractions of the negator *ne* and modals, such as *nill* (< *ne + will*), but they did not survive into EModE. Additionally, Denison (1993: 309) lists modal-negation contractions occurring since the fifteenth century in which some of the *modal* is elided or assimilated. Except for *shan't* and *won't*, which caught on, and actually combine elision/assimilation on the part of the modal with elision of the vowel in the negator, this seems to have been a relatively rare and short-lived phenomenon.

Daug (2021: 26) searches the fiction subcorpus of the Corpus of Historical American English (COHA) for *can/will* + negation and finds that both undergo the same shift: Early in the nineteenth century, contracted forms make up around 20 % of the data; this rate climbs steadily until it reaches around 80 % by the beginning of the twentieth century and then plateaus there. In more formal genres, the shift occurs later. In *TIME* magazine, for instance, contractions of *can*, *could*, *will*, *would* or *should* + *n't* only surpass the full forms by the 1990s (cf. Millar 2009: 211; see also Daug 2021: 44).

Nakamura (2023) provides evidence which suggests that the spread of the contraction may have been phonologically governed, with the contraction occurring first in forms where the resulting final consonant cluster consists exclusively of /nt/ (i.e. *don't*, *can't*, *won't* and *shan't*) and spreading later to forms where it results in longer consonant clusters (i.e. *doesn't*, *didn't*, *couldn't*, *wouldn't*, *shouldn't*, *needn't* and *mustn't*; see also data by Hejná and Walkden 2022: 77–79 which supports this claim). *Mayn't* seems to constitute an exception as with *may* the contraction never really caught on (cf. Zwicky and Pullum 1983: 507; Nakamura 2023).

There is further evidence that variation between contracted *n't* and full form *not* is systematic: Firstly, contracted *n't* is associated with informal spoken language and fiction (cf. e.g. Bergs 2008: 122; Biber et al. 1999: 1129–1132; Szmrecsanyi 2003: 302, 304). Secondly, speakers are much less likely to use *won't* in syntactically dependent contexts than in independent ones—an indication that speakers may opt for the full form in syntactically more complex environments, in line with Rohdenburg's (1996) complexity principle (cf. also Szmrecsanyi 2003: 309). Thirdly, the contracted forms *can't* and *won't* attract monosyllabic lexical verbs as well as verbs encoding mental processes (cf. Daug 2021: 32, 40). Finally, contractions are associated with different kinds of modality than the full forms (cf. Daug 2021: 40–41).

Due to its idiosyncrasies, Zwicky and Pullum (1983) classify *n't* as an inflectional affix. Daug (2021: 45, 47) interprets his findings as evidence that *can't* and *won't* may be separate constructions. Bergs (2008: 134) makes a similar argument for *shall not* versus *shan't*—he cautiously talks about new constructions developing “in some sense”. At the very least, this means that empirical analyses need to take into consideration that for several modals, there may be two kinds of NOT-negation, which are preferred in different contexts. Therefore, the last leg of the analysis investigates whether we see differences in the attraction between modals and contracted *n't* versus full-form *not*.

### 3 Data and coding

The present study is based on three of the Chadwyck-Healey corpora (Early English prose fiction 1997; Eighteenth-century fiction 1996; Nineteenth-century

**Table 1:** Historical corpora. Periods preceded by <sup>b</sup> are based on birth dates, the one preceded by <sup>p</sup> is based on publication dates.

Period	Corpus	Size	Total size
<sup>b</sup> 1460–1699	Early English Prose Fiction (EEPF)	10.0 million words	15.6 million words
	Eighteenth-Century Fiction 1 (ECF1)	5.6 million words	
<sup>b</sup> 1700–1799	Eighteenth-Century Fiction 2 (ECF2)	4.8 million words	16.6 million words
	Nineteenth-Century Fiction 1 (NCF1)	11.8 million words	
<sup>b</sup> 1800–1869	Nineteenth-Century Fiction 2 (NCF2)		27.8 million words
<sup>p</sup> 1960–1993	British National Corpus (wridom1)		19.4 million words

fiction 1999–2000) consisting of prose published in Great Britain between 1500 and 1899. Data for the twentieth century is supplied by the imaginative prose subdomain of the British National Corpus (1995; BNC wridom 1). Table 1 indicates the time-periods covered by the corpora as well as their size. While works were originally grouped by publication date, it has become customary to regroup the data by authors' birth dates as it can be assumed that a speaker's idiolect changes less over their lifetime than language itself changes in the same time (cf. e.g. Bailey et al. 1991).

The corpora were searched with Antconc (Anthony 2014) for all instances of *NOT*, including the variants *n't* and *'nt*.<sup>4</sup> After removal of false hits like *knot*, *snot*, *Huguenot*, *on't* ('on it') or *attain't* ('attained'), 799,604 tokens remained.

A second dataset was compiled by searching for all tokens of the core modals *can*, *could*, *shall*, *should*, *may*, *might*, *must*, *will*, *'ll* and *would* in the corpora. NCF1, NCF2 and wridom1 were also searched for the form *'d*.<sup>5</sup> In order not to miss any tokens due to spelling variation, the 500 most frequent types immediately preceding *NOT* in each of the six corpora were separately searched for spelling variants of the modals. I then compiled two separate search files per modal: one for the older corpora (EPPF–ECF2), in which variation is still extensive, and one for the newer corpora (NCF1–wridom 1) which contain fewer orthographic variants. These search files contained simple forms of the modals as well as variants with attached *NOT*. (9) shows the contents of the searchfile for *can* and *could* as an illustration of the large number of theoretically possible combinations.<sup>6</sup>

<sup>4</sup> Spelling variants of *not* with a macron representing <n> were also searched for, but not found in any of the corpora.

<sup>5</sup> In the older corpora, *'d* only occurs as a form of the past tense and past participle *-ed* suffixes.

<sup>6</sup> In order to keep false hits to a minimum, these forms were actually bracketed by boundary markers in the original searchfiles. This means that *canst* appeared as *\bcanst\b*, which is a way of telling Antconc to ignore this sequence of letters inside of other words, like *Americanstyle*.

- (9) canot, ca\_, ca\_not, ca\_ot, ca\_st, ca\_stnot, can, cannot, can't, cane, canne, cannot, canst, canstnot, canste, caud, con, connot, conn't, coud, coudnot, coudn't, coude, coudenot, couden't, coudst, couldstnot, could, couldnot, couldn't, couldde, couldnot, couldn't, couldes, couldesnot, couldest, couldestnot, couldst, couldstnot, coule, coulenot, coulen't, cowde, cowdenot, cowden't, culde, culdenot, culden't, cou'd, cou'dnot, cou'dn't, cou'd'st, cou'd'stnot, cou'd'stn't, cou'd'st, cou'dstnot, cou'dstn't, cannt

The resulting dataset of modals totals 1.4 million words. It was cleaned in a semi-manual process to weed out false hits. Non-standard spelling variants were checked to ensure that they were, in fact, variants of the modals in question and not other lexemes. Filters were used to inspect specific contexts, such as capitalised tokens, ones preceded by determiners or followed by punctuation. In this way, nouns like *can*, *cane* and *will* as well as proper names like *Will* or *May* were located and removed. Particular attention was paid to the 'd dataset as this contracted form can be a variant of *-ed*, *did* and *had* as well as of *would*. Tokens followed by a past participle (e.g. *he'd eaten*) were removed. Where the infinitive and past participle of the lexical verb were indistinguishable (e.g. *put*, *set*, *come*, *become*, *read*) the context was consulted to determine whether the form represented *would*. Tokens without a following lexical verb were discarded as this made it impossible to determine whether 'd represented *had* or *would*.

Despite the effort put into the clean-up, a dataset of this dimension will invariably contain some degree of noise as it is impossible to inspect each token manually. This also applies to the coding.

The following factors were semi-automatically coded for with the help of R (R Development Core Team 2009), due to the fact that the corpora are not POS-tagged.<sup>7</sup>

MODAL VERB—This factor provides the modal verb in standard modern spelling. The contracted forms 'd and 'll are treated as separate modals; *won't* is treated as *will + n't*.

NOT NEGATION—Modals co-occurring with *not* were coded as negated. This includes modal verbs followed by combinations such as *not only* or *not but*, despite the fact that, semantically, these cases are not negated. The tagger recognised tokens with an intervening adverb such as *rather* or *probably* as negated. In inverted clauses, i.e. where a single-word subject intervened between the modal and *not*, such as in (10), the modal VP was also recognised as negated.<sup>8</sup> Note that tokens like

<sup>7</sup> In a trial run even tagging with extensive retraining of a standard tagger provided unsatisfactory results (cf. Scherl 2019).

<sup>8</sup> Such single-word subjects were subject pronouns, including *thou* and *ye(e)* as well as *one*, *someone*, *anyone*, *something*, *this* and *that*. Retrieving inverted clauses from among the 23,000 tokens where the modal was followed by *you* and those where it was followed by a multi-word NP would not have been feasible.

*whether I could or not* were not coded as negated, as here the negator forms part of a coordinated elliptical VP.

(10) How could one not love him?

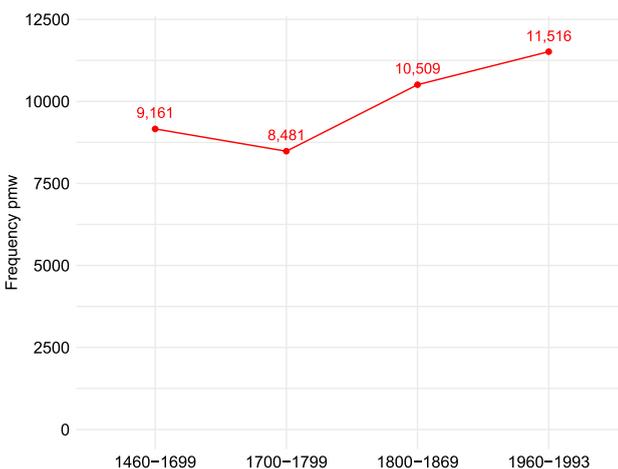
The same procedure was applied to tokens co-occurring with *n't* or *'nt*, except that these were coded *n't*-negated. Wherever a distinction between full-form *not* and contracted *n't* is not necessary, the two will be subsumed under the lemma NOT.

## 4 Analysis and results

### 4.1 Baseline frequencies

In order to determine whether modals and NOT-negation attract in English, we first need to know their frequency in each of the four periods under investigation. This first leg of the analysis provides baseline rates against which we can later compare the combined use of modals and NOT.

Figure 2 shows the relative frequency of NOT per million words, which seems to be on the rise. However, normalisation per million words may not be the most adequate solution as this does not account for changes in writing styles. Eighteenth and nineteenth century novels are known to be rich in nominal structures and poor in verb phrases, as exemplified by (11).



**Figure 2:** Relative frequency of NOT in British fiction.

- (11) A labyrinth of grandeur, less the property of an old family of human beings and their ghostly likenesses, than of an old family of echoings and thunderings which **start** out of their hundred graves at every sound, and **go** resounding through the building. (NCF1: Charles Dickens *Bleak House* 1853)

If we assume that most NOT-negation is clausal negation (i.e. scopes over the entire clause) and occurs in the verb phrase and, furthermore, that even constituent negation, which only scopes over a single constituent, such as in (12) and (13), rarely occurs more than once per clause, normalising the data per clause or verb phrase (VP) appears to be a more suitable option for diachronic comparisons.

- (12) [...] he Taught me **not only with Application**, but with admirable Judgement in the Teaching part [...] (ECF1: Daniel Defoe *Colonel Jack* 1723)
- (13) In the darkness they seemed **not the amusing and loveable clowns they were meant to be** but somehow menacing. (BNC: wridom1)

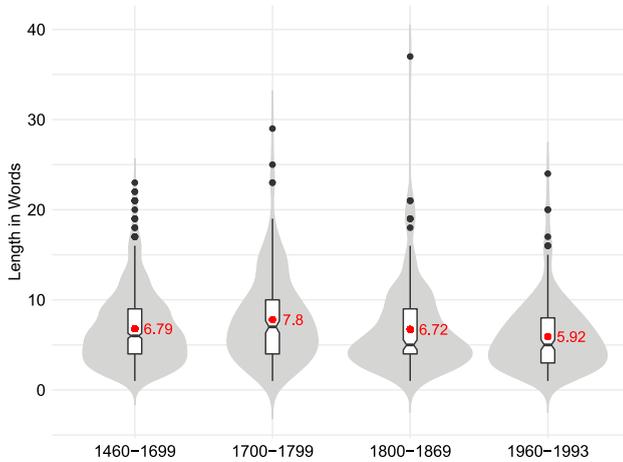
As the corpora are not parsed, the number of VPs had to be estimated. To do so, 200 tokens of lexical verbs (including copula) were randomly retrieved from each corpus. The context to the left of these sample verbs was cut off and the context to their right was cut off before the next lexical verb. Thus only the lexical verb and all words up to the next lexical verb in the text remained.<sup>9</sup> The result were four samples representing distances between lexical verbs.

Figure 3 visualises the distances between lexical verbs in the samples. Both the grey beans and the white boxes at their core provide information about the distribution of the data. The horizontal black line at the cinched waistline of the boxes indicates the median: 50 % of verbs were further apart and 50 % were closer together. Between the median and each end of the box lay 25 % of the data. Thus the box shows the inner 50 % of the data. In the case of the beans, girth indicates quantity: the wider the bean is in a particular area, the more data falls within that area. Overall, we see that the distance between lexical verbs is quite variable and that mean distances (indicated in red) have changed over time with lexical verbs in the eighteenth century on average being further apart than in other centuries.

Based on these samples, the number of VPs in each corpus could be extrapolated. This was done by doing 500 random draws of 100 words from each sample. The mean number of lexical verbs in these draws was used as the base for estimating the number of VPs. With the help of this estimate, the percentage of VPs which is negated was determined for each corpus; i.e. the frequency of NOT was

---

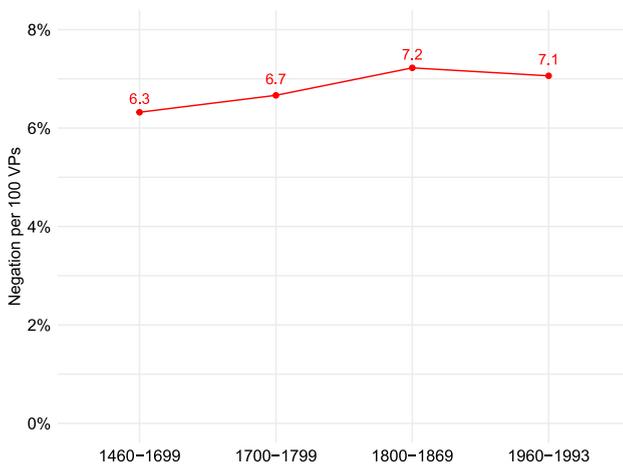
<sup>9</sup> My thanks go to my colleague Nicolás Rath for helping with this endeavour.



**Figure 3:** Bean plots indicating variation in the distance between lexical verbs in each period.

normalised per 100 VPs. This way of normalising the data rules out that increases in NOT are only due to clauses becoming shorter.

Figure 4 shows the result. Now, the slight dip in the negation rate in the eighteenth century, that was visible in Figure 2, has disappeared, indicating that it was due to eighteenth-century clauses being longer than those in the centuries before and after and authors therefore having fewer opportunities to use NOT. Figure 4 still suggests a slight upward trend in the negation rate until the nineteenth century, though.

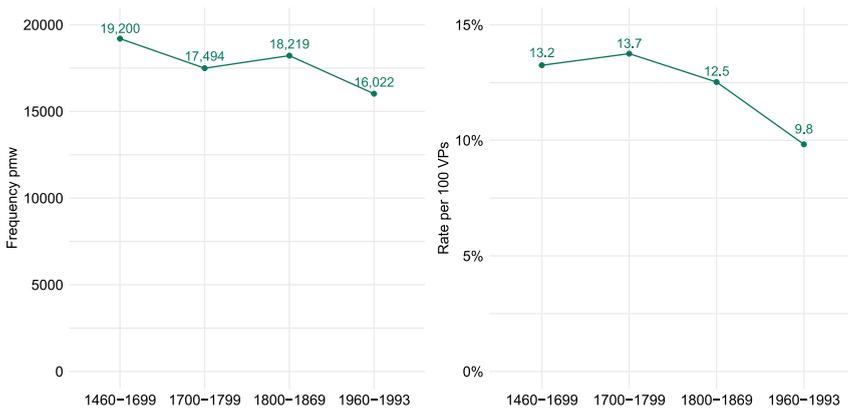


**Figure 4:** Frequency of NOT normalised per 100 verb phrases.

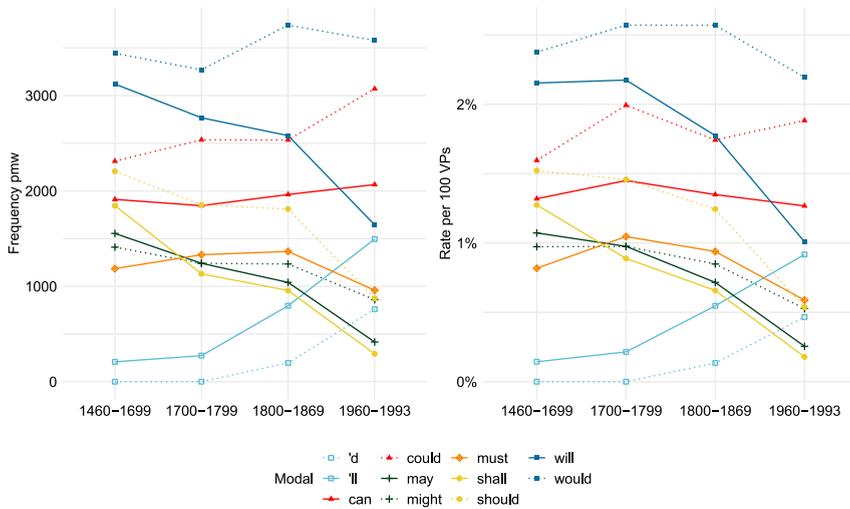
We can now turn to the modals. To render the data comparable to twentieth-century frequencies determined by Leech (2003, 2013) it is normalised per million words besides normalisation per 100 VPs. Figure 5 confirms the well-known decline in core modals (cf. Leech 2003, 2013) and provides evidence that it already began before the twentieth century. When normalised per million words (left panel), the data suggests a decline since EModE. When normalised per 100 VPs, however, the decline turns out to be a LModE phenomenon.

Yet the modals are unlikely to all exhibit a common behaviour and so need to be investigated separately. Figure 6 shows the development in the frequency of each modal. In the case of *shall*, *should* and *may*, usage frequency has been on the decline since EModE. *Might* seems to have begun its decline a century or two later. Not all core modals are on the decline, though. Frequencies of *can* are relatively stable and use of *could* even increased in the twentieth century. The most pronounced changes are those undergone by *will* and *would*. We witness their decline as well as the accompanying rise of the contracted forms 'll and 'd. In the case of 'll, its ascent comes at the expense of the full form of the modal (even with the two forms combined, by the twentieth century, the frequency of *will*/'ll is declining). By contrast, usage of full-form *would* remained unaffected by the rise of contracted 'd until the twentieth century and the combined frequency of 'd and *would* is stable (if normalised per 100 VPs) or even on the rise (if normalised per million words).

There are also some significant deviations from the frequencies reported by Leech (2003: 228). *Could*, *might* and *would* (if 'd is treated as a variant of *would* and not as a modal verb in its own right, which is how it was treated by Leech 2003: 226) are



**Figure 5:** Frequency of modal verbs normalised per million words (left) as well as per 100 verb phrases (right).



**Figure 6:** Frequency of core modals normalised per million words (left) as well as per 100 verb phrases (right).

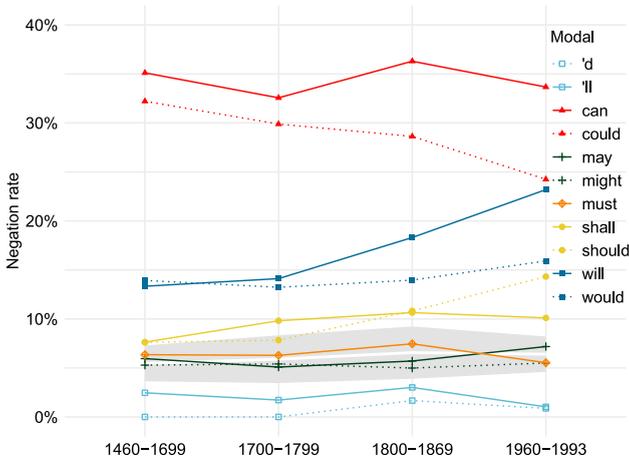
much more frequent here than in Leech’s data<sup>10</sup> while *may* occurs much less frequently in the novels than in the corpora used by Leech (2003) which comprise several genres (LOB/FLOB; see Leech 2003: 224).

## 4.2 Modal negation: frequencies

Let us now turn to the central question of whether modals and negation attract, which is addressed by Figure 7. At the bottom of the graph, just above the 5 % mark, is a white line that indicates the average negation rate for VPs which do not contain a core modal. This rate was estimated with the help of the two main datasets—the modal dataset and the NOT dataset—as well as the random draws providing information about the number of VPs in the corpora, described in Section 4.1. The shaded area around the mean is the standard deviation, determined with the help of Chang’s (2013) *Summary SE* function in R. The data indicates that mean NOT-negation rates have risen from 5.6 % in EModE to 6.3 % in the twentieth century.

Three of the curves fall on this band, indicating that the modals *may*, *might* and *must* have consistently had average negation rates and thus show no particular

<sup>10</sup> The combined frequency of *would/d* in the present data also exceeds the frequency determined by Biber et al. (1999: 488) for twentieth-century British fiction (c. 4000 tokens per million words vs. 4339 in the present data). The same is true for *will/ll* (c. 2600 per million words vs. 3140 in the present data).



**Figure 7:** NOT-negation rates of core modals.

attraction towards negation. The rest of the modals can be split into three further groups: Group one—the contracted modals *'ll* and *'d*—repels NOT-negation.<sup>11</sup> The second group is constituted by *shall*, *should*, *will* and *would*, which show elevated negation rates, often doubling or tripling the average rate. By the twentieth century, full-form *will* approximates the behaviour of the third and final group, where negation rates are more than four times that of the baseline. This group is made up of *can* and *could*.

Surprisingly, even the highest twentieth-century rates of NOT-negation lie consistently below those reported by Mindt (1995: 176; assuming that he treated *will* and *'ll* as one type). Yet they are higher than those found by Römer (2004: 189) in the spoken section of the BNC. Mindt (1995) does not report which part of his database was used to determine rates of modal negation. Therefore, I can only surmise that the differences are due to genre effects.

Figure 8 provides a different view of the data. Here, numbers are shown as frequencies per 100 VPs. It shows that changes in the negation rate can be a symptom of an increase or decrease in the usage frequency of affirmatives. The strong increases in the negation rate of *will*, for instance, are due to stronger losses in the use of affirmative *will* than of NOT-negated *will*.

<sup>11</sup> NOT-negation rates of *'ll* and *'d* are so low that they almost seem like positive polarity items. Yet, they actually attract *never* more strongly than other modals (cf. also Tagliamonte and Smith 2002: 268; Flach 2020: 752). In the present data, *never*-negation rates of twentieth-century *'d* and *'ll* are 3.3 % and 2.3 % respectively, while those of *would* and *will* are only 2.3 % and 1.1 % in the same period. Thus both contracted modals are significantly more strongly associated with *never* than their full-form counterparts (*will/'ll*:  $\chi^2 = 127.6$ ,  $p < 0.001$ ; *would/'d*:  $\chi^2 = 53.1$ ,  $p < 0.001$ )

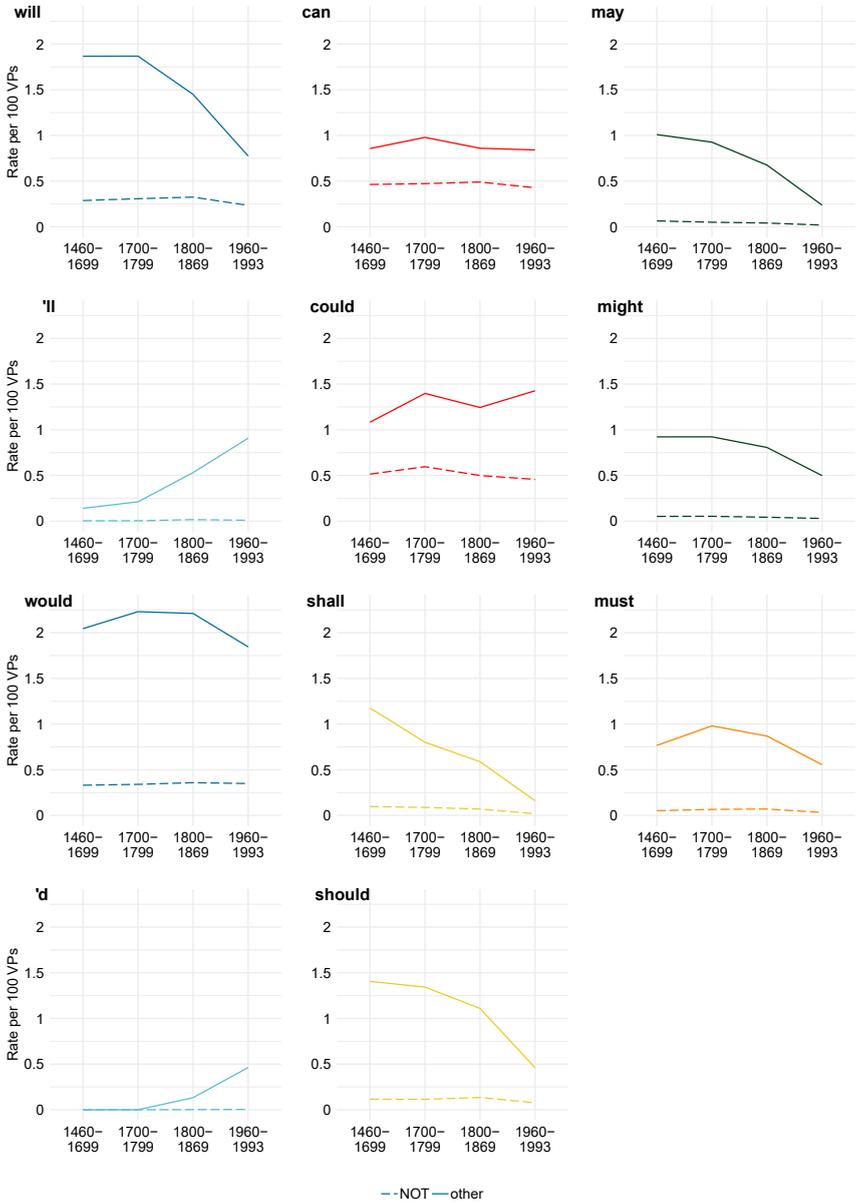


Figure 8: Modal negation normalised per 100 VPs.

What we have seen so far is that not all core modals seem to attract NOT—though predominantly the more frequent ones do. What remains is to statistically assess the strength of the attraction as well as to take the opposite perspective: does NOT attract modals? Both of these will be answered in the following section.

### 4.3 Modal negation: association strengths

Analyses of collocations (i.e. associations between two lexemes or word-forms) and collostructions (i.e. associations between a construction and a lexeme) are generally based on a contingency table, such as Table 2, which cross-tabulates the frequencies of the two elements (cf. e.g. Stefanowitsch and Gries 2003). Cell *a* is the one that is of particular interest to the researcher as it represents the condition where both elements are present, in our case, both the modal and the negator. *b*, *c* and *d* are used to calculate association measures which assess the strength of attraction or repulsion between two elements.

**Table 2:** A contingency table.

	word <sub>1</sub> is present	word <sub>1</sub> is absent	Row totals
word <sub>2</sub> is present	<i>a</i>	<i>b</i>	<i>a + b</i>
word <sub>2</sub> is absent	<i>c</i>	<i>d</i>	<i>c + d</i>
Column totals	<i>a + c</i>	<i>b + d</i>	<i>a + b + c + d</i>

Usually, word<sub>1</sub> is given more prominence in the analysis in that all relationships are described from the perspective of word<sub>1</sub>. In the present case, however, we are equally interested in the negator and the modal as both are Transitivity-lowering factors. This does neither mar nor complicate the analysis; it merely means that we will interpret some measures differently from the way they are commonly interpreted.<sup>12</sup>

**Table 3:** Schematic contingency table for modal negation.

	Modal present	Modal absent
NOT <b>present</b>	<i>a</i>	<i>b</i>
NOT <b>absent</b>	<i>c</i>	<i>d</i>

<sup>12</sup> Wherever possible, the analysis will follow the same steps as Schneider (2021a). For clarity, instead of Bates and MacWhinney's (1987) terminology used in Schneider (2021a), the terminology going back to Schmid (2000: 54–55) is used.

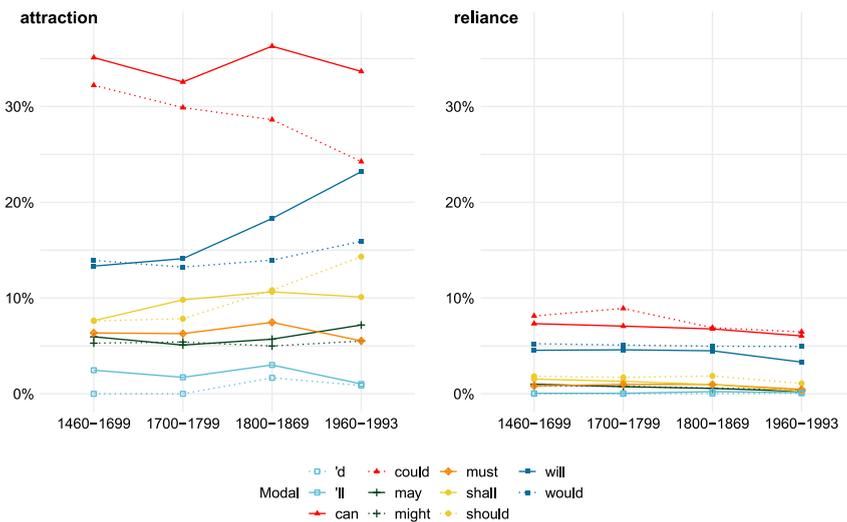
As a first step, we compute *attraction* (or forward transitional probability/cue availability, cf. Bates and MacWhinney 1987: 164). It is based on the formula (cf. e.g. Schmid and Küchenhoff 2013: 550; Levshina 2015: 228):

$$\text{attraction} = \frac{a}{a + c}$$

Once we fill the formula with values based on Table 3, we get a measure which indicates how strongly a core modal attracts *NOT*. Mathematically, attraction is the percentage of modal tokens which co-occur with *NOT*. It has already been discussed above; see Figure 7, which is repeated here as the left panel of Figure 9. Attraction is a directional measure of association as it quantifies how strongly the left element attracts the right one. The corresponding measure of attraction from right to left is often referred to as *reliance* (or backward transitional probability/cue reliability, cf. Bates and MacWhinney 1987: 164). It is calculated as follows (cf. e.g. Schmid and Küchenhoff 2013: 550; Levshina 2015: 228):

$$\text{reliance} = \frac{a}{a + b}$$

Usually, reliance is interpreted as a measure of how strongly word<sub>1</sub>, in our case the modal, relies on word<sub>2</sub>. It does that here, too; yet for the present purposes it is more useful to interpret it as a measure of how strongly *NOT* attracts the modal. Reliance is the percentage of *NOT* which is accompanied by a specific modal. We see it in the right panel of Figure 9.



**Figure 9:** Attraction and reliance.

Reliance is much lower than attraction—it was bound to be, due to the fact that we are dealing with a one-to-many relationship. This means that we are looking at one negator versus eleven core modals and each instance of NOT can only co-occur with a single modal. Therefore all reliance scores together cannot exceed 100 %. Coincidentally, the curve showing the combined reliance of all core modals on NOT would be almost identical to the curve displaying the attraction scores of *could* in the left plot—it has declined from 31.3 % in EModE to 23.4 % in the twentieth century.

Attraction and reliance are sometimes criticised for being hard to interpret, as they quantify associations but do not provide any information about significance (cf. Schmid and Küchenhoff 2013: 550). For significance testing, we have to compare the associations measured by attraction and reliance to distributions elsewhere in the language. This requires that cell *d* in the tables above be filled, which is not trivial (cf. e.g. Schmid and Küchenhoff 2013: 551–544). As modals and negators do not float through texts at random but occur only in syntactically restricted positions, the VP is once again a better unit than the word to determine corpus size.

We can now employ colostruational analysis as developed by Stefanowitsch and Gries (2003) to determine the mutual attraction between each modal and NOT. The standard metric for this type of analysis is *Fisher's Exact Test*. Yet, it may not be the most suitable measure in the present case. The test has been criticised for three things; firstly, for being difficult to compute—many applications return 0/infinity when the test is applied to large datasets;<sup>13</sup> secondly, for only providing information about significance and not of effect size and, thirdly, for being sensitive to sample size. The latter means that the same degree of deviation from expected values can result in a non-significant scoring when numbers are small while receiving significant scores when numbers are large (see discussion between Schmid and Küchenhoff [2013] and Gries [2015]).

To remedy these issues, Gries (2023) has recently proposed both a simplification and an expansion of colostruational analysis. In terms of simplification, he suggests replacing Fisher's Exact *p* with *chi<sup>2</sup> residuals*. The latter are obtained by performing a *chi<sup>2</sup>*-test on a complete table of all associations to be tested (Gries 2023: 358). R outputs the residuals with `chisq.test()$residuals`, but they can also easily be calculated:  $\text{observed} - \text{expected} / \sqrt{\text{expected}}$ . Gries (2023: 364–365) argues that this switch from one metric to another makes colostruational analysis faster and more accessible, while at the same time solving the 0/infinity issue. *Chi<sup>2</sup> residuals* and Fisher's Exact *p*

---

<sup>13</sup> Strong attraction in a large dataset results in a Fisher's Exact *p* so low that most implementations in R, such as `fisher.test()` or `pv.Fisher.collostr()` from Levshina's (2015: 232) package *Rling*, simply output a zero and when  $\log$  Fisher's Exact *p* is computed, which is often what is reported, this results in infinite values. Functions like `fisher.test.mpfr()` from Gries' (2021) script `coll.analysismpfr.r` solve this issue (cf. Gries 2015: 516–517).

are strongly correlated, which means that they provide users with the same information about associations (cf. Gries 2023: 359–360).

Gries’s (2023: 370–372) proposed expansion solves the second and third issues. He confirms that Fisher’s Exact  $p$  and  $\chi^2$  residuals “not only reflect frequency *and* association, but they also reflect frequency *more* than association” (Gries 2023: 372; emphasis in the original). He advocates that for some applications of collostructional analysis, frequency and association be kept separate (and that dispersion be added to the mix of dimensions taken into consideration). Thus, he proposes using *log Odds Ratio*, which is bidirectional and not sensitive to sample size. As an added perk, Odds Ratio is simple to compute (cf. Schmid and Küchenhoff 2013: 554):

$$\text{Odds ratio} = \frac{a/b}{c/d} \text{ or } \frac{a/c}{b/d}$$

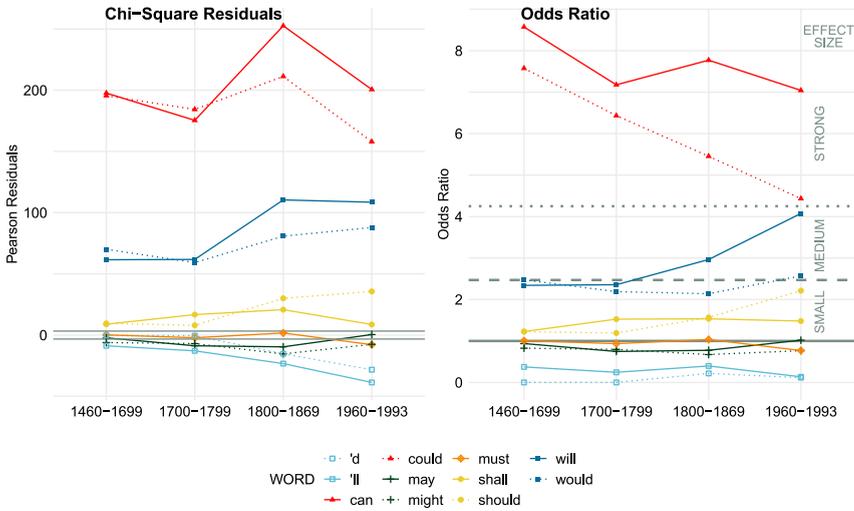
As its name implies, it compares two odds, which relate the probability of the observed to “the probability of what could also have happened given the full set of possibilities” (Schmid and Küchenhoff 2013: 553). Odds Ratio is a true effect size measure in the sense of Schmid and Küchenhoff (2013: 552–555).  $\chi^2$  residuals will therefore be combined with Odds Ratios.<sup>14</sup> I will provide more information about both measures as we discuss the results.

The left panel in Figure 10 shows the  $\chi^2$  (or Pearson) residuals. All curves above the  $x$ -axis belong to modals which co-occur more frequently with NOT than statistically expected, while all curves below the  $x$ -axis belong to modals which co-occur with NOT less frequently than statistically expected. Any residual higher than 3.25 or lower than  $-3.25$ , i.e. which falls outside of the narrow corridor around the  $x$ -axis which is indicated in the plot, corresponds to a significant result.<sup>15</sup>

The right panel in Figure 10 shows Odds Ratios. The following example illustrates how they can be interpreted: the Odds Ratio for *will not/won’t* in the twentieth-century data is four. This means that if we encounter a VP containing *will*, the odds that it also contains NOT are four times higher than the odds of a *will*-less VP to do so. And vice versa; if we encounter a NOT-negated VP, the odds that it also contains *will* are four times higher than the odds of a NOT-less VP to do so (cf. Schmid and Küchenhoff 2013: 554). An Odds Ratio of one indicates that there is no effect, while ratios below one indicate repulsion. The solid grey line added to the plot marks this dividing line between attraction and repulsion. The other lines indicate effect size. Anything below

<sup>14</sup> I use ‘plain’ Odds Ratios as proposed by Schmid and Küchenhoff (2013) instead of their logged cousin.

<sup>15</sup> This threshold already includes Bonferroni correction for the fact that we are performing 44 tests (one test per modal per period) on the same dataset (Gries 2009: 243, 2015: 520, fn.1): `sqrt(qchisq(c(0.05)/44, 1, lower.tail=F))`.



**Figure 10:** Association between modals and *not* assessed with the help of  $\chi^2$  residuals (also known as Pearson residuals) and Odds Ratio.

the dashed line is a small effect, points between the dashed and the dotted lines are medium effects and points above the dotted line represent large effects.<sup>16</sup>

The bidirectional tests largely confirm the impression we gleaned from the unidirectional tests above:

- Although declining, the attraction between *can/could* and *NOT* has been strong for the past 500 years (OddsRatio > 4.25,  $p < 0.001$ ).
- The attraction between *will/would* and *NOT* has been on the rise for the past 200 years. By the twentieth century, it reached medium strength (OddsRatio > 2.47,  $p < 0.001$ ).
- The attraction between *shall/should* and *NOT* has been consistently weak but still highly significant (OddsRatio < 2.47,  $p < 0.001$ ).

<sup>16</sup> These thresholds were determined by first converting Odds Ratio to Cohen's  $d$  with the help of the formula  $\ln(\text{OddsRatio})/1.81 = \text{Cohen's } d$ . Cohen states that effects with a  $d$  below 0.5 should be considered small, those with  $d$  between 0.5 and 0.8 medium and those with a  $d$  above 0.8 strong. We can determine which Odds Ratios correspond to these thresholds by converting the formula. As  $\ln$  is the natural logarithm, which has  $e$  (2.71828) as its base, the corresponding formulas are  $e^{0.5*1.81} = 2.47$  and  $e^{0.8*1.81} = 4.25$ .

The thresholds for repulsion effects have not been entered in the graph to avoid clutter. These are  $e^{-0.5*1.81} = 0.40$  and  $e^{-0.8*1.81} = 0.24$ .

- *Must* and *NOT* have shown neither attraction nor repulsion until the twentieth century, by which time they began to develop a (still very) weak repulsion (OddsRatio between 0.4 and 1,  $p < 0.001$ ).
- The relationship between *may/might* and *NOT* is one of weak repulsion (OddsRatio between 0.4 and 1), while the effect is consistently significant in the case of *might*, with *may*, it only reaches significance in the eighteenth and nineteenth centuries.
- The repulsion between the contracted modal *'ll* and *NOT* has increased. By the twentieth century, the effect is large (OddsRatio  $< 0.24$ ,  $p < 0.001$ ). The contracted modal *'d* and *NOT* have strongly repelled ever since the form emerged in the nineteenth century (OddsRatio  $< 0.24$ ,  $p < 0.001$ ).

#### 4.4 Contraction rates of *not*

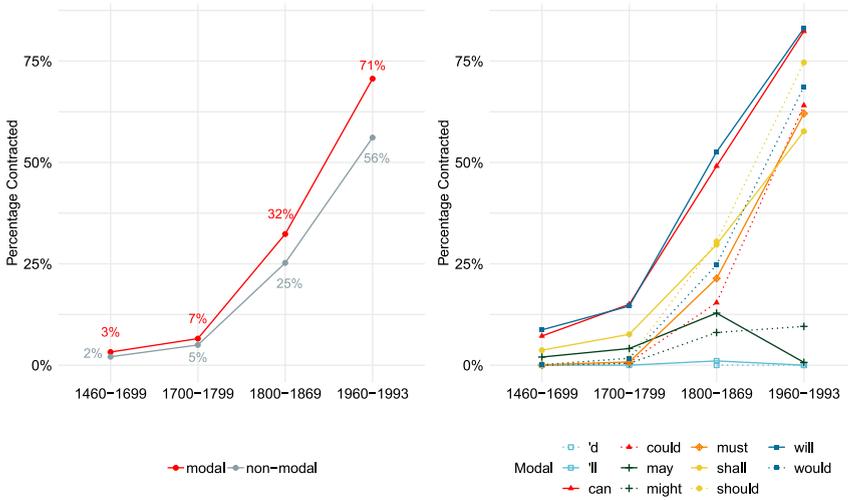
In this section, we take a closer look at the two variants of *NOT*, namely full-form *not* and contracted *n't* (assuming that *n't* and much rarer *'nt* are mere spelling variants). Figure 11 shows a familiar picture (cf. e.g. Daugš 2021: 26; Hejna and Walkden 2022; López-Couso and Pérez-Guerra 2023; Nakamura 2023): contraction of *not* was already an option in EModE, but did not really take off until the nineteenth century. While the present data suggests continuing rapid replacement of full-form *not* by *n't*, a decade-by-decade analysis of *NOT* contraction rates in the Corpus of Historical American English (COHA, Davies 2010–) by Daugš (2021: 26) suggest that—at least in American English—the development follows an S-curve, with the rate of replacement plateauing by the middle of the twentieth century. As the BNC only covers the second half of the twentieth century, it is not possible to tell from the present data, whether this is the case in British English, too.

Each curve in the left panel, however, represents a rather mixed bag of collocates. The grey curve mostly represents negation with forms of *BE*, *HAVE* and *DO*, which show strongly divergent contraction rates (compare, e.g., *\*amn't*, *aren't*, *isn't*; Hejna and Walkden 2022: 79).<sup>17</sup> The same is true for the modals, as becomes evident from the right panel.

A look at each modal separately (Figure 11, right) shows interesting similarities and differences. *Can't*, *won't*, *shan't* and *mayn't* were the pioneers in terms of contraction. Until the eighteenth century, they were the only modals with rates of *not*-contraction exceeding 2%. As predicted by Nakamura (2023), this is exactly the

---

<sup>17</sup> Particularly in the first periods, before operators were obligatory in negated contexts, we find a lot of negation of the type in *I know not*. Yet, even then, modals—including *need*, *dare/durst* and *ought*—as well as other auxiliaries and *do* make up the most frequent collocates of *NOT*, only *know not* and *knew not* rank among the top 20.



**Figure 11:** Contraction rates of NOT overall (left) and in combination with specific core modals (right).

group where the resulting final consonant cluster consists exclusively of /nt/. Their initial ranking (*won't* and *can't* above *shan't* and *mayn't*) suggests that within this group, frequency had an effect on early adoption rates of contraction—either the frequency of the modals by themselves or the combined frequency of modal + NOT (see Figures 6 and 7).

Except with *may*, *might* and the contracted modals *'ll* and *'d*, *n't* goes on to surpass full form *not*. With *will* and *can*, the cross-over happens first: here, contraction rates already reach c. 50 % by the nineteenth century (cf. also Daugs 2021: 26), while with all other modals, *n't* does not surpass *not* until the twentieth century. *Mayn't* and *mighn't*, on the other hand, are declining in popularity and, by the twentieth century, *mayn't* has become all but impossible (cf. also Leech et al. 2009: 81). These patterns suggest that from the nineteenth century onwards, frequency (of modal + NOT) had a broader influence on negative contraction rates among modals.

Overall, the results confirm patterns described elsewhere, particularly those by Hejná and Walkden (2022: 79).<sup>18</sup> Yet, the contraction rates themselves are higher than those based on other sources (see, for instance, contraction rates in Early English Books Online, EEBO, as reported by Nakamura 2023) presumably owing to the informality of the genre (cf. López-Couso and Pérez-Guerra 2023).

<sup>18</sup> This results partly from overlap between the texts contained in the Corpus of Late Modern English Texts used by Hejná and Walkden (2022) and the Chadwyck Healey collections used here.

## 5 Conclusions

This empirical and diachronic evaluation of the relationship between the two Transitivity-lowering parameters negation and modality—operationalised as NOT-negation and core modals—shows that the relationship between the two is complex. In terms of overall usage frequency, there is extensive covariation between the core modals and negation. When an English clause contains a core modal, the odds increase that NOT will also appear and vice versa. However, in terms of negation rates, the modals do actually not constitute one homogeneous block. While *can*, *could*, *will*, *would*, *shall* and *should* attract NOT, the modals *may*, *might* and *must* do not; *might* and *may* mostly even repel NOT-negation. The impression of general co-variation is created by the modals most strongly attracting NOT-negation also being the most frequent ones. Overall, we do not see a trend towards more consistent covariation or towards obligatory pairings between modals and NOT (cf. Hopper and Thompson 1980: 254, 279). There are some indications that modal–NOT clusters may attract further Transitivity-lowering features, such as verbs of perception and cognition, which lower Transitivity because they do not affect the patient (see Bybee 2010: 151–164; Budts 2022: 355).

The paper also discussed whether scope could be used as a predictor of negation rates, i.e. whether modals which allow both a narrow and a wide-scope reading of negation more frequently combine with negation than modals which only allow one scope interpretation. This would correctly predict the high negation rates of *will* and *would* as well as the lower rates of *must* and *shall* but would fail to predict the exceptionally high negation rates of *can* and *could*. Neither does it offer an explanation for why *may*, *might* and the contracted modals *'ll* and *'d* repel negation. Thus the results confirm that variability in scope has little explanatory power when it comes to negation rates.

Furthermore, the study set out to explore whether the solidification of DO-support in the EModE period had an influence on co-occurrence rates of modals and NOT. This could either mean that the modals became “distributionally more and more similar” to dummy DO in the seventeenth century (Budts and Petré 2020: 346) or theirs could be a “story [...] of divergence rather than attraction”, particularly that of *will* and DO (Budts 2022: 359). The results show no evidence of either process taking place: the transition from the EModE period to the eighteenth century is one with few pronounced changes. At this time, we see the strong association between *can/could* and NOT beginning to wane, but other than that, the collostructional analysis rather shows a great deal of stability. We can therefore tentatively conclude that while the modals may have influenced DO (see Budts 2022), they did not come out of the seventeenth century remarkably changed—at least not in terms of their affinity for negation.

The most substantial changes in the modal-negation system after 1500 were brought about by the establishment of the contractions *not* > *n't*, *will* > *'ll* and *would* > *'d* in the nineteenth century. The change from full *not* to the contracted variant was so sweeping that by the twentieth century all full-form modals, except *might* and *may* prefer *n't* over *not*. The availability of the contracted modals *'ll* and *'d* led to a functional split whereby *will* and *would* became much more strongly associated with negation while *'ll* and *'d* emerged as near positive polarity items (if it weren't for their elevated chance to co-occur with *never*).

A comparison with frequencies derived from other corpora suggests that the usage frequency of core modals, their propensity to occur in combination with NOT as well as contraction rates are genre-dependent, although general trends are similar across genres. For instance, Leech et al.'s (2009: 81–82) analysis on the basis of the Brown corpora, which contain a variety of genres, some more formal than the novels used here, finds a “sharp general increase in the use of contracted forms of verbs and negatives” but also that the “steeply declining modals”, *may*, *might*, *must*, *shall* and *ought to*, “have become scarcely usable with negative contractions towards the end of the twentieth century” and that *mayn't* has entirely disappeared. While negative contraction rates in the prose section of the BNC are higher than rates reported by Leech et al. (2009), the frequency-based ranking *mayn't* < ‘declining modals’ + *n't* < other core modals + *n't* is found in both corpora. The same is true for modal frequency and possibly also for modal–negation rates: rankings are similar, but frequencies are higher in prose than in more formal genres. These findings raise the question whether negated uses of modals fulfil specific functions in novels, such as, for instance, conveying subjective perspectives.

## References

- Anthony, Laurence. 2014. *AntConc* (Version 3.4.3m) [Computer Software]. Tokyo: Waseda University. Available at: <https://www.laurenceanthony.net/software>.
- British National Corpus. 1995. BNC Consortium/Oxford University Computing Services.
- Bailey, Guy, Tom Winkle, Jan Tillery & Lori Sand. 1991. The apparent time construct. *Language Variation and Change* 3. 241–264.
- Bates, Elizabeth & Brian MacWhinney. 1987. Competition, variation, and language learning. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 157–193. London/New York: Routledge.
- Bergs, Alexander. 2008. *Shall* and *shan't* in contemporary English – a case of functional condensation. In Graeme Trousdale & Nikolas Gisborne (eds.), *Constructional approaches to English grammar*, 113–143. Berlin: De Gruyter Mouton.
- Beukema, Frits & Wim van der Wurff. 2002. Modals, objects and negation in late Middle English. In Sjeff Barbiers, Frits Beukema & Wim van der Wurff (eds.), *Modality and its interaction with the verbal system*, 75–102. Amsterdam/Philadelphia: John Benjamins.

- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson.
- Budts, Sara. 2022. A connectionist approach to analogy. On the modal meaning of periphrastic do in Early Modern English. *Corpus Linguistics and Linguistic Theory* 18(2). 337–364.
- Budts, Sara & Peter Petr . 2020. Putting connections centre stage in diachronic Construction Grammar. In Lotte Sommerer & Elena Smirnova (eds.), *Nodes and networks in diachronic construction grammar*, 318–351. Amsterdam: John Benjamins.
- Bybee, Joan. 2010. *Language, usage, and cognition*. Cambridge: Cambridge University Press.
- Chang, Winston. 2013. *R graphics cookbook*. Beijing: O'Reilly. R scripts also found at: <http://www.cookbook-r.com> (accessed 23 September 2021).
- Coates, Jennifer. 1983. *The semantics of the modal auxiliaries*. London: Croom Helm.
- Daug, Robert. 2021. Contractions, constructions and constructional change. Investigating the constructionhood of English modal contractions from a diachronic perspective. In Martin Hilpert, Bert Cappelle & Ilse Depraetere (eds.), *Modality and diachronic construction grammar*, 13–51. Amsterdam/New York: John Benjamins.
- Davies, Mark. 2010. *The corpus of historical American English: 400 million words, 1810–2009*. Available at: <https://www.english-corpora.org/coha/>.
- Denison, David. 1993. *English historical syntax*. London/New York: Longman.
- Early English prose fiction*. 1997–2015. Ed. by Holger Klein, David Margolies & Janet Todd. Chadwyck-Healey. ProQuest LLC. Available at: [http://collections.chadwyck.com/marketing/list\\_of\\_all.jspf](http://collections.chadwyck.com/marketing/list_of_all.jspf).
- Eighteenth-century fiction*. 1996–2015. Ed. by Judith Hawley, Tom Keymer & John Mullan. Chadwyck-Healey. ProQuest LLC. Available at: [http://collections.chadwyck.com/marketing/list\\_of\\_all.jspf](http://collections.chadwyck.com/marketing/list_of_all.jspf).
- Elleg ard, Alvar. 1953. *The auxiliary do: The establishment and regulation of its use in English*. Stockholm: Almqvist and Wiksell.
- Flach, Susanne. 2020. Beyond modal idioms and modal harmony: A corpus-based analysis of gradient idiomaticity in MOD + ADV collocations. *English Language and Linguistics* 25(4). 743–765.
- Francis, Naomi & Sabine Iatridou. 2020. Modals and negation. In Viviane D prez & M. Teresa Espinal (eds.), *The Oxford handbook of negation*, 285–300. Oxford: Oxford University Press.
- Giv n, Talmy. 1984. *Syntax. A functional-typological introduction*, vol. I. Amsterdam/Philadelphia: John Benjamins.
- Gries, Stefan Th. 2009. *Quantitative corpus linguistics with R: A practical introduction*. London/New York: Routledge.
- Gries, Stefan Th. 2015. More (old and new) misunderstandings of collocation analysis: On Schmidt and K chenhoff (2013). *Cognitive Linguistics* 26(3). 505–536.
- Gries, Stefan Th. 2021. `coll.analysis.mpfr`. *Collostructional analysis: Computing the degree of association between words and words/constructions*. Software.
- Gries, Stefan Th. 2023. Overhauling collocation analysis: Towards more descriptive simplicity and more explanatory adequacy. *Cognitive Semantics* 9. 351–386.
- De Haan, Ferdinand. 1997. *The interaction of modality and negation*. London/New York: Routledge.
- Hankamer, Jorge. 2011. Auxiliaries and negation in English. In Rodrigo Guti rrez-Bravo, Line Mikkelsen & Eric Potsdam (eds.), *Representing language. Essays in honor of Judith Aissen*, 121–135. Santa Cruz: University of California, Linguistics Research Center.
- Hejna, Mi a & George Walkden. 2022. *A history of English*. Berlin: Language Science Press.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.
- Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: A. F. H st.

- Klima, Edward S. 1964. Negation in English. In Jerry A. Fodor & Jerrold J. Katz (eds.), *The structure of language: Readings in the philosophy of language*, 246–323. Englewood Cliffs, NJ: Prentice-Hall.
- Laing, Margaret. 2002. Corpus-provoked questions about negation in early Middle English. *Language Sciences* 24. 297–321.
- Lazard, Gilbert. 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia Linguistica* 36(3–4). 141–190.
- Leclercq, Benoît. 2022. From modals to modal constructions: An n-gram analysis of *can*, *could* and *be able to*. *Constructions and Frames* 14(2). 226–261.
- Leech, Geoffrey. 2003. Modality on the move: The English modal auxiliaries 1961–1992. In Roberta Facchinetti, Frank Palmer & Manfred Krug (eds.), *Modality in contemporary English*, 223–240. Berlin/New York: Mouton De Gruyter.
- Leech, Geoffrey. 2013. Where have all the modals gone? An essay on the declining frequency of core modal auxiliaries in recent standard English. In I. Marín-Arrese Juana, Marta Carretero, Jorge Arús Hita & Johan van der Auwera (eds.), *English modality. Core, periphery and evidentiality*, 95–115. Berlin/Boston: Mouton de Gruyter.
- Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in contemporary English*. Cambridge: Cambridge University Press.
- Levshina, Natalia. 2015. *How to do linguistics with R. Data explorations and statistical analysis*. Amsterdam/Philadelphia: John Benjamins.
- López-Couso, María José & Javier Pérez-Guerra. 2023. Promoting and inhibiting forces at work: A corpus-based analysis of negative contraction in the recent history of English. In Hendrik De Smet, Peter Petré & Benedikt Szmrecsanyi (eds.), *Context, intent and variation in grammaticalization*, 189–214. Berlin/Boston: De Gruyter Mouton.
- Malchukov, Andrej. 2006. Transitivity parameters and transitivity alternations. Constraining co-variation. In Leonid Kulikov, Andrej Malchukov & Peter de Swart (eds.), *Case, valency and transitivity*, 329–371. Amsterdam/Philadelphia: John Benjamins.
- Millar, Neil. 2009. Modal verbs in TIME: Frequency changes 1923–2006. *International Journal of Corpus Linguistics* 14(2). 191–220.
- Mindt, Dieter. 1995. *An empirical grammar of the English verb: Modal verbs*. Berlin: Cornelsen.
- Mondorf, Britta & Ulrike Schneider. 2016. Detransitivization as a support strategy for causative *bring*. *English Language and Linguistics* 20(3). 439–462.
- Nineteenth-century fiction*. 1999–2000. Ed. by Danny Karlin & Tom Keymer. Chadwyck-Healey. ProQuest LLC. Available at: [http://collections.chadwyck.com/marketing/list\\_of\\_all.jsp](http://collections.chadwyck.com/marketing/list_of_all.jsp).
- Nakamura, Fujio. 2023. Caught in the web of mysteries involved in the history of negative contractions: With special reference to the periods and chronological order of the establishment and diffusion of negative contractions. *Paper presented at ICEHL-22*, Sheffield, 3–6 July.
- Palmer, Frank. 1997. Negation and modality in Germanic languages. In Toril Swan & Olaf J. Westvik (eds.), *Modality in Germanic languages: Historical and comparative perspectives*, 133–149. Berlin: Mouton de Gruyter.
- Puskás, Genoveva. 2018. *To wish or not to wish: Modality and (metalinguistic) negation*. *Glossa: A Journal of General Linguistics* 3(1). 1–33.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London/New York: Longman.
- R Development Core Team. 2009. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>.
- Radden, Günter. 2009. Affirmative and negated modality. *Quaderns de Filologia: Estudis Linguistics* 14. 169–192.

- Roberts, Ian. 1985. Agreement parameters and the development of English modal auxiliaries. *Natural Language & Linguistic Theory* 3. 21–58.
- Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2). 149–182.
- Römer, Ute. 2004. A corpus-driven approach to modal auxiliaries and their didactics. In John Sinclair (ed.), *How to use corpora in language teaching*, 185–202. Amsterdam/Philadelphia: John Benjamins.
- Scherl, Magdalena. 2019. Language change as a challenge. *gitlab*. <https://mscherl.pages.gitlab.rlp.net/pos-website/> (accessed 3 August 2022).
- Schmid, Hans-Jörg. 2000. *English abstract nouns as conceptual shells. From corpus to cognition*. Berlin/New York: Mouton de Gruyter.
- Schmid, Hans-Jörg & Helmut Küchenhoff. 2013. Collostructional analysis and other ways of measuring lexicogrammatical attraction: Theoretical premises, practical problems and cognitive underpinnings. *Cognitive Linguistics* 24(3). 531–577.
- Schneider, Ulrike. 2021a. Loss of intersective gradience as the lifeboat of a dying construction. An analysis of the diachronic change of causative *bring*. *Folia Linguistica Historica* 42(2). 429–459.
- Schneider, Ulrike. 2021b. The syntax of metaphor. A diachronic case study of the transitivity of *derail*. *Yearbook of the German Cognitive Linguistics Association* 2021(9). 47–70.
- Schneider, Ulrike. 2023. Reflexive analytic causatives: A diachronic analysis of transitivity parameters. *English Language & Linguistics*. 1–25. First View. <https://doi.org/10.1017/s136067432300028x>.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
- Strang, Barbara M. H. 1970. *A history of English*. London: Methuen.
- Szmrecsanyi, Benedikt. 2003. BE GOING TO versus WILL/SHALL. Does syntax matter? *Journal of English Linguistics* 31(4). 295–323.
- Tagliamonte, Sali A. & Jennifer Smith. 2002. Either it isn't or it's not". NEG/AUX contraction in British dialects. *English World-Wide* 23(2). 251–281.
- Tottie, Gunnel. 1985. The negation of epistemic necessity in Presentday British and American English. *English World-Wide* 6. 87–116.
- Tsunoda, Tasaku. 1985. Remarks on transitivity. *Journal of Linguistics* 21. 385–396.
- Varela, Pérez & José Ramón. 2013. Operator and negative contraction in spoken British English: A change in progress. In Bas Aarts, Joanne Close, Geoffrey Leech & Sean Wallis (eds.), *The verb phrase in English: Investigating recent language change with corpora*, 256–285. Cambridge: Cambridge University Press.
- Visser, Fredericus Th. 1969. *An historical syntax of the English language—Syntactical units with two verbs*, vol. 3, first half. Leiden: E. J. Brill.
- Warner, Anthony R. 1993. *English auxiliaries. Structure and history*. Cambridge: Cambridge University Press.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English N'T. *Language* 59(3). 502–513.