



## Article

Robert Daugš\* and David Lorenz

# A radically usage-based, collostructional approach to assessing the differences between negative modal contractions and their parent forms

<https://doi.org/10.1515/cllt-2024-0051>

Received May 2, 2024; accepted May 10, 2024; published online July 16, 2024

**Abstract:** Starting from the premise that English negative modal contractions constitute partly variable patterns of associations that include both the preceding subject and the following verb infinitive, the study sets out to investigate distributional differences between *can't*, *shouldn't*, and *won't* and their corresponding uncontracted parent forms. Given that some configurations are assumed to correlate with specific modal meanings (e.g. inanimate subjects and stative verbs > 'epistemic prediction'; first person subjects > '(un)willingness' or 'commissive modality'), roughly 200,000 trigrams from COCA are submitted to *distinctive covarying collexeme analysis* in order to uncover if these contractions and their full forms are conventionalized and entrenched differentially enough to merit their separate treatment on both conceptual and methodological grounds. The results point to probabilistic tendencies, suggesting a cline where *won't* and *can't* appear to be more emancipated from their respective full-form analogue than *shouldn't*. Furthermore, the study showcases how collostructional methods can be applied fruitfully to case studies embedded in Schmid's (Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*. Oxford: Oxford University Press) *Entrenchment and Conventionalization Model*.

**Keywords:** negative modal contractions; distinctive covarying collexeme analysis; EC-Model

---

\***Corresponding author: Robert Daugš**, Englisches Seminar, Kiel University, Leibnizstraße 10, 24118 Kiel, Schleswig-Holstein, Germany, E-mail: [daugs@anglistik.uni-kiel.de](mailto:daugs@anglistik.uni-kiel.de)

**David Lorenz**, English Studies, Joint Faculties of Humanities and Theology, Lund University, Lund, Skåne, Sweden, E-mail: [david.lorenz@englund.lu.se](mailto:david.lorenz@englund.lu.se)

# 1 Introduction

A growing body of research indicates that several contractions in present-day English (e.g. *gonna*, *'ll*, *I dunno*) are past the stage of being mere online phonetic reductions of their corresponding full forms (here *BE going to*, *will*, *I do not know*) and rather constitute emancipated constructions (Bresnan 2021; Bybee 2010; Daugas 2022; Krug 2000; Lorenz 2013a, 2013b; Nesselhauf 2014; Scheibman 2000; Schmidtke-Bode 2009; to name a few). Similar claims have been made for the negative modal contractions *can't*, *won't*, and *shan't*, due to their idiosyncratic morphosyntax, (discourse-)functional preferences, and unpredictable distributional behavior (Bergs 2008; Daugas 2021). Moreover, their invariant phonological and graphemic representation in present-day English (PDE) attests to both their lexical storage as (parts of) entrenched patterns into the minds of speakers and a high degree of conventionality. In addition, at least *can't* and *won't* have clearly ousted their parent forms in terms of their overall frequency of use and have diffused in various registers and genres, which corroborates the impression that they are not just colloquial pronunciation variants. Nor does it seem to be the case that the contractions simply replace the full forms. Despite expected overlaps in their use, each form has a distinctive collexemic profile from which functional preferences can be inferred (cf. Hilpert 2008, 2016).

The present investigation builds on these findings, delving deeper into the collostructional biases of *can't*, *shouldn't*, and *won't*. Our aim is to test if and how speakers distinguish between the three contractions and their uncontracted counterparts *cannot*, *should not*, and *will not*, by relying on their probabilistic knowledge of language. To assess this, data retrieved from the *Corpus of Contemporary American English* (COCA; Davies 2008) are submitted to *distinctive covarying collexeme analysis* (Stefanowitsch and Flach 2020), which is designed to compare variable patterns with multiple slots – here, SUBJ {*can't*|*won't*|*shouldn't*} v versus SUBJ {*cannot*|*will not*|*should not*} v. By going beyond verbal collexemes, which has been the default in previous studies on modal constructions (e.g. Daugas 2020, 2021; Gries and Stefanowitsch 2004; Hilpert 2008, 2016), we seek to do justice to (i) the fact that modals are not only linked to the following but also their previous co-text via syntagmatic associations and (ii) the established correlations between specific modal interpretations and properties of the subject (e.g. ANIMACY, PERSON) (Coates 1983). Since each of the contractions at hand belongs to a different semantic cluster (namely ‘ability/possibility’, ‘obligation/necessity’, and ‘volition/prediction’, respectively; Coates 1983: 27–29) and is entrenched and conventionalized to different degrees, *can't* and *won't* arguably more so than *shouldn't*, the findings

are, in sum, expected to contribute to understanding negative contractions specifically and contractions more generally.

The paper is structured as follows. Section 2 outlines the theoretical background on modal contractions as constructional patterns in an associative network, drawing heavily on tenets from Construction Grammar (CxG; Goldberg 2006; Hilpert 2021) and the *Entrenchment-and-Conventionalization Model* (EC-Model; Schmid 2015, 2020). An overview of the data, methods, and findings from our corpus study will be presented in Section 3. Following this, Section 4 brings together the findings and measures them against the theoretical backdrop of how modal contractions are stored and processed in an associative network. Finally, we will reflect on some conceptual and methodological issues when it comes to assessing the status of contractions using constructional methods.

## 2 The theory: negative modal contractions as patterns of associations and conventionalized utterance types

Modal expressions have recently gained serious interest within usage-based, constructionist frameworks (see e.g. Depraetere et al. 2023; Hilpert et al. 2021; Hohaus and Schulze 2020 and the respective articles therein). At first glance, it appears as though treating modal verbs as constructions (i.e. entrenched pairings of form and meaning; Goldberg 2006) adds hardly any substance to understanding their intricacies, let alone modality as such. After all, as words, they represent constructions by definition and, if combined with their infinitival collocates, they are (usually) semantically transparent and predictable from the general behavior of auxiliary constructions. However, a compelling argument for a constructionist approach is provided by Hilpert (2016). Although any MOD + V combination is theoretically possible, each modal may be used with some verbal collexemes more (or less) often than would be expected by chance, which results in distinctive collocational profiles. Such probabilistic preferences are assumed to be part of the tacit knowledge of speakers and, what is more, they cannot be predicted based on any other knowledge of language. While this view already emphasizes the role of connections in an associative network rather than the role of constructions as nodes, it can be advanced even further and applied fruitfully to the present case.

From the more dynamic, radically usage-based perspective of the EC-Model, negative modal contractions can simply be viewed as patterns of associations

(cf. Schmid 2020: 27–28).<sup>1</sup> The main advantage of this approach is that it circumvents the issue of having to postulate whether they constitute constructions (i.e. nodes) or not, based on properties such as unpredictability or sufficient frequency; see Hilpert (2019) for a comprehensive overview of different diagnostics to test constructionhood.<sup>2</sup> Instead, differences in the use and meaning of negative modal contractions and their parent forms are best understood as differences in the strength of the different kinds of associations (i.e. connections) they evoke (cf. Schmid 2015: 11–16, 2020: Ch. 4, 11). Essentially, this allows for a unified treatment of all three contractions *can't*, *shouldn't*, and *won't* but also acknowledges their differences. *Shouldn't*, for example, represents a pattern of symbolic (i.e. the form-meaning correspondence), syntagmatic (i.e. the mutual attraction between neighboring elements), paradigmatic (i.e. the probabilistic knowledge of onomasiological competitors), and pragmatic associations (i.e. the contextual contingency) no less than *can't* and *won't*, but their respective associations are differentially routinized (cf. Schmid 2020: 45–48, 229–234).

To illustrate, in terms of its complexity, *shouldn't* constitutes a bimorphemic word where the constituent elements *should* and *n't* are linked by strong syntagmatic associations. Yet, it remains formally segmentable and semantically transparent, which means its components are linked to those of other formal and functional relatives like, for example, *mustn't* ('obligation/necessity') or *wouldn't* ('epistemicity/hypothesis') through paradigmatic associations. The symbolic associations evoked by *shouldn't* are thus likely to largely operate compositionally, that is, pertain to its individual parts. Apart from its pragmatic associations with more informal genres, *shouldn't* may thus be routinized quite similarly to *should not*. By contrast, the suppletion and coalescence exhibited by *can't* and *won't* render active parsing rather improbable (due to a lack of autonomous forms \**ca* and \**wo*). In their case, both the symbolic and paradigmatic associations are likely evoked holistically without

---

1 The term *dynamic* refers to the inherent nature of living languages, which are conceived of as complex-adaptive systems that are continuously updated through the interplay of social processes, cognitive processes, and usage activities of speakers. While this assumption may be (implicitly) shared by virtually all cognitive-functional, usage-based strands of grammar, the EC-Model vehemently doubles down on it by advertising the importance of fluid links in an associative network over seemingly stable, stored mental representations to account for persistence, variation, and change in language (Schmid 2020: 46, 339–345, et passim). *Radically usage-based* suggests both highly redundant, exemplar-based storage of information and a minimal abstractionist view (i.e. abstractions may only be postulated if necessary and empirically supported; cf. Hilpert 2013; Dausgs 2022), which stands in contrast to complete inheritance models that generally thrive for economy and maximum elegance.

2 The constructional status of *can't* and *won't* is actually rather uncontroversial, considering their idiosyncratic morphosyntax (for a more detailed discussion, see Dausgs 2021). *Shouldn't*, on the other hand, represents a more challenging case from the 'traditional', node-oriented perspective.

depending on syntagmatic associations (similar to monomorphemic words, but probably unlike *cannot* and *will not*; cf. Schmid 2020: 242).

Also, based on the assumption that every utterance that is perceived leaves traces in memory upon first exposure (Goldberg 2019: 54), there is no practical reason for assessing whether something qualifies as a construction or not based on sufficient frequency, unless a principled distinction is made between stored ‘pre-constructional’ patterns and constructions proper. Goldberg’s (2019) idea of clusters of ‘lossy’ memory traces or Schmid’s (2015, 2020) proposal of patterns of associations, however, avoid this problem altogether and provide a more parsimonious take on the matter. Repeated (or recessive) activation simply strengthens (or weakens) different kinds of associations or memory traces, which leads to a higher (or lower) degree of entrenchment thereof.

In sum, the question whether a given contraction is frequent enough or can be predicted based on any other construction(s) (here, probably the respective non-negative forms and a more abstract *xn’t* construction) is of no immediate importance. From the dynamic view, differentially entrenched connections capture any differences between the mental representations of contractions and their respective full forms without the need to determine their constructionhood (cf. Schmid 2020: 233). The crucial question then is whether (some) contractions are routinized differently enough from their parent forms to merit a separate treatment.

In the present study, we will conceive of negative modal contractions as partly variable, lexico-grammatical patterns. Next to the contraction itself as the pivotal element, each pattern also includes two variable elements, namely the preceding subject and the following verb infinitive.<sup>3</sup> The same applies to the respective parent forms, against which the contractions are measured; compare the three pairs in (1) that are under scrutiny here.

- (1) a. SUBJ *can’t* v    b. SUBJ *shouldn’t* v    c. SUBJ *won’t* v  
       SUBJ *cannot* v    SUBJ *should not* v    SUBJ *will not* v

It stands to reason that any differences in the use of contracted and uncontracted forms (e.g. SUBJ *can’t* v vs. SUBJ *cannot* v) will be considerably more subtle than, for example, between different contractions (e.g. SUBJ *can’t* v vs. SUBJ *won’t* v). With but a few highly idiomatized exceptions like *Can’t say I VP* (‘mitigation, agreement, or approval’), *You shouldn’t have* (‘expression of gratitude’) or *That dog won’t hunt* (‘certainty that a plan is not going to materialize’), negative contractions and full

<sup>3</sup> Obviously, this restriction automatically excludes any other syntactic configurations the contractions might be used in (e.g. inversion, adverb modification, ellipsis), which potentially diminishes the generalizability of our results to some extent. This step was, however, deemed necessary to reduce noise in the data and assess the variation between contracted and uncontracted forms more accurately, as both contractions and full forms are assumed to be generally interchangeable in this environment. Also, the configuration selected is by far the most frequent one.

forms, when used with the same subjects and verbal collexemes, convey more or less the same meaning; compare the examples in (2)–(4):

- (2) ‘circumstantial impossibility/permission not granted’
- a. Basically, **we can’t allow** a district or a school to stay on priority improvement or a turnaround plan for more than five years. [COCA, Denver, 2012]
  - a’. For the sake of our people and our economy **we cannot allow** gridlock to prevail. [COCA, CBS\_NewsMorn, 2011]
- ‘epistemic impossibility’
- b. But of course **it can’t be** the same lion. That was forty years ago in northern Botswana. I’ve never heard of a lion living more than twenty-five years. [COCA, Bk:FeverDream, 2010]
  - b’. While some students claimed that advertising had influenced their buying decisions, **it cannot be** the case, as the 14 chosen products are rarely advertised in traditional media outlets. [COCA, CollegStud, 2011]
- ‘inability’
- c. Where did we ever get the ridiculous notion that **children can’t understand** spiritual things? [COCA, Atlanta, 2002]
  - c’. **Children cannot understand** or get the notion of something unknown. [COCA, People, 1993]
- (3) ‘weak obligation/recommendation’
- a. You cannot eat a bowl of cereal with milk in a car while you’re driving, on a train, or **you shouldn’t eat** it on a train. [COCA, CBS\_Morning, 2018]
  - a’. In your book you say that **you should not eat** bacon or sausage with sodium nitrate. How important is it to stay away from sodium nitrate? [COCA, CNN\_King, 2004]
- ‘epistemic inference’
- b. [W]ith all the talk about circular firing squads inside the Obama White House, **it shouldn’t come** as a surprise that some [...] believe that chief of staff Rahm Emanuel is badmouthing his colleagues [...]. [COCA, AmericanSpectator, 2010]
  - b’. Due to the campus’ relatively close proximity to Hartford, **it should not come** as a surprise that over a dozen potential informants had been identified before interviewing commenced. [COCA, SportBehavior, 2007]
- ‘hypothesis (contrasting with *wouldn’t/would not*)’
- c. Walk around if you like. **I shouldn’t think** it would matter, sir. [COCA, Analog, 2005]

- c'. As to his learning, **I should not think** it an objection if he resembled Sir Roger de Coverly's chaplain in some respects [...]. [COCA, ChurchHistory, 2005]
- (4) 'permission not granted'
- a. **You won't marry** me off to the highest bidder. Father said I could marry as I pleased. He promised. [COCA, Bk:TemptedByYour, 2002]
- a'. You'll do no such thing. **You will not marry** just to get your hands on those account ledgers. [COCA, Bk:AllINeedIsYou, 1998]
- 'epistemic prediction'
- b. You see, without food, the **medicine won't work**. [COCA, CNN\_YourWorld, 2005]
- b'. Antibiotic **medicine will not work** on viruses. They are only effective against bacteria. [COCA, ChildLife, 2001]
- 'unwillingness'
- c. So raising the awareness level and making it clear **we won't tolerate** it is one of the clear things we have to do. [COCA, CNN\_Event, 2005]
- c'. This announcement should send a firm and unequivocal message to anyone who would engage in dishonest or illegal financial activity that the Justice Department doesn't and **we will not tolerate** such activities. [COCA, PBS\_Newshour, 2014]

The examples corroborate the impression that negative contractions and their parent forms share a common meaning potential, which, arguably, casts doubts on the validity of a differential treatment beyond register connotations. But the crucial point is that the distinction need not be a categorical one. Rather each pattern will show probabilistic tendencies as to which subjects and verbal collexemes are attracted to it (cf. Hilpert 2021: 75). This translates into differentially entrenched syntagmatic associations, which, depending on their relative strength, may manifest in relatively distinctive collexemic profiles that provide insights into the sense distribution of a given pattern and potentially give access to hidden semantic structures.

Finally, it is probably uncontroversial that the negative modal contractions constitute more or less strongly conventionalized utterance types at the communal level; compare, *can't*, *shouldn't*, and *won't* (relatively high degree of conventionality) with, for example, *shan't*, *mightn't*, or *mayn't* (relatively low degree of conventionality). The reduction and fusion of *x not* to *xn't* was likely completed before the 16th century and the forms remain phonemically and graphemically stable in PDE (cf. Jespersen 1917: 117; McElhinney 1992: 369). But despite the general increase in usage frequency of negative modal contractions over the last centuries and their diffusion

to other genres (see e.g. Dausgs 2020, 2021; Leech et al. 2009; Mair 1997), they are still often distinguished from their parent forms based on how speakers use them in order to conform to the regularities expected in different genre-specific contexts, for example, the normative avoidance of contracted over uncontracted forms in formal writing. We will address this issue further in Section 3.

What remains to be seen is whether a case can be made for contractions and their parent forms being conventionalized in a way that makes their choice also contingent on onomasiological and semasiological conformity. In other words, is it more conventional to use, for example, *will not* over *won't* to reach the communicative goal of expressing 'unwillingness', as in (4c) and (4c'), because their respective sense distributions make *will not* the more probable candidate in this context?

### 3 The corpus study: distributions and idiomatic combinations

#### 3.1 Corpus data

We used *R* (R Core Team 2019) to retrieve the instances of the contractions and full forms at hand from a commercial version of COCA (~450 million words, 1990–2012). The search was limited to selected pronominal subjects (personal pronouns, existential *there*, *this*, *that*, *who*, and *which*). Noun phrases were excluded because subject properties like ANIMACY, which can provide important cues for specific modal interpretations, are considerably more difficult to detect from simple trigrams. Consider the example in (5).

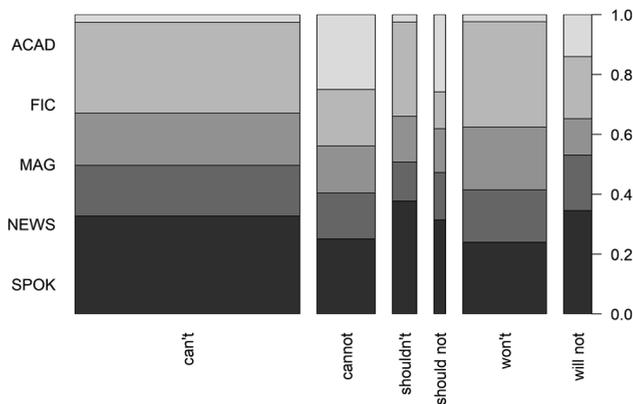
- (5) The hitch is the **company won't guarantee** calculations are correct with the standard version. They do with the deluxe version. [COCA, USAToday, 1999]

Without the immediate co-text, it would arguably not be immediately clear that the noun *company* is used metonymically in the SUBJ *won't* v pattern in (5), as is obvious from the subject *they* in the following sentence. Beyond that, nominal subjects usually have a relative proclivity to co-occur with uncontracted rather than contracted forms, which also justifies their exclusion.

No restrictions on time span and genre were imposed. For two reasons, this is not unproblematic: (i) the 22-year period covered by the corpus introduces a potentially confounding diachronic component, while the varied sources that feed into it may create artefacts, actual grammatical variation and change notwithstanding, and (ii) contractions, as markers of colloquial speech, are usually assumed to have a

distributional bias towards less formal genres. Concerning the first point, while short-term grammatical changes in AmE have been reported previously by, for example, Leech et al. (2009), these pertain first and foremost to changes in overall usage frequency, rather than changes in collexemic preferences, which the present study seeks to investigate.<sup>4</sup> Furthermore, the status of modals as grammatical elements, which are generally rather evenly dispersed across texts (Hilpert and Correia Saavedra 2017), as well as the overall size and balanced structure of COCA (by year and genre) give reason to believe that source variability will not have much impact on the overall results.

With regard to the second point, Figure 1 shows the presumed relative underuse of negative modal contractions in formal, academic contexts. However, it also corroborates previous findings that modals (both contracted and uncontracted) are generally relatively more frequent in spoken(-like) genres (Leech et al. 2009: 77), at least when used with pronominal subjects. Also, a preliminary investigation of the



**Figure 1:** Genre-specific relative frequency distribution of English negative modal contractions and their parent forms; COCA, AmE, 1990–2012 ( $N_{\text{tokens}} = 199,317$ ).

<sup>4</sup> Leech et al. (2009: 79–89, 284) looked at changes in the sense distribution of specific modals as well, also *should (not)*, which requires taking verbal collexemes into consideration, among other things. However, based on their data, a generalized linear fixed-effects model with MODAL (should.1961 vs. should.1991) as the dependent and SENSE (epistemic vs. root vs. quasi-subjunctive vs. hypothetical vs. unclear) as the independent variable ( $\text{LR}\chi^2 = 27.80$ ;  $df = 4$ ;  $p < 0.0001$ ;  $R^2 = 0.022$ ;  $C = 0.548$ ) indicates that only the quasi-subjunctive use of *should*, which was quite rare to begin with, has declined significantly over a 30-year period in AmE, not to mention the overall very poor model discrimination. We therefore think it unlikely that the 22-year period covered in the present study would reveal any dramatic reshuffling of the overall sense distributions or collexemic preferences of the modal expressions at hand.

collexemic preferences of the patterns investigated here showed that the rankings remain highly stable regardless whether the ACAD data are included or not (in all three cases, Spearman's  $\rho > 0.94$ ;  $p < 0.0001$ ); for a similar observation regarding modal enclitic + ADV combinations, see Flach (2021a: 756).

The final data set contains 199,317 observations of SUBJ {*can't*|*cannot*|*shouldn't*|*should not*|*won't*|*will not*} v, with 12 pronoun types and 2,935 verb types, which still amounts to about 72 % of all tokens and 88 % of all verb types of nominal and pronominal subjects in the six patterns combined. SUBJ *can't* v is by far the most frequent pattern (accounting for 52 % of all observations), followed by SUBJ *won't* v (19 %), SUBJ *cannot* v (13 %), SUBJ *will not* v (7 %), SUBJ *shouldn't* v (6 %), and finally SUBJ *should not* v (3 %).<sup>5</sup>

### 3.2 Distinctive covarying collexeme analysis

Distinctive covarying collexeme analysis (DCCA; Stefanowitsch and Flach 2020) is the most recent addition to the family of collostructional methods. Conceptually, it represents a combination of distinctive collexeme analysis (Gries and Stefanowitsch 2004) and covarying collexeme analysis (Stefanowitsch and Gries 2005), which makes it possible to compare two or more constructions with multiple slots regarding the elements that are associated with them. Mathematically, DCCA is a configural frequency analysis (CFA; von Eye 1990), through which observed combinations are tested for their above- or below-chance occurrence. CFA typically outputs *a*-corrected chi-square-based rankings, thereby determining *types* (i.e. combinations that are statistically significantly more frequent than expected) and *anti-types* (i.e. combinations that statistically significantly less frequent than expected). In a collostructional context, types would correspond to items (strongly) attracted to a given pattern, whereas anti-types would correspond to items that are (strongly) repelled.

Applied to the present cases, DCCA assesses which combinations of a pronominal subject and a verbal collexeme are distinctive for either a contracted or uncontracted pattern. In other words, it tests all possible configurations of SUBJ  $\times$  *xn't*|*not*  $\times$  v for their relative over- and underuse. The analysis was performed using a modified version of Flach's (2021b) R package {collostructions}. To obtain the rankings, we used  $G^2_{\text{simple}}$  (*simple log-likelihood*; Evert 2009), a mathematical approximation of the more widely known and commonly used  $G^2$  (*log-likelihood ratio*) that only requires the observed and expected frequencies of the respective target node (i.e. the top

---

<sup>5</sup> Note that in the following analyses, we collapsed the pronouns *she* and *he*, as there are no differences between them in terms of PERSON and ANIMACY.

leftmost cell in a contingency table) as input without cross-classification.<sup>6</sup> Like  $G^2$ ,  $G^2_{\text{simple}}$  is an *evidence-based* measure, which means that it favors high-frequency (i.e. more evidence) over low-frequency (i.e. less evidence) observations and can underestimate the actual strength of an association (Evert 2009: 1227–1228). In fact, Gries (2022) argues that  $G^2$  (and, as a consequence,  $G^2_{\text{simple}}$ ) is hardly an association measure at all because it not only conflates frequency and association but actually correlates more strongly with the former. Since rigorous testing of different measures on larger than  $2 \times 2$  contingency tables in a collostructional context is to our knowledge still pending, we will continue to use  $G^2_{\text{simple}}$  but also check it against ‘true’ (i.e. frequency-independent) association and directionality by obtaining the *surprisal*-values (i.e. the negative  $\log_2$ -transformed forward transition probabilities) for each configuration in a pattern (cf. Schmid 2020: 52).<sup>7</sup> Figure 2 plots the relationship between  $G^2_{\text{simple}}$  and surprisal based on generalized additive models (GAMs) for the top 100 distinctive covarying collexemes of each pair.

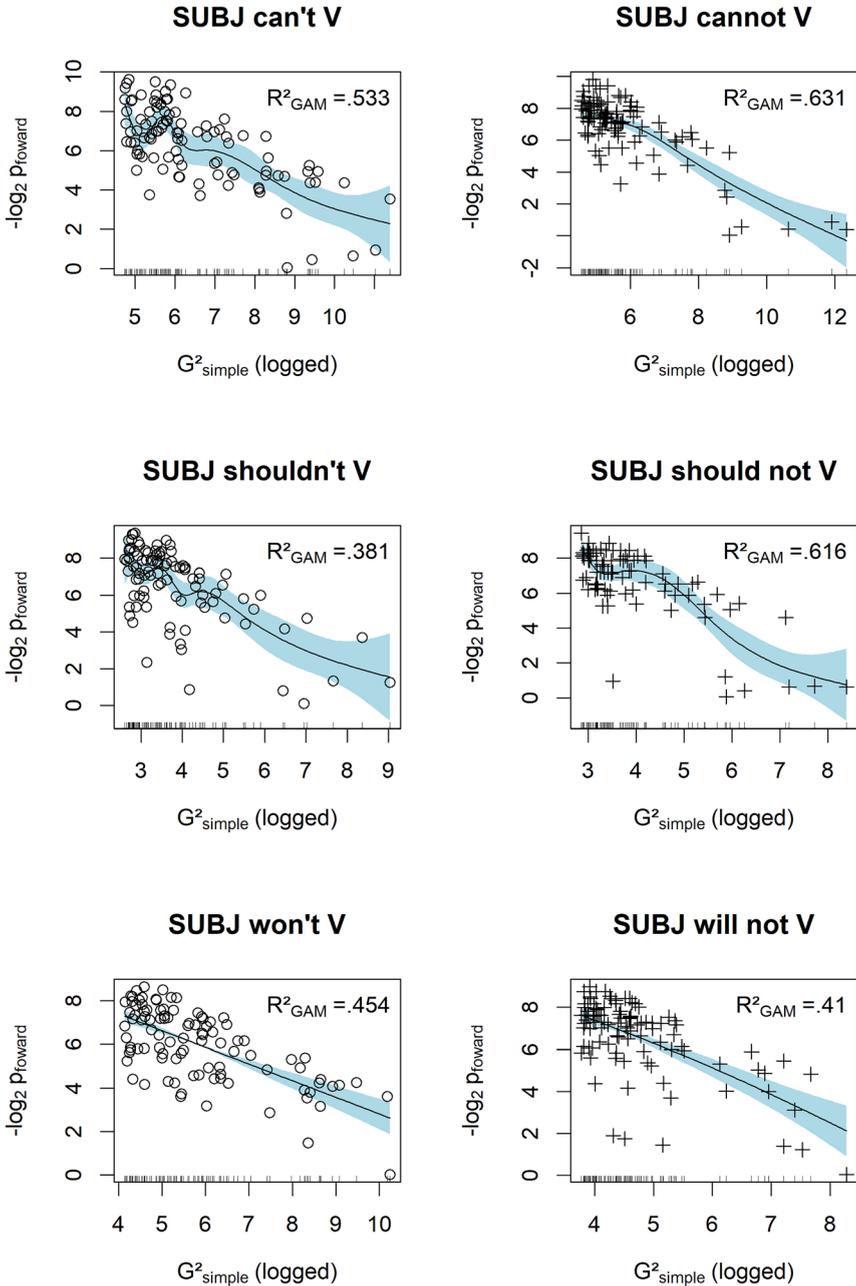
Depending on the pattern, the deviance explained by the GAMs is between 38.8 and 64.2 %, which confirms that the two measures do not completely reflect the same thing (cf. Gries 2022). But the plots also show a negative correlation indicating that higher  $G^2_{\text{simple}}$ -values predict stronger forward syntagmatic associations and thus faster reading times at least to some extent. We will therefore continue to focus on  $G^2_{\text{simple}}$  but also report surprisal-values anecdotally, mainly as a safeguard to check whether both measures completely deviate from one another. Finally, in Section 4, we discuss the implications of choosing one measure over the other in the present context.

### 3.3 Results

By grouping the results according to one of the words occurring in the variable slots of two related patterns (i.e. contraction vs. full form), it is possible to highlight their onomasiological competition (or lack thereof). Figure 3 illustrates this for selected SUBJ + MOD<sub>neg</sub> combinations.

<sup>6</sup> Simple log-likelihood is defined as follows:  $G^2_{\text{simple}} = 2 \cdot \left( \text{Obs} \cdot \log_{\text{Exp}} \frac{\text{Obs}}{\text{Exp}} - (\text{Obs} - \text{Exp}) \right)$ . We thank Stephanie Evert for her consult on this measure.

<sup>7</sup> Surprisal is calculated as  $-\log_2 p(w_{n+2}|w_n w_{n+1}) = -\log_2 p(v|\text{SUBJ MOD}_{\text{neg}}) = -\log_2 \frac{\text{Obs}(\text{SUBJ MOD}_{\text{neg}} v)}{\text{Obs}(\text{SUBJ MOD}_{\text{neg}})}$ . The  $-\log_2$ -transformation renders surprisal-values proportional to reading times (Linzen and Jaeger 2014: 10): the higher the transition probability of an item given a context, the lower its degree of surprisal and cognitive effort, which translates into faster reading times. In other words, the smaller the value of  $-\log_2 p_{\text{forward}}$  for a given stimulus, the faster speakers are expected to react to that stimulus.



**Figure 2:** The relationship between  $G^2_{\text{simple}}$  (logged) and surprisal ( $-\log_2 p_{\text{forward}}$ ) as measures of collostructional strength used in DCCA for the top 100 distinctive covarying collexemes of each negative contraction/full form pair.

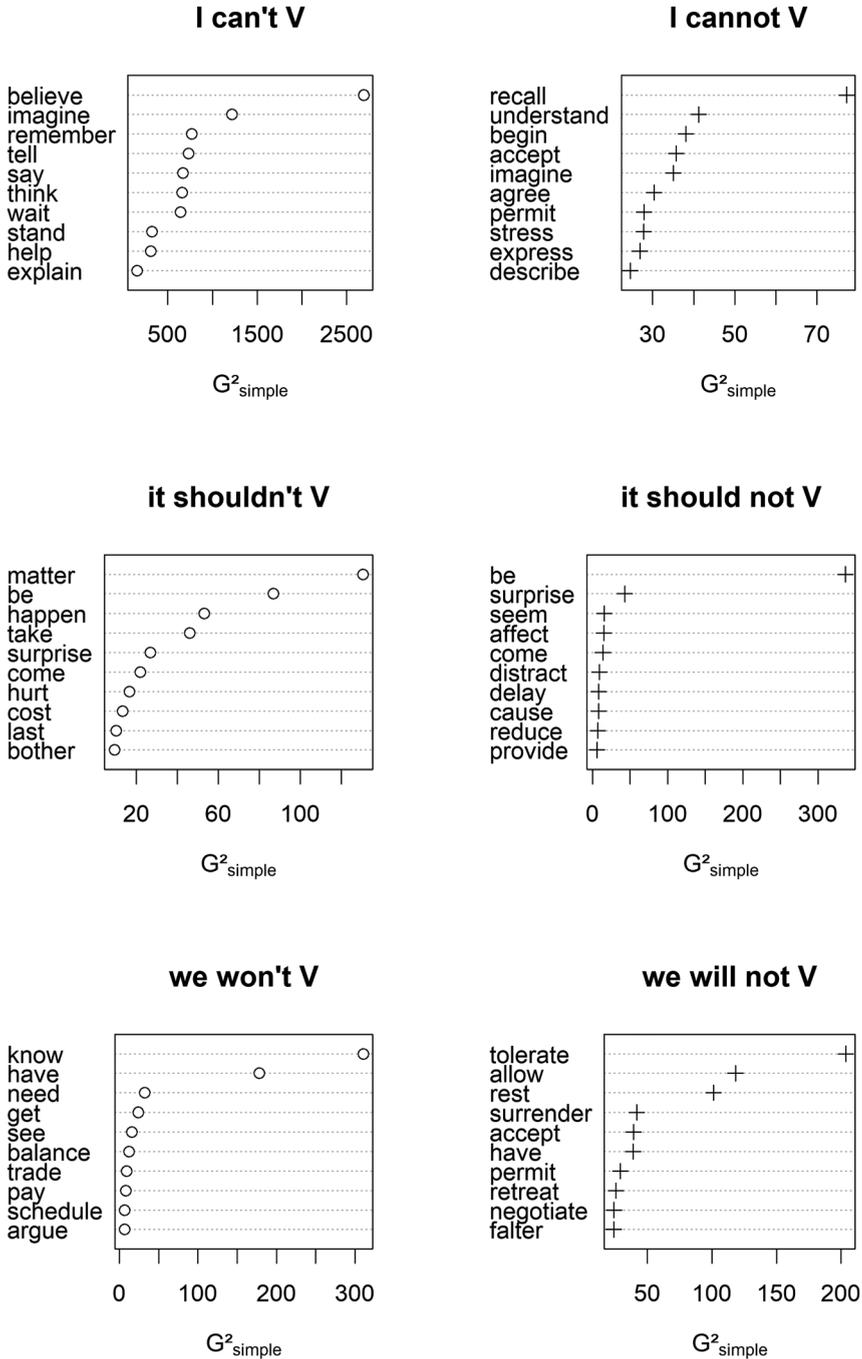


Figure 3: Top ten verbal collexemes of selected SUBJ + MOD<sub>neg</sub> combinations; COCA, AmE, 1990–2012.

Several observations can be made from these plots. For *I can't* v versus *I cannot* v, it is clear that the contracted pattern attracts both cognition verbs (e.g. *believe* [rank 1], *imagine* [2], *remember* [3]) and communication verbs (e.g. *tell* [4], *say* [5], *explain* [10]), which, in combination with the subject *I*, seem to be rather formulaic than compositional (e.g. *I can't believe* 'surprise or emphasis', *I can't tell (you how ...)* 'emphasis'). It has also been noted by Bybee (2010: Ch. 9) that the pattern *can't* +  $v_{\text{cognition}}$  can show noteworthy frequency asymmetries, i.e. some combinations are more frequent than the affirmative variant with *can*, which further attests to their status as a prefabs, as affirmatives are generally more frequent than negatives cross-linguistically. Similarly, the verb *help* (9) when used in the utterance *I can't help but*  $v_{\text{P}}$  forms part of a complex pattern that holistically evokes the symbolic association 'compulsion to do something'. By contrast, the pattern *I cannot* v inter alia attracts verbs like *accept*, *agree*, *permit*, or *stress*, which can be expected in permissive contexts where the interlocuter refers to the impossibility of a situation or refusal rather than an intrinsic inability (cf. Daugs 2021: 32).

In the case of *it shouldn't* v versus *it should not* v, such (fully) idiomatized sequences are missing from the data, which is perhaps unsurprising, considering the relatively lower frequency of *shouldn't* compared to, for example, *can't*. Instead, we see more overlap in the use of the contraction and its parent form, evidenced by the share of stative verbs among the top collexemes (e.g. *be*, *seem*, *matter*, *happen*, *surprise*). The collexemic demarcation between *shouldn't* and *should not* in this particular configuration seems thus less clear.

Finally, with regard to *we won't* v and *we will not* v, the sharply demarcated semantic range exhibited by the verbs in the uncontracted pattern is noteworthy, as nine out of the top ten of them (except *have*) invite an 'unwillingness' or 'commissive' interpretation that can be expected in contexts where seriousness and determination are typically (and emphatically) conveyed, for example, political discourse (cf. Daugs 2020: 29–30, 2021: 32–33). The contracted pattern, on the other hand, has among its most attracted verbal collexemes, stative verbs like *know*, *have*, *need*, and *see*, which point to an 'epistemic prediction' reading. In terms of their collocational profiles, these two patterns appear to be the most distinctive.

This procedure can obviously be extended to all other subjects. Alternatively, the patterns can be grouped by specific verbs and tested for the kinds of subjects they combine with more or less than would be expected, which would substantiate the claim that the subject preferences are crucial, at least from a distributional perspective. To illustrate, while *believe* is highly attracted to *can't* in combination with *I* (obs: 3,055; exp: 1,005.7;  $G^2_{\text{simple}}$ : 2,690.3; surprisal: 3.5), it is repelled in combination with *you* (obs: 255; exp: 905.9;  $G^2_{\text{simple}}$ : -655.4; surprisal: 6.9). Conversely, *you can't* attracts *get* (obs: 1,963; exp: 1,316.3;  $G^2_{\text{simple}}$ : 275.6; surprisal: 4.1), whereas *I can't* repels it (obs: 1,356; exp: 1,461.3;  $G^2_{\text{simple}}$ : -7.8; surprisal: 4.7).

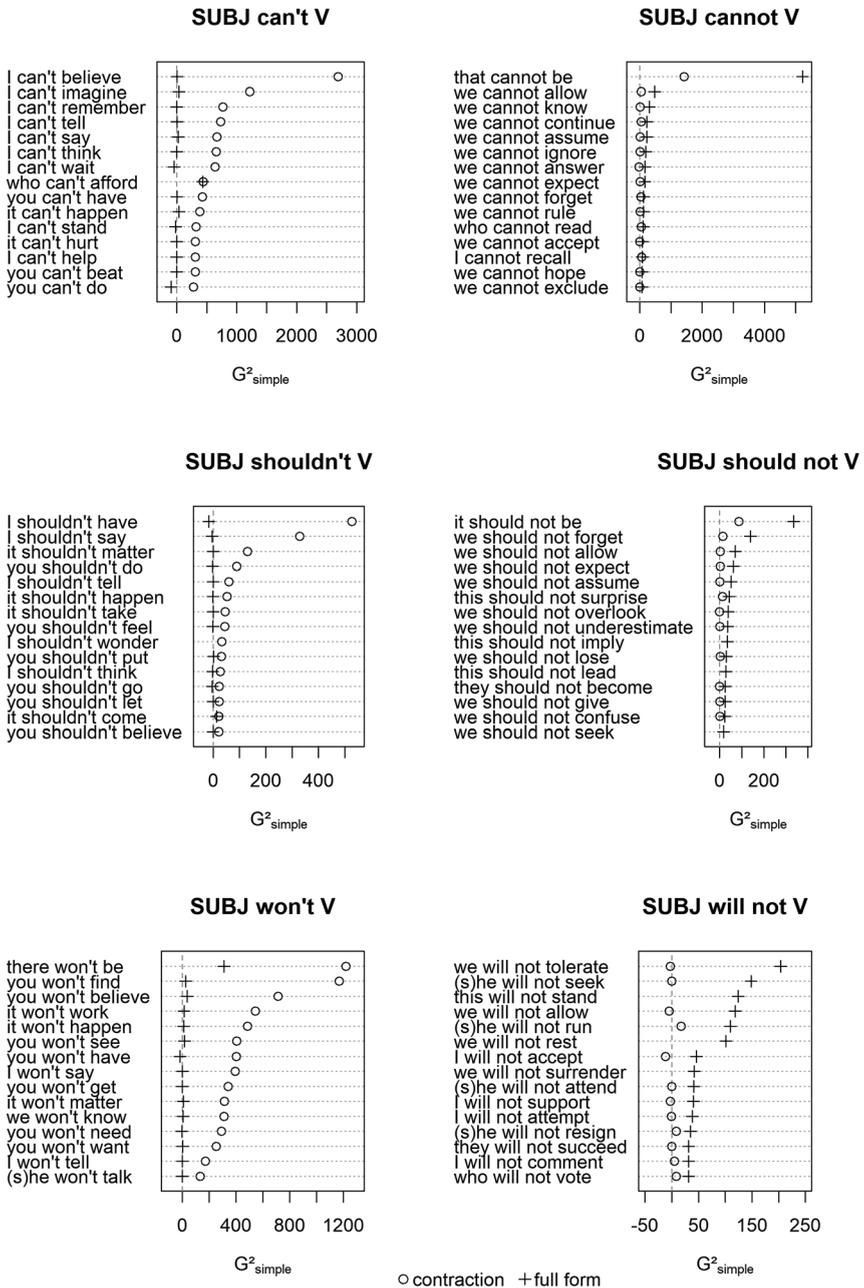
To account for such subject preferences and, at the same time, adopt a broader perspective on each pattern, we first ranked all configurations of a given pair by their  $G^2_{\text{simple}}$ -score and selected the top 15 trigrams. Verbal duplicates for each pattern were removed and, in order to avoid doubling across patterns, preference was given to whichever pattern had a higher  $G^2_{\text{simple}}$ -score in cases of overlap among the top-ranked collexemes. The results are shown in Figure 3.

Unlike the original implementation of distinctive collexeme analysis (Gries and Stefanowitsch 2004), which completely separates the data for a given pair (Stefanowitsch 2006), DCCA allows for overlaps, namely that a SUBJ *v* combination might be attracted/repelled for both patterns in a pair. This seems to be more realistic in the present case given the degree of relatedness between contractions and their parent forms; see, for example, *who can't afford* and *who cannot afford* (first panel, left). Still, from Figure 4, we learn that each pattern shows some degree of demarcation from its corresponding relative and, although our approach relativizes strong collexemic preferences for specific verbs (e.g. without duplicate removal, the first five configurations of SUBJ *cannot v* would be *that, it, which, this, and there* in combination with *be*), it allows us to identify potential semantic clusters.

Compared to the results discussed in Daus (2021), where the list of verbal collexemes of *can't* (based on a bigram analysis) was full of dynamic action verbs, the data shown in the top plots of Figure 4 are surprising at first sight. Here, only *you can't beat* [rank 14] and *you can't do* [15] would fall in that category, whereas the highly idiomatic triples with *I* and verbs of cognition and communication are ubiquitous and clearly occupy the top. Yet, the notion that *can't* is predominantly used to express intrinsic inability on part of the subject is solidified here, nonetheless, as verbs like *believe, remember, or think* coupled with *I* relate to inherent properties (cf. Coates 1983: 89–91; Daus 2021: 40).<sup>8</sup> The pattern with *cannot*, on the other hand, shows a different distribution in that 80 % of all triples listed here co-occur with *we* and mainly verbs conveying a situational impossibility/refusal (e.g. *allow, expect, accept*). Such combinations seem to onomasiologically compete with the 'self-imposed obligation/recommendation' sense of *should not* rather than referring to inherent properties of the subject. Also note the overwhelmingly distinctive combination, *that cannot be*, which can confidently be predicted to convey epistemic meaning. In sum, the findings corroborate the impression that *can't* is relatively preferred in 'ability' contexts and *cannot* in other contexts.

---

<sup>8</sup> Other dynamic action verbs are also identified as distinctive for the contraction by the DCCA, but they are not ranked among the top 15 collexemes. Compare, for example, *you can't go* (obs: 1,098; exp: 718.5;  $G^2_{\text{simple}}$ : 172.3; surprisal: 4.9) versus *you cannot go* (obs: 108; exp: 185.8;  $G^2_{\text{simple}}$ : -38.4; surprisal: 5.4) or *they can't get* (obs: 605; exp: 392.1;  $G^2_{\text{simple}}$ : 98.9; surprisal: 3.7) versus *they cannot get* (obs: 78; exp: 101.4;  $G^2_{\text{simple}}$ : -5.9; surprisal: 5.3).



**Figure 4:** The 15 most distinctive subj v combinations in contracted and uncontracted negative modal patterns; COCA, AmE, 1990–2012.

Consider the middle panel next. There seems to be a slight tendency for the contraction to attract dynamic verbs (*do, take, put, go, come*), yet, given the subject diversity of both *shouldn't* and *should not* it is difficult to identify a cluster. Nonetheless, as in the case of *cannot*, *should not* shows an inclination to co-occur with *we*. These cases arguably convey a warning or a moral, possibly camouflaged obligation imposed on the subject not to engage in the activity expressed by the main verb. By using *we* instead of, for example, *you*, the speaker might engage their interlocutor in such a way that the overtness of the obligation is mitigated. Previous research has linked such a strategy to a *democratization* trend in English (see e.g. Fairclough 1992; Farrelly and Seoane 2012; Leech et al. 2009). What is striking though is that democratization has also been argued to be linked to *informalization* (Hiltunen and Loureiro-Porto 2020). While both *shouldn't* and *should not* are less overt markers of authority than, for example, deontic *must*, it would be expected that the contraction, as a marker of colloquialness and informality, rather than the full form would attract *we*. Aside from that, both patterns co-occur with inanimate subjects (*it, this*) and stative verbs (e.g. *be, matter, happen*), thereby likely expressing epistemic necessity. It thus appears as though the patterns with *shouldn't* and *should not* are still more intertwined than the other two pairs.

Finally, the relative preference of *we will not v* to convey 'unwillingness', as discussed earlier, prevails even when other subjects are considered. It might even be extended to a more general SUBJ<sub>1ST.PERS</sub> *will not v* pattern, since the lower right-hand panel of Figure 3 also has *I* in combination with similar verbs (e.g. *accept, support*). Interestingly, the only triple with an inanimate subject among the combinations with *will not* is *this will not stand* 'not accepting a given situation', which, counter to the inanimacy-epistemicity correlation, also signals 'determination' rather than 'prediction'.<sup>9</sup> By contrast, *won't* seems to be favored over *will not* in combination with stative verbs and inanimate subjects (e.g. *there won't be, it won't work, it won't happen, it won't matter*), which points to a preference for 'epistemic prediction'. Additionally, combinations of *won't* with second person *you*, which are completely missing from the pattern with *will not*, appear to form another cluster of low-dynamicity actions, namely *you won't {find|believe|see|have|get|need|want}*, where the speaker seems to make a (confident) epistemic prediction about the subject.

---

<sup>9</sup> This particular triple is inter alia part of a speech by former US President George H. W. Bush from 1990, "This will not stand, this aggression against Kuwait", and was picked up in the Cohen brothers' (1998) crime comedy *The Big Lebowski*, which, given the movie's pop-cultural relevance, might have arguably furthered its usualization and diffusion.

## 4 Discussion: negative modal contractions and their full forms in an associative network

In this section, we will interpret the findings within the framework of the EC-Model laid out by Schmid (2015, 2020). We will also discuss the contribution DCCA can bring to the table when it comes to identifying differences and similarities in alternations. The aim of this study was to revisit the status of negative modal contractions compared to their corresponding parent forms. While it has already been shown that, at least, *can't* and *won't* seem to have emancipated themselves from *cannot* and *will not* respectively (Dausgs 2021), a refinement of these findings is in order. According to the radically usage-based underpinnings of the EC-Model, stored mental representations are best understood as “more or less routinized latent pattern[s] of associations” (Schmid 2020: 46). Furthermore, the model assumes a minimal abstractionist view, that is, units like schemas are only postulated if necessary (Schmid 2020: 55, 215).

With this in mind, we first need to take a step back. Traditionally, it is assumed that modals like *can* or *will* subsume different forms, namely *can(not)*, *can't* and *will(not)*, *'ll*, *won't* respectively. Given that all of these represent conventionalized utterance types, they all have the potential to regularly license usage events. The higher this potential for a given utterance (which is contingent on conformity), the more likely it will be used in speech situations and will thus have a higher chance to be activated as a pattern of associations in the minds of the speakers (Schmid 2020: 6–7). Regarding the negative contractions and their full forms, speakers would have to choose between them (and obviously others) to reach their communicative goal(s) of expressing modal notions like ‘inability’, ‘circumstantial impossibility’, ‘prediction’, ‘obligation’, or ‘unwillingness’. In turn, each pattern can potentially convey more than one conventionalized and entrenched meaning or function; see, for example, Dausgs (2023) for a discussion on modals as many-to-many mappings. These two perspectives capture the symbolic associations (i.e. the form-meaning connections) as well as the paradigmatic associations (i.e. the semasiological and onomasiological choices) that the patterns at hand evoke. The internal interaction between the elements within each pattern, namely *SUBJ*, *MOD<sub>neg</sub>*, and *v*, represents the syntagmatic associations, which connect the contraction (as well as its full form) to its co-text, while pragmatic associations link each pattern to a given context.

The stance taken here is that the processing of negative modal contractions and their parent forms unfolds in line with the EC-Modal in the following way, exemplified by the utterance in (6).

- (6) (Rep) Steven Rothman: We're moving the country forward. **We will not allow** this country to go back to the policies that brought us to the brink of disaster. [COCA, PBS\_NewsHour, 2010]
- (7) You may appeal to the administrator's sense of good education, but **it won't hurt** to mention the legal implications of insufficient planning and participation. [COCA, MusicEduc, 1991]

The communicative setting in (6) is established by the speaker's position as a member of the U.S. House of Representatives, which activates the corresponding pragmatic associations with POLITICAL DISCOURSE. The utterance of interest starts with the deictic subject *We* (here: reference to the Democratic Party and the Obama administration), which establishes expectations about a following verb phrase. The negated modal auxiliary *will not* is then picked over *won't* (and other competitors, e.g. *should not* or *cannot*) to reach a specific communicative goal, which is to express the 'unwillingness' to accept certain political strategies, while arguably, at the same time, conveying 'seriousness' and 'determination'. Also, in anticipation of the following main verb, *allow*, which can be expected in these contexts, the onomasiological choice of *will not* is reinforced, since the whole pattern forms a more cohesive sequence that conforms well with the established conventions in such specific usage events. That is not to say that *won't* or other competitors are not available here at all, only that the use of *will not* may be more probable.<sup>10</sup> The data from the DCCA actually back this up, as the different frequency-related measures available all suggest *will not* to be the more suitable candidate in this sequence; compare *we will not allow* (obs: 59; exp: 9.3;  $G^2_{\text{simple}}$ : 118.3; surprisal: 4.9) versus *we won't allow* (obs: 17; exp: 27.6;  $G^2_{\text{simple}}$ : -4.7; surprisal: 7.6). Consider yet another example. In (7), we have the sequence *it won't hurt*, which expresses the notion of 'prediction'. The speaker suggests a strategy that will not likely result in a negative outcome but will rather be beneficial to the subject. Once *it* is uttered, it becomes highly improbable to expect anything but an 'epistemic' interpretation, given the strong correlation between inanimate subjects and this particular modal function (Coates 1983: 181). This would of course be the same for the full form. However, the relative frequency of the sequence with *won't* alone points to a highly idiomatized expression that could easily outmatch other expressions competing for selection in such a communicative

---

<sup>10</sup> It is quite interesting to note that the previous utterance in (6) includes the contraction *We're*, which obviously can be expected in spoken discourse (almost as the default). The fact the speaker does not continue with contracted *won't* here fuels the suspicion that *will not* shows more conformity when it comes to conveying the communicative goal at hand, even regardless of a possible priming effect.

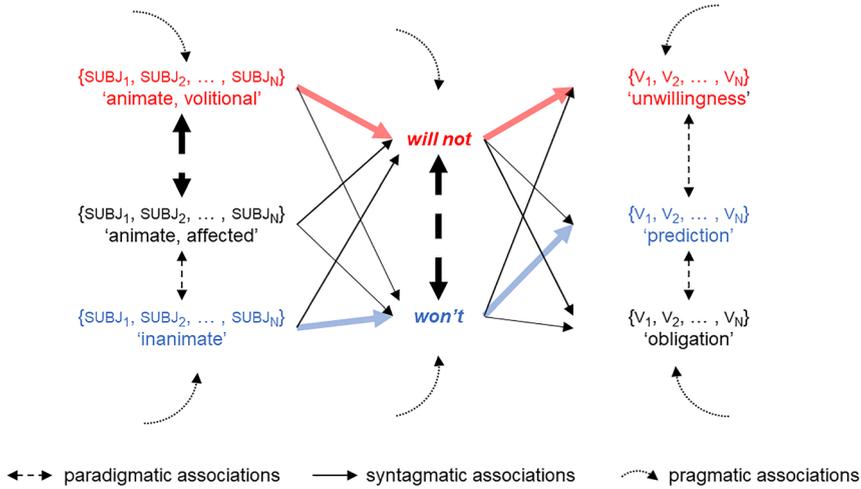
setting.<sup>11</sup> Again, the DCCA data corroborates this impression; compare *it won't hurt* (obs: 118; exp: 39.6;  $G^2_{\text{simple}}$ : 100.9; surprisal: 5.7) versus *it will not hurt* (obs: 9; exp: 13.4;  $G^2_{\text{simple}}$ : -1.6; surprisal: 7.8).

Based on these observations and the quantitative findings, we endorse our initial proposal that negative modal contractions and their parent forms deserve to be treated separately, though with some important caveats. First, there is no doubt that any contraction and its corresponding full form will show notable overlap in their respective use. The historical relationship between a contracted form and its uncontracted relative simply produces such behavior. In the case of *will not* and *won't*, the forms might be different enough to develop further idiosyncrasies and finally become fully autonomous. Second, when used with the same subject and the same verb infinitive, there is no reason to believe that the contracted and the uncontracted pattern convey a different meaning or communicative function. However, we suggest probabilistic tendencies, that is, even though, for example, *I can't help* and *I cannot help* are possible, they are entrenched and conventionalized to different degrees. Third, the differences uncovered here are highly contingent not only on specific verbs but also specific subject choices. *Won't* does not universally convey 'epistemic prediction' (see the example in 4a), nor does *can't* always express 'intrinsic inability' (see the example in 2a). Rather there are some  $\text{SUBJ} + \text{V}$  combinations that clearly favor the contracted pattern, whereas others favor the uncontracted one. Fourth, not all contractions behave alike in the sense that a differential treatment is feasible in every case. The previous sections have shown that in terms of formal as well as distributional properties, *can't* and *won't* seem to be much more emancipated than *shouldn't*. It is thus an empirical question to decide whether assuming a common underlying representation in the form of a schema is merited or not (for a similar discussion on English modal enclitics, see Dausgs 2022; cf. also Barðdal 2008; Hilpert 2013).

Against the backdrop of the EC-Model, it was suggested here that contractions and their parent forms represent differentially conventionalized utterance types at the level of the speech community. As such, they license usage events and will do so more effectively, the more a specific combination of  $\text{SUBJ} + \text{MOD}_{\text{neg}} + \text{V}$  conforms with the established conventions. Also, contractions and their parent forms represent differentially entrenched, variable patterns of associations at the level of the individual speakers. As such, they are part of an emergent, dynamic, associative network, where specific combinations of  $\text{SUBJ} + \text{MOD}_{\text{neg}} + \text{V}$  are more strongly routinized than others. In any case, each contraction as well as each full form (and all other models for that matter) can potentially convey different meanings. The polysemous nature of modal expressions has long been accepted and studied meticulously. In line with the

---

<sup>11</sup> An exception here might be the sequence *it doesn't hurt*.



**Figure 5:** Idealized associative network of SUBJ *won't* v and SUBJ *will not* v; line widths represent the relative strength of the respective association; color codings emphasize the relative preference of a pattern for conveying a specific modal function, as identified by the DCCA.

distributional hypothesis (Firth 1957), the sense distribution of a contraction (or its parent forms) is contingent on the degree of entrenchment of the different kinds of associations, which, in turn, is contingent on *usage intensity*.<sup>12</sup> Figure 5 summarizes our claims in a highly simplified, idealized associative network for the cases of *will not* and *won't*.

The visualization aims to capture the importance of connections over nodes, as it is essentially connections (i.e. the types of associations) that are entrenched to different degrees (Schmid 2020: 342–343). Based on sequences like *I/we will not {accept, allow, permit, tolerate, ...}* uncovered by the DCCA, we could postulate a SUBJ<sub>1,pers</sub> *will not* v pattern that distinctively conveys ‘unwillingness’. From sequences like *it/that/this won't {happen, matter, cost, ...}*, a SUBJ<sub>3,pers.inanimate</sub> *won't* v pattern that is preferably used to express ‘epistemic prediction’ is also conceivable. In a CxG framework, these more specified patterns would correspond to meso-constructions (and, in some cases, even micro-constructions). If such lower-level instances give indication about the meaning differences between contractions and full forms, there is no reason beyond elegance to maintain the idea of a more schematic generalization that captures all the properties of both contractions and their parent forms. In fact, there is no reason to assume a general full form or contraction pattern. Instead,

<sup>12</sup> The term *usage intensity* is adopted from Stefanowitsch and Flach (2017). It subsumes different corpus-based measures as operationalizations of entrenchment (and conventionalization).

we can simply focus on the connections, which would allow us to stay close to the data, as corpus distributions provide insights about variation and change in the links rather than the nodes of a network; see Hilpert (2021: Ch. 3) for an intriguing discussion of the ‘fat node problem’ in CxG.

Finally, there are some conceptual and methodological issues to be considered. While it is generally acknowledged that usage intensity is crucial to both the cognitive entrenchment of patterns of associations and their conventionalization as utterance types in the speech community, the exact nature of this force in general and how to assess it from corpus and experimental data specifically continues to be subject of lively debates (see e.g. Blumenthal-Dramé 2012, 2017; Deshors and Gries 2022; Divjak and Caldwell-Harris 2015; Flach 2020; McConnell and Blumenthal-Dramé 2022; Schmid 2020; Stefanowitsch and Flach 2017). Given the multifacetedness of both entrenchment and conventionalization, different measures will account for different aspects of either. We are not going to rehearse these debates in full but instead briefly focus on two issues.

First, seeing that entrenchment is a process that happens in the minds of individual speakers (Schmid 2017: 10; Schmid and Mantlik 2015: 584; Schmid et al. 2021), aggregated corpus data as the default empirical source pose a conceptual as well as a methodological challenge (see also Petré and Anthonissen 2020 and references therein). Individual speaker variation is acknowledged here, but still our data represent somewhat of a black box in this regard. Any claim towards entrenchment must thus be taken with a grain of salt, since we are dealing with language data averaged across different speakers without knowing what and how much each speaker has actually contributed. Second, we need to review how we have operationalized degrees of entrenchment and conventionality in the present study. Essentially, the DCCA was used to test for cohesive combinations in three different  $\text{SUBJ} \times \text{XN'T}|\text{not} \times \text{V}$  patterns, which basically translates into assessing the strength of syntagmatic associations, while accounting for paradigmatic associations at the same time. For this, we used two measures of collostructional strength:  $G^2_{\text{simple}}$  and surprisal (cf. Section 3.2). Schmid (2020: 52) argues that the degree of entrenchment (and arguably conventionalization) of syntagmatic associations can be approximated using transitional probabilities, but they need to be cross-checked against statistical measures that incorporate information about the overall corpus frequencies of the elements within a sequence as well as the corpus size (e.g.  $G^2$ ). However, since the goal of collostructional analysis is to uncover clusters of semantically similar collexemes, we opted to give preference to  $G^2_{\text{simple}}$ , seeing that the rankings obtained were much easier to interpret. In fact, if preference had been given to surprisal, all patterns would have listed combinations of an inanimate or existential subject (e.g. *it, that, which, there*) and the stative verb *be* among the top configurations. What all this shows is that these subjects are simply an extremely strong syntagmatic signpost

for activating expectations about the soon-to-follow stative verb because there are not very many highly frequent alternatives, regardless of the negative modal that is used along with them; see Coates (1983) on correlations and syntactic co-occurrence patterns of modal verbs in general.  $G^2_{\text{simple}}$  at least incorporates information about the paradigmatic alternations of a given pattern, which, among other factors, has been argued to determine the strength of syntagmatic associations during processing (cf. Schmid et al. 2021). Importantly, the configurations that turned out on top of the list based on  $G^2_{\text{simple}}$ -values usually also outperformed their competing variant in terms of surprisal-values.

To conclude, we have demonstrated that negative modal contractions (*can't*, *shouldn't* and *won't*) and their parent forms (*cannot*, *should not*, *will not*) show some interesting distributional differences in that each pattern seems to have a relative preference for co-occurring with specific subjects and verb infinitives, thereby conveying a specific modal meaning (i.e. 'deontic', 'dynamic', 'epistemic'). We explicitly adopted a radically usage-based perspective by treating each contraction and its corresponding full form as conventionalized utterances and differentially entrenched variable patterns of associations. In order to assess the strength of the associations that each pattern triggers, our aim was to identify cohesive SUBJ MOD<sub>neg</sub> V combinations by means of DCCA, the latest member of the family of collostructional methods. The major advantages of DCCA are that it allows the researcher to adopt a broader perspective on alternating patterns by incorporating more co-text and that it also allows for distributional overlaps, which, specifically in the present case, seems to capture the relationship between such closely related patterns more accurately. In sum, our data indicate that differences in meaning or function between contractions and full forms seem to manifest in more specified, less schematic patterns, which is in accordance with the dynamic, minimal abstractionist view of exemplar-based, network-oriented models (cf. also Daus 2022).

## References

- Barðdal, Jóhanna. 2008. *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.
- 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*, 113–144. Berlin: De Gruyter.
- Blumenthal-Dramé, Alice. 2012. *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Berlin: De Gruyter.
- Blumenthal-Dramé, Alice. 2017. Entrenchment from a psycholinguistic and neurolinguistic perspective. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 129–152. Berlin: De Gruyter.

- Bresnan, Joan. 2021. Formal grammar, usage probabilities, and auxiliary contraction. *Language* 97(1). 108–150.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Coates, Jennifer. 1983. *The semantics of the modal auxiliaries*. London: Croom Helm.
- Dausgs, Robert. 2020. Revisiting global and intra-categorical frequency shifts in the English modals: A usage-based, constructionist view on the heterogeneity of modal development. In Pascal Hohaus & Rainer Schulze (eds.), *Re-assessing modalising expressions: Categories, co-text, and context*, 17–46. Amsterdam: John Benjamins.
- Dausgs, 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*, 12–52. Amsterdam: John Benjamins.
- Dausgs, Robert. 2022. English modal enclitic constructions: A diachronic, usage-based study of 'd and 'll. *Cognitive Linguistics* 33(1). 221–250.
- Dausgs, Robert. 2023. *Modality, usage and diachrony: Constructional changes in the modal domain in American English*. Kiel: Kiel University PhD dissertation.
- Davies, Mark. 2008. *The Corpus of Contemporary American English (COCA)*. Available at: <https://www.english-corpora.org/coca/>.
- Depraetere, Ilse, Bert Cappelle, Martin Hilpert, Ludovic De Cuyper, Mathieu Dehouck, Pascal Denis, Susanne Flach, Natalia Grabar, Cyril Grandin, Thierry Hamon, Clemens Hufeld, Benoît Leclercq & Hans-Jörg Schmid. 2023. *Models of modals: From pragmatics and corpus linguistics to machine learning*. Boston: De Gruyter.
- Deshors, Sandra C. & Stefan Th. Gries. 2022. Using corpora in research on second language psycholinguistics. In Aline Godfroid & Holger Hopp (eds.), *The Routledge handbook of second language acquisition and psycholinguistics*, 164–177. New York: Routledge.
- Divjak, Dagmar & Catherine L. Caldwell-Harris. 2015. Frequency and entrenchment. In Ewa Dabrowska & Dagmar Divjak (eds.), *Handbook of cognitive linguistics*, 53–75. Berlin: De Gruyter.
- Evert, Stefan. 2009. Corpora and collocations. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, vol. 2. Berlin: De Gruyter.
- Fairclough, Norman. 1992. *Discourse and social change*. Cambridge: Polity Press.
- Farrelly, Michael & Elena Seoane. 2012. Democratization. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *The Oxford handbook of the history of English*, 392–401. Oxford: Oxford University Press.
- Firth, John R. 1957. *Papers in linguistics*. Oxford: Oxford University Press.
- Flach, Susanne. 2020. Schemas and the frequency/acceptability mismatch: Corpus distribution predicts sentence judgments. *Cognitive Linguistics* 31(4). 609–645.
- Flach, Susanne. 2021a. 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.
- Flach, Susanne. 2021b. *Collostructions: An R implementation for the family of collostructional methods* (v.0.2.0). Available at: <https://sfla.ch/collostructions/>.
- Goldberg, Adele. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldberg, Adele. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press.
- Gries, Stefan Th. 2022. What do (some of) our association measures measure (most)? Association? *Journal of Second Language Studies* 5(1). 1–33.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on “alternations”. *International Journal of Corpus Linguistics* 9(1). 97–129.

- Hilpert, Martin. 2008. *Germanic future constructions: A usage-based approach to language change*. Amsterdam: John Benjamins.
- Hilpert, Martin. 2013. *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.
- Hilpert, Martin. 2016. Change in modal meanings: Another look at the shifting collocates of *may*. *Constructions and Frames* 8(1). 66–85.
- Hilpert, Martin. 2019. *Construction grammar and its application to English*, 2nd edn. Edinburgh: Edinburgh University Press.
- Hilpert, Martin. 2021. *Ten lectures on diachronic construction grammar*. Leiden: Brill.
- Hilpert, Martin, Bert Cappelle & Ilse Depraetere (eds.). 2021. *Modality and diachronic construction grammar*. Amsterdam: John Benjamins.
- Hilpert, Martin & David Correia Saavedra. 2017. Why are grammatical elements more evenly dispersed than lexical elements? Assessing the roles of text frequency and semantic generality. *Corpora* 12(3). 369–392.
- Hiltunen, Turo & Lucía Loureiro-Porto. 2020. Democratization of Englishes: Synchronic and diachronic approaches. *Language Sciences* 79. 101275.
- Hohaus, Pascal & Rainer Schulze (eds.). 2020. *Re-assessing modalising expressions: Categories, co-text, and context*. Amsterdam: John Benjamins.
- Jespersen, Otto. 1917. *Negation in English and in other languages*. Copenhagen: Ejnar Munksgaard.
- Krug, Manfred. 2000. *Emerging English modals: A corpus-based study of grammaticalization*. Berlin: De Gruyter.
- Leech, Geoffrey, Marianne Hundt, Christian Mair & Nicholas Smith. 2009. *Change in contemporary English: A grammatical study*. Cambridge: Cambridge University Press.
- Linzen, Tal & Florian Jaeger. 2014. Investigating the role of entropy in sentence processing. In *Proceedings of the fifth workshop on cognitive modeling and computational linguistics*, 10–18. Baltimore, Maryland, USA: Association for Computational Linguistics.
- Lorenz, David. 2013a. *Contractions of English semi-modals: The emancipating effect of frequency*. Freiburg: NIHIN/Universitätsbibliothek Freiburg.
- Lorenz, David. 2013b. From reduction to emancipation: Is *gonna* a word? In Hilde Hasselgård, Jarle Ebeling & Signe Oksefjell Ebeling (eds.), *Studies in corpus linguistics*, vol. 57, 133–152. Amsterdam: John Benjamins.
- Mair, Christian. 1997. Parallel corpora: A real-time approach to the study of language change in progress. In Magnus Ljung (ed.), *Corpus-based studies in English: 200 Papers from the 17th International Conference on English Language Research on Computerized Corpora (ICAME17)*. Stockholm, 15–19 May 1996, 195–209. Amsterdam: Rodopi.
- McConnell, Kyla & Alice Blumenthal-Dramé. 2022. Effects of task and corpus-derived association scores on the online processing of collocations. *Corpus Linguistics and Linguistic Theory* 18(1). 33–76.
- McElhinney, Bonnie. 1992. The interaction of phonology, syntax and semantics in language change: The history of modal contraction in English. In Costas P. Canakis, Grace P. Chan & Jeannette Marshall Denton (eds.), *Chicago Linguistic Society 28: Papers from the 28th Regional Meeting of the Chicago Linguistic Society 1992*, 367–381. Chicago: Chicago Linguistic Society.
- Nesselhauf, Nadja. 2014. From contraction to construction? The recent life of *'ll*. In Marianne Hundt (ed.), *Late Modern English syntax*, 77–89. Cambridge: Cambridge University Press.
- Petré, Peter & Lynn Anthonissen (eds.). 2020. Constructionist approaches to individuality in language. [Special Issue]. *Cognitive Linguistics* 31(2). <https://doi.org/10.1515/cog-2020-frontmatter2>.
- R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at: <http://www.R-project.org>.

- Scheibman, Joanne. 2000. *I dunno*: A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics* 32(1). 105–124.
- Schmid, Hans-Jörg. 2015. A blueprint of the Entrenchment-and-Conventionalization Model. *Yearbook of the German Cognitive Linguistics Association* 3(1). 3–26.
- Schmid, Hans-Jörg. 2017. A framework for understanding linguistic entrenchment and its psychological foundations. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 9–35. Berlin: De Gruyter.
- Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*. Oxford: Oxford University Press.
- Schmid, Hans-Jörg & Annette Mantlik. 2015. Entrenchment in historical corpora? Reconstructing dead authors' minds from their usage profiles. *Anglia* 133(4). 583–623.
- Schmid, Hans-Jörg, Quirin Würschinger, Sebastian Fischer & Helmut Küchenhoff. 2021. *That's cool*. Computational sociolinguistic methods for investigating individual lexico-grammatical variation. *Frontiers in Artificial Intelligence* 3. 547531.
- Schmidtke-Bode, Karsten. 2009. *Going to V and gonna V* in child language: A quantitative approach to constructional development. *Cognitive Linguistics* 20(3). 509–538.
- Stefanowitsch, Anatol. 2006. Distinctive collexeme analysis and diachrony: A comment. *Corpus Linguistics and Linguistic Theory* 2(2). 257–262.
- Stefanowitsch, Anatol & Susanne Flach. 2017. The corpus-based perspective on entrenchment. In Hans-Jörg Schmid (ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge*, 101–127. Berlin: De Gruyter.
- Stefanowitsch, Anatol & Susanne Flach. 2020. Too big to fail but big enough to pay for their mistakes: A collocation analysis of the patterns [too AD] to V] and [AD] enough to V]. In Gloria Corpas Pastor & Jean-Pierre Colson (eds.), *Computational phraseology*, 248–272. Amsterdam: John Benjamins.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2005. Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1(1). 1–43.
- von Eye, Alexander. 1990. *Introduction to configural frequency analysis: The search for types and antitypes in cross-classification*. Cambridge: Cambridge University Press.