

I can't say it's just the verbs: using distributional and experimental data to predict speakers' choices of negative modal contractions over their full forms

Robert Dausgs, Kiel University, daugs@anglistik.uni-kiel.de

David Lorenz, University of Rostock, david.lorenz2@uni-rostock.de

A growing body of research indicates that several contractions in present-day English (e.g. *gonna*, *'d*, *I dunno*) are past the stage of being mere online phonetic reductions of their corresponding full-forms (here BE *going to*, *would*, *I do not know*) and rather constitute emancipated constructions that take on a life of their own (Bresnan 2021; Dausgs 2021; Krug 2000; Lorenz 2013; Nesselhauf 2014; Scheibman 2000; Schmidtke-Bode 2009; to name a few). This view largely rests on distributional data obtained from corpora, showing that contractions, like their full forms, have syntactic, collocational, and functional preferences.

The present investigation builds on this research, delving deeper into the collocational biases of the negative modal contractions *can't*, *won't*, and *shouldn't* using both corpus and experimental methods. Our pilot study, based on data obtained from the Corpus of Contemporary American English (COCA; Davis 2008) and a g-maze task (Forster et al. 2009), revealed that speakers are indeed sensitive to the verb infinitives with which negative modal contractions co-occur, as reaction times for distinctive collexemes were lower than for repelled ones. By contrast, the verbal collexemes identified for the full forms did not exhibit such a distinctiveness effect.

A possible, previously neglected confounding factor are preferences regarding the syntactic subject. Considering established correlations between specific modal meanings and specific subject properties (cf. Coates 1983), we updated the corpus data analysis by utilizing distinctive covarying collexeme analysis (DCCA; Stefanowitsch & Flach 2020), which allows one to compare constructions with multiple slots: in this case, SUBJ [*can|will|should*] *not* V vs SUBJ [*ca|wo|should*]*n't* V.

The results indicate that the verbal collexemes identified in the pilot study are essentially only distinctive in combination with specific subjects. To illustrate, *allow* ranks among the top verbal collexemes of *cannot* in general, but according to the DCCA, it is only distinctive in combination with the subject *we*, whereas *they* is actually repelled by *cannot allow*. Similar findings have been reported for English modal enclitics (cf. Dausgs 2022). To measure the true distinctiveness effect in the contraction/full-form alternation, we therefore argue that trigrams (e.g. *we cannot allow*, *you won't believe*) should be preferred over bigrams (e.g. *cannot allow*, *won't believe*). An experimental study with the design adjusted accordingly is currently in progress, testing the effect of trigram frequencies on the recognition of contractions and full forms. This should account more accurately for the different forms' syntagmatic and paradigmatic associations speakers have routinized (cf. Schmid 2020).

References

- Bresnan, Joan. 2021. Formal grammar, usage probabilities, and auxiliary contraction. *Language* 97(1). 108–150. doi:[10.1353/lan.2021.0003](https://doi.org/10.1353/lan.2021.0003).
- 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*, 12–52. Amsterdam: John Benjamins. doi:[10.1075/cal.32.02dau](https://doi.org/10.1075/cal.32.02dau).
- Daug, Robert. 2022. English modal enclitic constructions: a diachronic, usage-based study of 'd and 'll. *Cognitive Linguistics* 33(1). 221–250. doi:[10.1515/cog-2021-0023](https://doi.org/10.1515/cog-2021-0023).
- Davies, Mark. 2008. *The Corpus of Contemporary American English (COCA)*. Available online at <https://www.english-corpora.org/coca/>.
- Forster, Kenneth, Christine Guerrero & Lisa Elliot. 2009. The maze task: Measuring forced incremental sentence processing time. *Behavior Research Methods* 41(1). 163–171.
- Lorenz, David. 2013. *Contractions of English semi-modals: The emancipating effect of frequency*. Freiburg: NIHIN / Universitätsbibliothek Freiburg.
- 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. doi:[10.1017/CBO9781139507226.007](https://doi.org/10.1017/CBO9781139507226.007).
- 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. doi:[10.1016/S0378-2166\(99\)00032-6](https://doi.org/10.1016/S0378-2166(99)00032-6).
- Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*. Oxford: Oxford University Press. doi:[10.1093/oso/9780198814771.001.0001](https://doi.org/10.1093/oso/9780198814771.001.0001).
- Schmidtke-Bode, Karsten. 2009. *Going to V and gonna V in child language: A quantitative approach to constructional development*. *Cognitive Linguistics* 20(3). doi:[10.1515/COGL.2009.023](https://doi.org/10.1515/COGL.2009.023).
- Stefanowitsch, Anatol & Susanne Flach. 2020. Too big to fail but big enough to pay for their mistakes: A collostructional analysis of the patterns [too ADJ to V] and [ADJ enough to V]. In Gloria Corpas Pastor & Jean-Pierre Colson (eds.), *Computational phraseology*, 248–272. Amsterdam: John Benjamins. doi:[10.1075/ivitra.24.13ste](https://doi.org/10.1075/ivitra.24.13ste).