

A Database of North American Multiple Modals from YouTube

Steven Coats

University of Oulu

E-mail: steven.coats@oulu.fi

Abstract

The use of multiple modal auxiliary verbs in a single verbal phrase is a feature of dialectal spoken English that has attracted substantial research attention, but due to its rareness, most studies of its syntax and geographical distribution have relied on non-naturalistic data with a limited geographical scope. This paper presents a database of sequences of two modal verbs that have been manually annotated as either authentic double modals (i.e. two modal verbs in a single verbal phrase) or as instances of correction/self-repair in which a speaker substitutes one modal form for another within a single speech turn (https://stcoats.github.io/dm_table1.html). The data, from the Corpus of North American Spoken English (CoNASE, Coats 2021), is linked to the corresponding video at the time of utterance. The resource attests new locations and new combinations of double modals and may prove to be useful for research into the syntax and geographical distribution of an emblematic North American English dialect feature.

In addition, the database may be suitable for corpus-based approaches to pragmatics, for example by providing data for the study of corrections or self-repair phenomena in speech. Self-repairs have been seen in the field of Conversation Analysis as a device for speaker turn regularity maintenance or local meaning negotiation (Fox et al., 1996; Schegloff et al., 1977), but have mainly been researched on the basis of small numbers of conversations. In this data, self-repairs could be considered from a quantitative perspective in terms of (for example) geographic distributions or the discourse content in which they occur. In addition, the methods used to create the database can be implemented for the creation of time-aligned multimodal corpora from publicly accessible social media videos suitable for the investigation of a wide range of syntactic, lexical, pragmatic, and interactional phenomena.

Keywords: dialectology, modal verbs, corpus linguistics, repair, conversation analysis

References

- Coats, S. (2021). Corpus of North American Spoken English (CoNASE). <http://cc.oulu.fi/~scoats/CoNASE.html>
- Fox, B. A., Hayashi, M. & Jaspersen R. (1996). Resources and repair: A cross-linguistic study of syntax and repair. In E. Ochs, E. A. Schegloff & S. A. Thompson (Eds.), *Interaction and Grammar*, pp. 185–237. Cambridge: Cambridge University Press.
- Schegloff, E. A., Jefferson, G. & Sacks H. (1977). The preference for self-correction in the organization of repair in conversation. *Language* 53, pp. 361–382.