

***'She said {that/Ø} she couldn't take a complement':
Complementizer that omission in American English***

Isaac Bleaman, Daniel Duncan, Shelley Feuer, Gregory Guy
Zachary Jagers, and Matthew Stuck (New York University)

Complementizer *that* is well-known to occur variably in English: complement clauses can be produced with or without overt *that*: e.g. “Mary said {that/Ø} she was coming tonight.” This variation has some social significance: Kroch & Small 1978 find more overt *that* used by the hosts of a radio program than by callers. Other variationist studies (e.g. Tagliamonte & Smith 2005, Torres Cacoullos & Walker 2009) have also identified significant linguistic constraints on this variable, including features of discourse structure, syntactic complexity, and properties of the matrix verb. These studies offer conflicting interpretations of these constraints: K&S argue that grammatical ideology favors more explicit syntax, while T&S and TC&W appeal to ongoing grammaticalization of certain collocations (e.g. *I think*) as discourse markers. Several studies mention lexical frequency of the matrix verb as a constraint, while K&S also use an etymological classification, contrasting Germanic and Latinate verbs. The present study seeks to compare and test several alternative analyses of these constraints.

Data are drawn from sociolinguistic interviews with 16 speakers from the Buckeye corpus, divided evenly among male and female speakers, in two age groups (older and younger, as grouped in the corpus.) All tokens (n>1200) of relevant complement clause constructions were extracted. Multivariate analyses of the data were conducted using Goldvarb. Considerable individual variability is found, with speakers ranging from 47% to 92% deletion rates, but this variation did not systematically correlate with age or gender in our corpus, although there is evidence of stylistic correlates.

Matrix verb effects were coded three ways: by lexical root (with lower frequency items clustered by etymology), lexical frequency, and root length (in syllables). Of these, the best predictor is lexical item/etymology, with *think* and *mean* strongly favoring Ø, and Germanic verbs generally favoring more omission than Latinate verbs. Monosyllabic verbs also favor deletion over polysyllables. Lexical frequency, although a significant predictor on its own, does not improve the model when these other two factor groups are included. First person singular subjects also favor deletion, consistent with the hypothesis of the grammaticalization of collocations like *I think*.

Factors of grammatical complexity are also significant. Overt complementizers are favored when the complementizer position is separated from the verb by syntactic elements, and complex verb tenses with an extra lexeme (auxiliaries or *do*-support) favor *that* over simple verb tenses (present and past). Similarly, *that* is favored when matrix and complement clauses have non-coreferential subjects. Such results suggest that overt *that* provides a processing cue to sentence structure which is more often employed in utterances that are more complex or less canonical; this is consistent with Rohdenburg's (1996) claims about grammatical explicitness and cognitive complexity.

The results indicate that, although many instances of zero complementizer may be phrasal forms with discursive implications, there are nonetheless general syntactic and processing constraints at play governing complementizer usage in English.

References

- Kroch, Anthony, and Cathy Small. 1978. Grammatical ideology and its effect on speech. In *Linguistic variation: Models and methods*, ed. David Sankoff, 44–55. New York: Academic Press.
- Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2):149-182.
- Tagliamonte, Sali, and Jennifer Smith. 2005. No momentary fancy! The zero ‘complementizer’ in English dialects. *English Language and Linguistics* 9:289–309.
- Torres Cacoullos, Rena, and James Walker, 2009. On the persistence of grammar in discourse formulas: a variationist study of *that*. *Linguistics* 47:1-43.