The typology of veridicality inferences

Aaron Steven White
University of Rochester

Kyle Rawlins
Johns Hopkins University

Introduction. The semantics literature has long been concerned with the complex array of veridicality inferences that clause-embedding predicates trigger about their complements (Kiparsky & Kiparsky, 1970; Karttunen, 1971a,b; Horn, 1972; Karttunen & Peters, 1979; Heim, 1992; Simons, 2001, 2007; Simons et al., 2010; Abusch, 2002, 2010; Gajewski, 2007; Egré, 2008; Anand & Hacquard, 2013, 2014; Spector & Egré, 2015; Baglini & Francez, 2016; Nadathur, 2016). Yet, we still know relatively little about how the structure of a predicate's clausal arguments modulates these inferences. Consider:

(1) a. Jo **remembered** that *Bo left*.
b. Jo didn't **remember** that *Bo left*.

(2) a. Jo **remembered** *Bo to have leave*.
b. Jo didn't **remember** *Bo to have leave*.

(3) a. Bo **remembered** ~~*Bo*~~ *to leave*.
b. Bo didn't **remember** ~~*Bo*~~ *to leave*.

(4) a. Bo left.
b. Bo didn't leave.

Why should (1a), (1b), and (3a) trigger (4a) while (3b) triggers (4b) and (2a) and (2b) trigger neither (see van Leusen, 2012; White, 2014)? And why does *hope* not show a similar inference pattern?

(5) a. Jo **hoped** that *Bo left*.
b. Jo didn't **hope** that *Bo left*.

(6) a. Bo **hoped** ~~*Bo*~~ *to leave*.
b. Bo didn't **hope** ~~*Bo*~~ *to leave*.

A major challenge for progress in this domain is that we do not know the full range of inference patterns that are extant across lexical items and complement structures. Surmounting this challenge is necessary for constructing a general theory of how veridicality inferences arise from linguistic representations, and it may furthermore help uncover the semantics of embedding more generally.

Contribution. We approach this challenge by extending White & Rawlins's (2018) MegaVeridicality dataset – which contains veridicality judgments for all English clause-embedding verbs with finite complements – to all English clause-embedding verbs with a variety of infinitival complements. We use these data to ask (i) which inference patterns are extant across verbs and complement structures; and (ii) which underlying lexical properties determine veridicality inferences in different syntactic contexts. We find twelve inference patterns in our data, which appear to be driven by two interacting axes of variation: whether the verb takes an internal DP argument and what kind of modality it involves.

Experiment. Following White & Rawlins's (2018) method for finite complements (7), we selected 603 verbs from the publicly available MegaAttitude dataset (White & Rawlins, 2016) based on those verbs' acceptability with various infinitival complements (8) – on average, a rating of 4 out of 7 or better.

(7) a. Someone thought that something happened. [NP _ed that S]
b. Someone was told that something happened. [NP was _ed that S]

(8) a. Someone needed for something to happen. [NP _ed for NP to VP]
b. Someone wanted someone to {do, have} something. [NP _ed NP to VP[+/-EV]]
c. Someone was overjoyed to {do, have} something. [NP was _ed to VP[+/-EV]]
d. Someone managed to {do, have} something. [NP _ed to VP[+/-EV]]

We manipulated eventivity of the embedded predicate ([+/EV]) in order to capture any effects of infinitival tense/aspect (see Stowell, 1982; Pesetsky, 1991; Bošković, 1996, 1997; Martin, 1996, 2001; Grano, 2012; Wurmbrand, 2014) on veridicality inferences (see White, 2014).

This yielded the sample sizes in the table to the right. These verbs were then slotted into contexts like those in (8) except that the underlined *someone*s and *something*s were replaced with *a particular person* and *a particular thing*, as in (9), and matrix negation was manipulated.

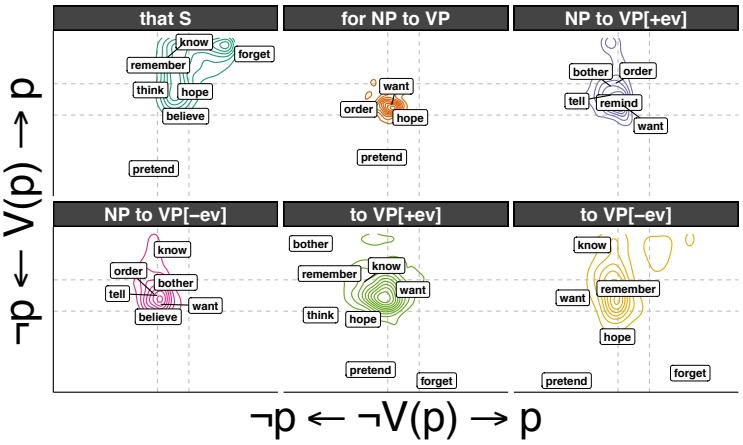
Frame	# verbs
NP _ed that S	375
NP was _ed that S	169
NP _ed for NP to VP	184
NP _ed NP to VP[+EV]	197
NP _ed NP to VP[-EV]	128
NP was _ed to VP[+EV]	278
NP was _ed to VP[-EV]	256
NP _ed to VP[+EV]	217
NP _ed to VP[-EV]	165

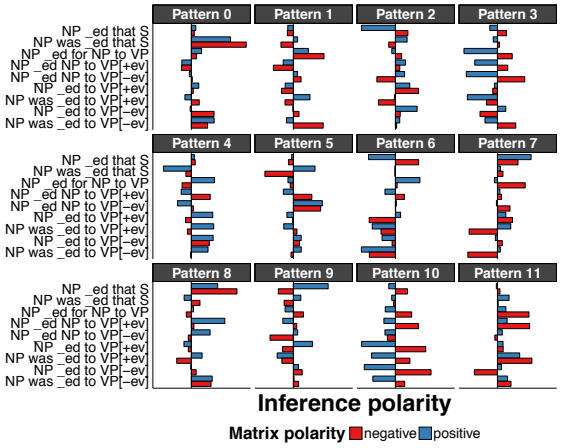
(9) Someone {needed, didn't need} for a particular thing to happen.

This yielded 2,850 items, which we randomly partitioned into 50 lists.

Given sentences such as those in (9), the task was to answer the question *did that person do that thing?*, *did that person have that thing?*, or *did that thing happen?* (depending on the sentence at hand). There were three response options: *yes*, *maybe or maybe not*, and *no* (cf. Karttunen et al. 2014). We recruited participants through Amazon's Mechanical Turk, gathering 10 ratings per item given by 10 different participants. To control for differences in participants' use of the response scale, we applied a ordinal mixed model-based normalization to their responses (following White & Rawlins 2018).

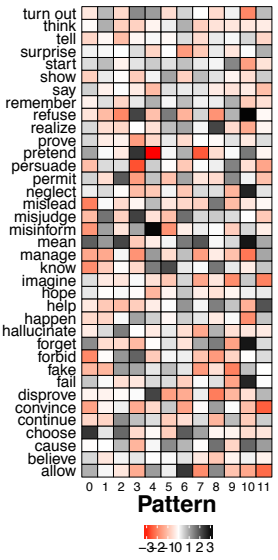
Results. The figure to the right plots the distribution of normalized responses for contexts with negative matrix polarity (x-axis) against those for contexts with positive matrix polarity (y-axis). Each facet shows a complement type (collapsing across the passivization manipulation), and the dotted lines show where the majority response changes, averaging across participants. (The *that S* facet plots White & Rawlins's data.) Example verbs are those found in the active frames only. A verb being toward the top means that it entails its embedded clause, while a verb being towards the bottom means that it entails the negation of its embedded clause. A verb being toward the right means that its embedded clause content projects through negation, while a verb being towards the left means that the negation of its embedded clause content projects through negation. We see that our experiment successfully distinguishes between factives (upper right), antifactives (lower left), veridicals (upper middle), implicatives (upper left), antiimplicatives (lower right), and nonveridicals (middle).





Analysis. We break our analysis into two parts: (i) discovery of the extant inference patterns in our dataset; and (ii) discovery of the properties that that drive those patterns. To carry out the first part, we fit a factor analysis with varying numbers of latent factors to the normalized data. Each latent factor is associated with a particular inference pattern (left plot), and each verb is associated with each latent factor to different extents (below plot). (In a cross-validation based procedure, we find 12 to be optimal.) As an example, PATTERN 10 shows a pattern wherein matrix negation inverts the polarity of an inference – i.e. an implicative pattern. This is reflected in the verb plot by the fact that *fail* loads positively on this pattern, and *manage* loads negatively.

To discover properties that drive these inference patterns, we conduct Canonical Correlation Analysis between the verb representation (right plot) and the distributional data present in the MegaAttitude dataset (not shown). The idea here is to discover generalizations about the veridicality patterns that correspond as closely as possible to syntactic distributions, thus ensuring that these generalization are plausibly linguistically represented. Though we cannot cover all findings for reasons of space, we note three generalizations from this analysis: (i) taking experiencers (*surprise*, *delight*, and *shock*) is the most important factor in determining veridicality; (ii) future-oriented predicates give rise to different inference patterns based on whether they involve epistemic (*expect*), bouletic (*want*), or teleological (*intend*) modality; and (iii) predicates with lexical negation give rise to different inference patterns based on whether they involve epistemic (*misjudge*, *mislead*) or root modality (*refuse*, *forbid*).



Conclusion. We extended White & Rawlins's (2018) MegaVeridicality dataset to all English clause-embedding verbs with a variety of infinitival complements and used these data to ask (i) which inference patterns are extant across verbs and complement structures; and (ii) which underlying lexical properties determine veridicality inferences.