The role of veridicality and factivity in clause selection*

Aaron Steven White & Kyle Rawlins

University of Rochester, Johns Hopkins University

1. Introduction

This paper investigates the relationship between clausal selection and two lexical semantic properties that have been claimed to be important in determining such selection: *factivity* (Hintikka 1975) and *veridicality* (Egré 2008). Specifically, these properties have been suggested to be connected to *responsivity*: selection of both declarative and interrogative complements. Our investigation builds on the MegaAttitude dataset of White & Rawlins (2016), which contains experimentally-collected acceptability judgments for effectively every English verb that embeds clauses, in a large set of frames. Based on these acceptability judgments, we collect a new dataset of *veridicality judgments* for all English verbs that embed declarative clauses.[1] This allows a systematic, large-scale comparison of veridicality judgments and selectional patterns. We show that factivity and veridicality do not correlate with responsivity when considering the entirety of the lexicon but that they do correlate with other selectional patterns, in particular, selection of DP direct and indirect objects.

We begin in §2 by introducing specific hypotheses and assumptions about factivity, veridicality, and responsivity. We then introduce our methods and resulting dataset in §3-4, and in §5 develop both confirmatory analyses to test the hypotheses introduced in §2, and exploratory analyses to discover new relationships among the variables in our datasets.

2. Background and hypotheses

2.1 Factivity and veridicality

Veridicality and factivity are closely related lexical semantic properties that concern the inferences one can draw about the truth of a predicate's embedded clause. Specifically, a

*We are grateful to audiences at JHU, UMD, the University of Rochester, ESSLLI 2017, and NELS 48, as well as Rachel Dudley and Valentine Hacquard, Rachel Rudinger, and Ben Van Durme. This work was partly funded by NSF INSPIRE BCS-1344269 (*Gradient symbolic computation*), the JHU Science of Learning Institute, and the DARPA AIDA program (*Multi-modal Decompositional Semantics*).

[1]The MegaVeridicality dataset is available at `http://megattitude.io`.

verb V is veridical iff DP V S entails S. Thus, verbs like *know* and *prove* are both veridical because, from sentences like those in (1), we infer that Bo left (Karttunen 1971a, Egré 2008, Karttunen 2012, Spector & Egré 2015).[2]

(1) *Cartoon scenario*: Jo is watching a cartoon with an animal named Bo; it is hard to tell what kind of animal Bo is, though the show's writers intend a particular animal.
   a. Jo knew that Bo was a cat. $\rightsquigarrow$ Bo was a cat.
   b. Jo proved that Bo was a cat. $\rightsquigarrow$ Bo was a cat.

The classic first approximation is that a verb V is factive iff it is veridical and DP V S presupposes S – i.e. presupposes the truth of its declarative complement (Kiparsky & Kiparsky 1970, Karttunen 1971b *et seq*). This can be operationalized using standard projection diagnostics for presupposition – e.g. whether the inference shows up when embedded under negation (our focus in this paper), questioning, or the antecedent of conditionals. A factive verb like *know* still implies the content clause even when negated. However, *know* contrasts with *prove*, which is neutral about the truth of the complement under negation. We therefore call *prove* veridical but not factive, and *know* both veridical and factive.

(2) a. Jo didn't know that Bo was a cat. $\rightsquigarrow$ Bo was a cat.
   b. Jo didn't prove that Bo was a cat. $\not\rightsquigarrow$ Bo was a cat.

It is important to note that the category of factives is not so clean as this baseline description suggests (Karttunen 1971b). Many verbs that appear factive at first glance have variable behavior depending on the diagnostics: they are what Karttunen terms *semi-factives*. For example, projection in conditionals varies by verb and context.

(3) a. If I {realize, find out} later that I had not told the truth, I will confess.
   $\not\rightsquigarrow$ I didn't tell the truth.
   b. If I {regret, am bothered} later that I had not told the truth, I will confess.
   $\rightsquigarrow$ I didn't tell the truth.

While the messy reality of semi-factivity has not so far been important in the proposed generalizations about responsivity we discuss below, we consider the possibility that responsivity may interact with semi-factivity, aiming to explicitly measure it in our experiments.

2.2 Responsivity

We focus here on the selectional pattern of clause-embedding verbs that Lahiri (2002) terms *responsivity* (see also Karttunen 1977, Groenendijk & Stokhof 1984, Ginzburg 1995 *et seq*). A verb is *responsive* if it takes both interrogative and declarative complements.

[2] In the exact form we use it, the term *veridical* comes from Spector & Egré 2015, fn. 7, where it is based on a slightly different use from Lahiri 2002, *veridical-responsive*. Lahiri (2002) uses this term primarily for verbs in combination with interrogative clauses, for cases where some relationship with the true answer is entailed. For example, *tell* is veridical-responsive in that *tell-wh* entails that the speaker told the truth.

(i) Jo told me whether Bo was a cat. $\rightsquigarrow$ if Bo is a cat, Jo told me so; and if Bo isn't a cat, Jo told me so.

(4) a. Jo knew that Bo was a cat.
b. Jo knew whether Bo was a cat.

There is a large set of responsive verbs (to be explored further) including many cognitive factives (e.g. *know*, *find out*, *discover*), emotive factives (e.g. *shock*, *surprise*, amaze)[3], and communication predicates (e.g. *tell*, *report*, *advertise*). There are of course many non-responsive verbs as well, and Hintikka (1975) originally introduced the puzzle about what underlies responsivity by contrasting *know* with *believe*. (We will be concerned here only with non-responsives that do embed declarative clauses.)

(5) a. Jo believed that Bo was a cat.
b. *Jo believed whether Bo was a cat.

Understanding what, if anything, the generalization is concerning which predicates are responsive has been an important project in the linguistics literature on question-embedding, intertwined with the analysis of responsivity. That is, is there a single core entry for *know* that can handle both of these complement types, in combination with specific compositional assumptions? And if so, can such an analysis generalize across this large set of verbs?

A central observation is that there seems to be some relationship between responsivity and other semantic properties to do with whether the truth of a declarative complement is implied by the verb. In particular, Hintikka (1975) observed that many factive verbs are responsive, suggesting (though not with this terminology) that the lack of factivity provides an explanation for why *believe* does not take questions.[4] This idea has been pursued or adopted by much subsequent literature, though it has equally often been rejected on the basis of obvious counterexamples (see Berman 1991, Ginzburg 1995, Beck & Rullmann 1999, Lahiri 2002, Egré 2008, George 2011, Uegaki 2012, 2015, Theiler 2014, Spector & Egré 2015 for discussion). Relaxations of factivity as the generalizing factor have also been considered: in particular, what we have termed veridicality (Egré 2008).

The main reason for rejecting either hypothesis, promising though they initially seem, is that there are counterexamples, and the most-discussed counterexamples are of course communication verbs. Though communication verbs are veridical in Lahiri's sense when they embed interrogatives, they do not at first glance appear to be either factive or veridical with declarative complements.

(6) Jo told me that Bo was a cat. $\not\rightsquigarrow$ Bo was a cat.

Spector & Egré (2015) argue against this quick judgment, however, providing examples to show that at least some communication verbs appear to be factive in certain contexts (see also Schlenker 2010, Anand & Hacquard 2014).

[3]With the caveat that, in standard descriptions, emotive factives do not take plain polar interrogative complements (Karttunen 1977; see Cremers & Chemla 2017 for a recent overview).

[4]While, in initial discussion, Hintikka implies that the relationship among attitude verbs is factivity iff responsivity, which is often how his proposal is described, he also provides one early counterexample that would seem to defeat the account of *believe*: "It is particularly instructive to see that some verbs, e.g. 'guess', which normally do not have success presuppositions acquire one when they occur in a wh-construction." Hintikka does not consider communication verbs, presumably taking them not to be attitude verbs at all.

(7) Did Sue tell anyone that she is pregnant? $\rightsquigarrow$ Sue is pregnant. (S&E ex. 29)

They suggest that from this one should conclude that such verbs are ambiguous between factive and non-factive variants, and that it is the factive variant that their responsivity rests on. This provides a rather elegant explanation for why communication verbs are veridical-responsive in interrogative frames as well.[5]

There are also counterexamples in the realm of cognitive verbs, and the ambiguous-factivity hypothesis doesn't fare so well with these cases. One well-explored example is the class of selection predicates, such as *decide* (Egré 2008, White & Rawlins 2017). These tend to be responsive, not veridical, and is not so easy to fit them into the veridical-responsive category in its interrogative frame.

(8) a. Jo decided that Bo was a cat. $\not\rightsquigarrow$ Bo was a cat.
b. Jo decided whether Bo was a cat. $\not\rightsquigarrow$ Jo's decision was correct.[6]

Building on a substantial literature connecting the content of attitude reports to the event structure of the attitude (Kratzer 2006, Hacquard 2006, 2009, Moulton 2009, Anand & Hacquard 2013, 2014, Rawlins 2013, White 2014, Bogal-Allbritten 2016), White & Rawlins (2017) suggest that the explanation for responsivity here follows from the event structure. Specifically, *decide* (along with other selection predicates) is a change-of-state predicate involving a pre-state, where the content is inquisitive, and a post-state where the content is not inquisitive (a decision has been made).[7] The prediction here is that responsivity will correlate with whatever verb properties are involved in event structure, such as aspectual properties like telicity. On this sort of hypothesis, there is no expectation of a relationship between responsivity and factivity or veridicality.

2.3 Hypotheses

This leads to two competing, unsettled hypotheses, which we state here in a strong form.

(9) a. *Factivity*: a verb is responsive iff it is factive.
b. *Veridicality*: a verb is responsive iff it is veridical.

Though these hypotheses have been explored with relatively small sets of verbs, they are hypotheses about the lexicon as a whole, and are therefore good candidates for testing as such (White & Rawlins 2016). We quantitatively test the hypotheses in (9) by evaluating the correlation between a measure of responsivity and measures of factivity and veridicality, described below. We also consider weaker versions of these hypotheses, and in an ex-

[5]One further important reason why such a generalization would be important is that it provides a potentially easy cue for a child faced with the challenge of learning which verbs are factive, given that direct cues to learn factivity are not present in the data (Dudley et al. 2017); see Dudley (2017) for further discussion.

[6]Unless Jo has some authority to make this decision, e.g. is the writer for the show.

[7]Beyond this, there is a wide range of possible accounts for selectional constraints that interact with the putative generalization that we will not directly consider here. To take one example, returning to Hintikka's original claim about *believe*, Theiler et al. (2017) propose that this verb (and many others) cannot take interrogatives (they are *anti-rogative*) because they are neg-raising, leading to a further, potentially distinct explanation for at least part of the data.

ploratory analysis we ask whether there are any previously unnoticed relationships among these properties and verbs' syntactic distributions – e.g. Spector & Egré's (2015) suggestion that factivity might not be easy to detect in communication verbs, but that among non-communication verbs we should find a correlation between factivity and responsivity.

3. Data

To quantitatively test the hypotheses in (9), we evaluate the correlation between two measures: (i) verbs' acceptability with both declarative and interrogative complements – i.e. a measure of syntactic distribution – and (ii) those same verbs' factivity and veridicality. For our measure of syntactic distribution, we use the pre-existing MegaAttitude dataset of White & Rawlins (2016). For our measure of veridicality and factivity, we collect a new dataset – the MegaVeridicality dataset – covering all verbs that embed declarative clauses.

3.1 A measure of syntactic distribution

The MegaAttitude dataset of White & Rawlins (2016) contains ordinal acceptability judgments for 1,000 English clause-embedding verbs – effectively all clause-embedding verbs in English – in 50 different syntactic frames. White & Rawlins's frames instantiate configurations of syntactic features that have been suggested to be correlated with various aspects of clause-embedding verb semantics: COMPLEMENTIZER, EMBEDDED TENSE, MATRIX OBJECT, MATRIX PP, EMBEDDED SUBJECT, and PASSIVIZATION.

An important aspect of White & Rawlins's methodology, which we carry forward into our new dataset, is the way in which they construct items. In order to avoid typicality effects and thus ensure that participants' judgments are only about a clause-embedding verb's compatibility with the syntactic structure of a particular complement, items in MegaAttitude are constructed using a *low lexical content* format, wherein all DP arguments are indefinite pronouns and all verbs besides the one being tested are *do*, *have*, or *happen*.

For the purposes of testing the hypotheses in (9), we focus on the acceptability of different verbs in six of these low lexical content frames, exemplified in (10).

(10) a. Someone knew that something happened. [DP ___ed that S]
b. Someone knew whether something happened. [DP ___ed whether S]
c. Someone knew which thing happened. [DP ___ed which thing S]
d. Someone was told that something happened. [DP was ___ed that S]
e. Someone was told whether something happened. [DP was ___ed whether S]
f. Someone was told which thing happened. [DP was ___ed which thing S]

The reason for including the latter three frames, which involve a passivized direct object, is to capture verbs that take a direct object in addition to a clausal complement, regardless of whether they take a referential external argument, like *tell* and *promise*, or an expletive external argument, like *surprise* and *bother*.

3.2 Two measures of veridicality and factivity

For our measures of veridicality and factivity, we construct a new dataset of veridicality judgments for all English verbs that take *that*-clause complements. We use the MegaAttitude dataset to select 517 verbs from MegaAttitude based on their acceptability in the [DP __ed that S] and [DP was __ed that S] frames, (10) (a) and (d) above. A verb is included in this set if it received an average rating of 4 out of 7 or better after normalizing participants' acceptability judgments for differences in ordinal scale use (see below for the normalization procedure). This yields a sample of 348 verbs that only occur in the active frame, 142 that only occur in the passive frame, and 27 that occur in both.

To construct items, we use White & Rawlins's (2016) low content method. We implement this by slotting verbs into contexts illustrated in (11) based on whether they were acceptable in the active, passive, or both. This yields 1,088 items, which were randomly partitioned into 16 lists of 68. Given sentences such as those in (11), the task was to answer the question (12) using one of three response options: *yes*, *maybe or maybe not*, and *no* (see Karttunen et al. 2014 for a related method).

(11) a. Someone {thought, didn't think} that a particular thing happened.
b. Someone {was, wasn't} told that a particular thing happened.

(12) Did that thing happen?

In constructing the low-content items, we used *a particular thing* rather than the more bleached *something* in order to avoid extremely non-specific readings for the embedded subject, allowing reference back to the subject in the following question.

In addition to the above task (Task 1), we deploy a slightly modified task (Task 2) that uses the same response options but places all items (11) tested in Task 1 within the antecedent of a conditional question whose consequent is (12).

(13) a. If someone didn't know that a particular thing happened, did that thing happen?
b. If someone wasn't told that a particular thing happened, did that thing happen?

The aim of Task 2 is to distinguish semi-factives, like *know* and *discover*, from true factives, like *love* and *hate*, by inducing local accommodation within the conditional antecedent for the semi-factives, since conditional antecedents are a common environment to find non-projection behavior. We demonstrate the relative success of this aim – at least for presuppositions that normally project through negation – in §4.

As part of both tasks, we also collect acceptability judgments on a seven-point ordinal scale for every item. (In Task 2, the item itself was bolded, and participants were asked to judge the bolded sentences.) These judgments were collected because, though we have estimates of the acceptability for each item with positive matrix polarity, we do not for those items with negative matrix polarity, and we expected at least some interactions.

For each task, we recruited 160 unique participants through Amazon's Mechanical Turk, yielding 10 ratings per item per task, each given by 10 different participants.

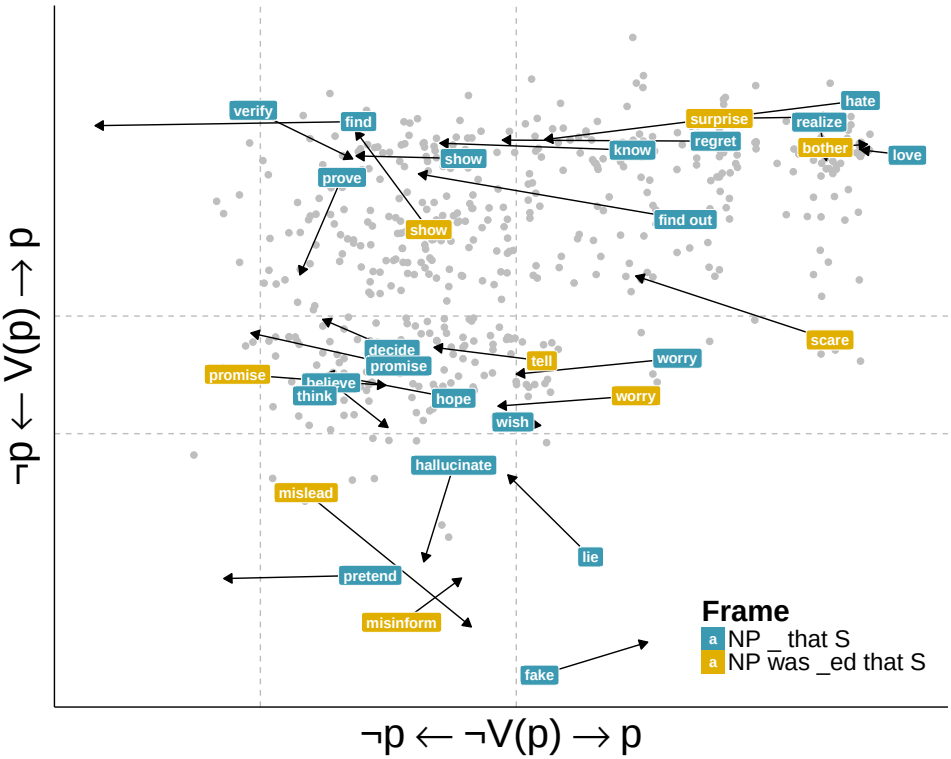
To control for differences in participants' use of the response scale, we apply an ordi-

nal mixed model-based normalization to both their veridicality responses and their acceptability responses prior to analysis.[8] For the veridicality responses, this procedure yields a normalized score for each item (verb-frame-polarity tuple), where more positive means more likely to entail the embedded clause content, and more negative means more likely to entail the negation of the embedded clause content. For the acceptability responses, a more positive score means more acceptable, and a more negative score means less acceptable.

4. Results

Figure (14) plots the normalized responses for contexts with negative matrix polarity (*x*-axis) against those for contexts with positive matrix polarity (*y*-axis). Each grey point corresponds to a verb, with labels on (a superset of) verbs mentioned in the paper. The placement of the points and labels corresponds to the normalized responses for verbs in Task 1. The arrows attached to each labeled verb point to the position for that verb in Task 2.

(14) *Normalized responses for contexts with negative matrix polarity (x-axis) against those for contexts with positive matrix polarity (y-axis)*



The dotted lines show where the majority response changes, averaging across participants. Verbs toward the top entail their embedded clause, while verbs toward the bottom entail the

[8] In particular, we fit a model with unconstrained cutpoints for each participant and a random intercept for each item (verb-frame-polarity tuple). The normalized rating for a particular item is then the Best Linear Unbiased Predictor for that item's random intercept. This method can be thought of as a generalization of *z*-scoring that better accounts for how participants make ordinal judgments (see White et al. 2018 for quantitative evidence in the domain of acceptability).

negation of their 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 by inspecting individual verbs that our experiment successfully distinguishes between standard cases of factives (upper right), non-factive veridicals (upper middle), nonveridicals (middle), and antiveridicals (lower middle). Taking the majority response boundaries as a guide, the vast majority of verb-frame pairs are nonveridical (115 verbs), non-factive veridical (177 verbs), or factive (199 verbs), with far fewer being antiveridical in either positive or negative frames. However, there are not necessarily clear dividing lines between these classes present in the data, suggesting that speaker's inferences about veridicality are generally quite gradient and likely influenced by the fine-grained semantics of particular verbs. We do appear to see a near-categorical *lack* of certain logically possible combinations for the verbs and frames in this study: very few verbs are antiveridical in either the positive and negative case – e.g. *hallucinate* – and very few verbs that are neutral in the positive frame but veridical in the negative frame – e.g. *worry*.[9]

These numbers shift substantially in Task 2, in both dimensions in Figure (14). Generally speaking, verbs that are factive by Task 1 become less so in Task 2, showing the overall pattern that we'd expect given observations about semi-factivity in the literature. Generally, factives with some emotive content are also less likely to shift, which would be expected from typical characterizations of them as hard triggers. For present purposes we will take Task 2 in comparison to Task 1 to provide a reasonable overall measure of semi-factivity, leaving a full evaluation and exploration of this sub-dataset for the future. The immediate questions that Task 2 factors into is whether the category of factives as discussed into the literature on responsivity needs to take semi-factivity into account.

5. Analysis

Having established the success of our veridicality tasks in capturing various forms of lexical inference, we now turn to an analysis of the relationship between the measures gathered in these tasks and the measures of syntactic distribution present in MegaAttitude. We begin with a confirmatory analysis that aims to test the hypotheses coming from prior work, listed in (9). We then discuss an exploratory analysis aimed at extracting further potential generalizations about the relationship between veridicality and syntactic distribution.

5.1 Confirmatory analysis

To test the hypotheses in (9), we construct four derived measures from MegaAttitude and MegaVeridicality for each verb.

(15) a. QUESTION-TAKING: the maximum normalized acceptability of [DP (was) ___ed whether S] and [DP (was) ___ed which thing S] from MegaAttitude

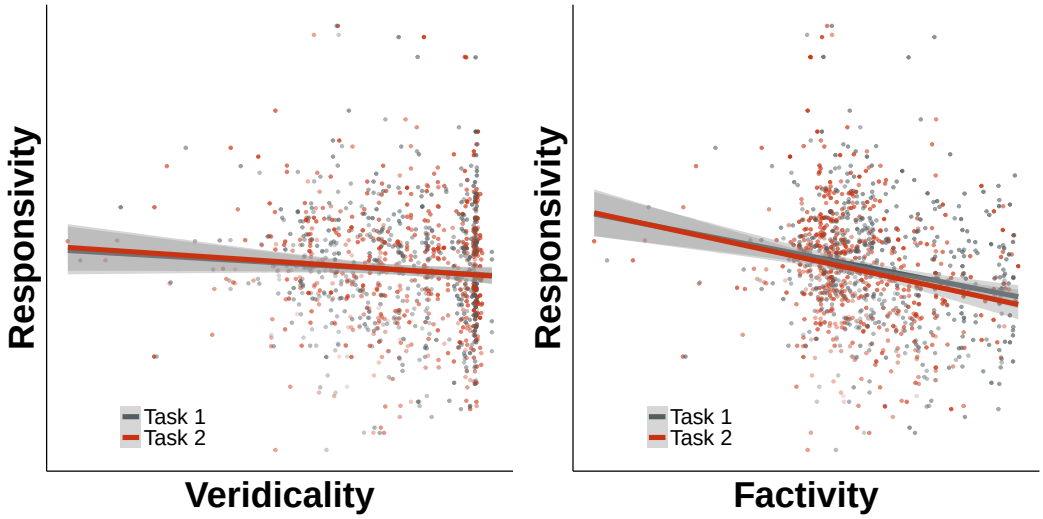
[9]This study did not generally include infinitival frames or implicatives such as *manage*, which we expect would instantiate some of the missing anti-veridical/negative cases. We leave these for future work.

b. RESPONSIVITY: the minimum of QUESTION-TAKING and the normalized acceptability of [DP (was) ___ed that S] from MegaAttitude
c. VERIDICALITY: the normalized veridicality for the positive matrix polarity frame from MegaVeridicality
d. FACTIVITY: the minimum of the normalized veridicality for positive matrix polarity frame and negative matrix polarity frame from MegaVeridicality

Because we have two veridicality judgment tasks, there are two versions of VERIDICALITY and FACTIVITY, which we report separately in our analyses.

Figure (16) plots RESPONSIVITY against the two VERIDICALITY measures (right) and the two FACTIVITY measures (right). The two tasks are distinguished by color – Task 1 in grey and Task 2 in red – and the lines give linear regression fits weighted by acceptability.[10]

(16) *Correlation of* RESPONSIVITY *with* VERIDICALITY *(left) and* FACTIVITY *(right)*



These plots provide a visual indicator of the fact that, rather than being positive, as we would expect under the hypotheses in (9), they are very near zero or perhaps negative.

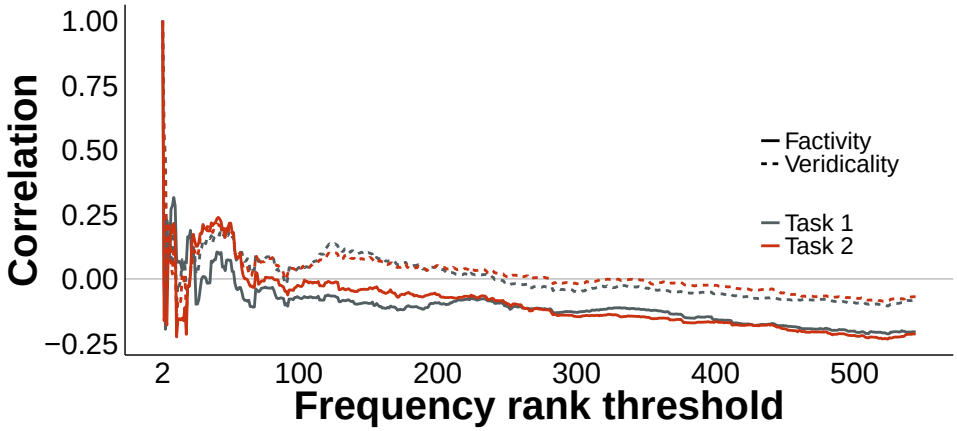
Further, we see a wide dispersion of points around the regression lines, suggesting that very little information about RESPONSIVITY is captured by either VERIDICALITY or FACTIVITY. This is corroborated by a linear regression with simple effects for VERIDICALITY and FACTIVITY from both tasks as well as all two-, three-, and four-way interactions and a loss weighted by acceptability. In a 10-fold cross-validation, we find that, this regression explains only 5% of the variance in RESPONSIVITY (95% CI = [-2%, 13%]), weighting by acceptability and averaging across folds.[11] This finding is problematic for proposals that rest on generalizations like those in (9), since these proposals should predict, if not perfect explanation of the variance in the data, that at least some nontrivial amount of this variance can be reliably explained.

[10]For veridicality, weighted by acceptability in the positive frame, and for factivity, weighted by the lower acceptability of the two for positive vs. negative frames.

[11]Using a linear regression with only the simple effects listed above, yields a 3% gain in the mean variance explained, so it is not likely that the low explained variance is only due to overfitting caused by the inclusion of higher-order interactions.

One question that arises here is why the hypotheses in (9) seemed so appealing in prior work on this topic. A potential explanation for the appearance of plausibility is that prior work only investigates a small number of the most frequent verbs in English: *know*, *think*, *believe*, *tell*, *say*, *see*, *hear*, etc. Figure (17) plots the correlation between responsivity and veridicality and factivity when looking at only the top *N* most frequent verbs.

(17) *Correlation between responsivity and veridicality and factivity when looking at only the top N most frequent verbs based on SUBTLEX-US (Brysbaert & New 2009)*



If one looks at only the most frequent clause-embedding verbs, the correlations between responsivity and veridicality and factivity look promising. But as we move past the 100 most frequent verbs, the correlations begin to waver around zero, sinking well into negative correlations. This may indicate that previous analyses overfit to a small number of frequent verbs that are not indicative of any more general lexical pattern. Such overfitting would be problematic, since in other domains of linguistic theorizing, we know that lower frequency items tend to be the most rule-governed, with higher frequency items tending to be more idiosyncratic (see O'Donnell 2015 for a recent review). Thus, the pattern observed in Figure (17) is especially worrying for veridicality/factivity-based hypotheses about responsivity because we see the opposite of what me might expect if these generalizations were in fact true – i.e. the correlations should only get *stronger* as the frequencies shrink.

5.2 Exploratory analysis

In light of our failure to find quantitative evidence of the generalizations in (9), we now ask whether there are weaker generalizations about the relationship between syntactic distribution and veridicality or factivity that can be made on the basis of these data. That such weaker generalizations might exist has been countenanced in prior work. For instance, though Egré (2008) aims to defend (9b), he suggests that, in the long run, it may be more tenable to back off to a weaker version – specifically: *a verb is responsive if it is veridical.*

This is an instance of a more general strategy wherein subsets of the data are sought for which some generalization might hold. It is important to tread carefully here, though, since such regimentation of the data risks the possibility of making spurious generalizations. With this in mind, we take an alternative, data-driven approach to constructing such generalizations using Canonical Correlation Analysis (CCA; Hotelling 1936).

CCA can be thought of as a method for synthesizing the information in two *views* – here, syntactic distributions and veridicality inferences – of some set of objects – here, verbs – in order to find latent features of those objects (*canonical variables*) that best explain both views. This implements the idea of searching for a regimentation – the canonical variables – that simultaneously explains the entire dataset as well as possible. The canonical variables are furthermore ordered (from most to least) by the amount of information they simultaneously explain about the distributional properties of both views.

If veridicality interacts with responsivity in some way that is noticeable at the scale of the lexicon, we would expect this analysis to find it. From this particular exploratory analysis, we find generalizations about veridicality and factivity, but not about responsivity.

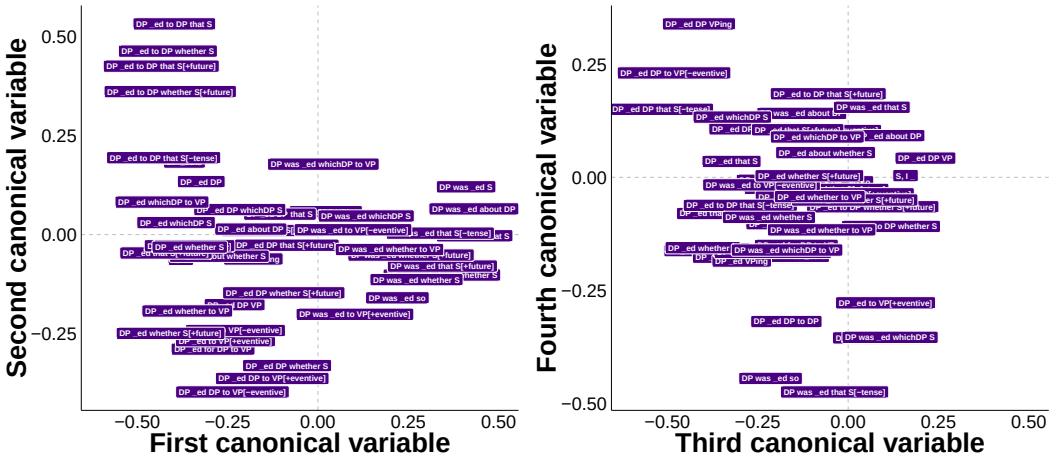
Table (18) shows the relationship between the resulting canonical variables and the veridicality measures from our two tasks. The first two variables correspond roughly to the FACTIVITY and VERIDICALITY measures defined in (15), while the third variable appears to capture true factivity, and the fourth, semi-factivity.

(18) *Relationship between first two canonical variables and veridicality*

CV	Positive (Task 1)	Negative (Task 1)	Positive (Task 2)	Negative (Task 2)
1	0.37	0.70	0.31	0.70
2	0.74	-0.01	0.73	-0.05
3	-0.49	-0.09	0.83	0.27
4	0.00	0.72	0.20	-0.67

Figure (19) plots the relationship between these canonical variables and all the syntactic frames found in MegaAttitude.

(19) *Relationship between canonical variables and syntactic distribution*



The general pattern we see here is that DP-taking and PP-taking behavior dominates the first two variables, with the first canonical variable nearly perfectly tracking whether the frame involves a passivized direct object and the second canonical variable tracking whether the frame involves a PP headed by *to*. We take this to suggest that some aspect of the semantics relevant to DP- and PP-taking is also relevant to factivity and veridicality.

For the first canonical variable, the relevant syntactic fact appears to be that the verb

takes a direct object, but not a referential external argument, since only passivized direct objects are found to load positively on that variable. This likely suggests that this variable is capturing the relationship between emotivity – or more generally, whatever aspect of a verb's semantics determine that an entity argument is realized in object position – and factivity, corroborated by analysis of the verb scores for this component (not shown).

For the second canonical variable, the relevant syntactic fact appears to be that the verb takes an optional recipient argument. This likely suggests that this variable is capturing an interesting and surprising relationship between communicativity – or more generally, transfer – and veridicality. This is interesting in light of the fact that generalizations about the relationship between veridicality or factivity and syntactic distribution are generally stated in terms of the nonveridicality of communicatives. In analyzing the verb scores for this variable, we find many veridicals – *verify*, *prove*, *show*, etc. – loading positively on this variable. In conjunction with the fact that communicatives tend to be responsive on the whole (White & Rawlins 2016), we take this to indicate a further reason why the responsivity hypothesis seemed plausible: the true semantic property controlling the distribution (communicativity) is potentially confounded with responsivity. This supports Spector & Egré's overall point that communicatives are more veridical than they might initially seem, though from our perspective it leads to the opposite conclusion about veridicality.

The third and fourth canonical variables are similarly dominated by DP-taking behavior, though there appears to be some interaction with (lack of) tense in the embedded clause present. We leave a full exploration of these variables to future work, but a tentative suggestion is that these variables indicate some relationship between the aspectual structure of the embedding verb and factivity. This is suggestive of an important next step: our data considers only declarative complements, but implicative verbs, such as *manage*, which take only infinitival complements and are veridical but not factive, are an important piece of the picture. Thus, a full treatment of the distribution of factivity and veridicality would require expanding the dataset to include judgments about non-finite complements.

6. Conclusion

We have provided a large-scale empirical investigation of the relationships between the selectional property of responsivity, factivity, and veridicality – which, since Hintikka (1975), have often been suggested to be linked in some way. In our confirmatory analysis, we showed both a negative and a positive result. The negative result disconfirms a positive connection at the scale of the entire lexicon: there is, if anything, a negative correlation between responsivity and both factivity and veridicality. The positive result (such as it is) is that there is a correlation to be found when only looking at the $\sim$100 most frequent verbs.

We then used canonical correlation analysis to explore the question of how veridicality and factivity might be linked, if not to responsivity specifically, to any other aspect of syntactic distribution. Here, the clearest finding is that the possibility or requirement of a DP direct or indirect object is a major factor. This finding may suggest that particular verb classes mediate the relationship between veridicality and syntactic distribution, and it is consonant with proposals, such as White & Rawlins's (2017), that link clausal selection to event structural properties, such as *telicity* and *change-of-state*. This possibility suggests

a reconceptualization of the set of semantic properties that determine clausal selection: rather than attempting to explain clausal selection in terms of properties that can only be defined for clause-embedding predicates – e.g. *veridicality*, *factivity*, *neg-raising*), etc. – it may be more fruitful to search for properties that are definable for the entire lexicon – e.g. *dynamicity*, *durativity*, *transfer*, *telicity*, etc.

References

Anand, Pranav, & Valentine Hacquard. 2013. Epistemics and attitudes. *Semantics and Pragmatics* 6:1–59.

Anand, Pranav, & Valentine Hacquard. 2014. Factivity, belief and discourse. In *The Art and Craft of Semantics: A Festschrift for Irene Heim*, ed. Luka Crnič & Uli Sauerland, volume 1, 69–90. Cambridge, MA: MIT Working Papers in Linguistics.

Beck, Sigrid, & Hotze Rullmann. 1999. A flexible approach to exhaustivity in questions. *Natural Language Semantics* 7:249–298.

Berman, S. 1991. The semantics of open sentences. Ph.D. dissertation, UMass Amherst.

Bogal-Allbritten, Elizabeth A. 2016. Building Meaning in Navajo. Doctoral dissertation, University of Massachusetts, Amherst.

Brysbaert, Marc, & Boris New. 2009. Moving beyond Kučera and Francis. *Behavior research methods* 41:977–990.

Cremers, Alexandre, & Emmanuel Chemla. 2017. Experiments on the acceptability and possible readings of questions embedded under emotive-factives. *Natural Language Semantics* 25:223–261.

Dudley, Rachel. 2017. The role of input in discovering presuppositions triggers: Figuring out what everybody already knew. Doctoral dissertation, University of Maryland, College Park, MD.

Dudley, Rachel, Meredith Rowe, Valentine Hacquard, & Jeffrey Lidz. 2017. Discovering the factivity of *know*. In *Proceedings of SALT 27*.

Egré, Paul. 2008. Question-embedding and factivity. *Grazer Philosophische Studien* 77:85–125.

George, Benjamin Ross. 2011. Question Embedding and the Semantics of Answers. Doctoral dissertation, University of California Los Angeles.

Ginzburg, Jonathan. 1995. Resolving questions, II. *Linguistics and Philosophy* 18:567–609.

Groenendijk, Jeroen, & Martin Stokhof. 1984. Studies on the Semantics of Questions and the Pragmatics of Answers. Doctoral dissertation, University of Amsterdam.

Hacquard, Valentine. 2006. Aspects of Modality. Doctoral dissertation, Massachusetts Institute of Technology.

Hacquard, Valentine. 2009. On the interaction of aspect and modal auxiliaries. *Linguistics and Philosophy* 32:279–315.

Hintikka, Jaakko. 1975. Different Constructions in Terms of the Basic Epistemological Verbs: A Survey of Some Problems and Proposals. In *The Intentions of Intentionality and Other New Models for Modalities*, 1–25. Dordrecht: D. Reidel.

Hotelling, Harold. 1936. Relations between two sets of variates. *Biometrika* 28:321–377.

Karttunen, Lauri. 1971a. Implicative verbs. *Language* 340–358.

Karttunen, Lauri. 1971b. Some observations on factivity. *Papers in Linguistics* 4:55–69.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.

Karttunen, Lauri. 2012. Simple and phrasal implicatives. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics*, 124–131. Association for Computational Linguistics.

Karttunen, Lauri, Stanley Peters, Annie Zaenen, & Cleo Condoravdi. 2014. The Chameleon-like Nature of Evaluative Adjectives. In *Empirical Issues in Syntax and Semantics 10*, ed. Christopher Piñón, 233–250. CSSP-CNRS.

Kiparsky, Paul, & Carol Kiparsky. 1970. Fact. In *Progress in Linguistics: A collection of papers*, ed. Manfred Bierwisch & Karl Erich Heidolph, 143–173. The Hague: Mouton.

Kratzer, Angelika. 2006. Decomposing attitude verbs.

Lahiri, Utpal. 2002. *Questions and Answers in Embedded Contexts*. Oxford University Press.

Moulton, Keir. 2009. Natural Selection and the Syntax of Clausal Complementation. Doctoral dissertation, University of Massachusetts, Amherst.

O'Donnell, Timothy J. 2015. *Productivity and reuse in language: A theory of linguistic computation and storage*. Cambridge, MA: MIT Press.

Rawlins, Kyle. 2013. About 'about'. *Semantics and Linguistic Theory* 23:336–357.

Schlenker, Philippe. 2010. Local contexts and local meanings. *Philosophical Studies* 151:115–142.

Spector, Benjamin, & Paul Egré. 2015. A uniform semantics for embedded interrogatives: An answer, not necessarily the answer. *Synthese* 192:1729–1784.

Theiler, Nadine. 2014. A multitude of answers: Embedded questions in typed inquisitive semantics. MSc thesis, University of Amsterdam.

Theiler, Nadine, Floris Roelofsen, & Maria Aloni. 2017. What's wrong with *believing whether*? In *Proceedings of SALT 27*.

Uegaki, Wataru. 2012. Content nouns and the semantics of question-embedding predicates. *Proceedings of Sinn und Bedeutung 16* 613–626.

Uegaki, Wataru. 2015. Content nouns and the semantics of question-embedding. *Journal of Semantics* 33:623–660.

White, Aaron Steven. 2014. Factive-implicatives and modalized complements. In *Proceedings of the 44th annual meeting of the North East Linguistic Society*, ed. Jyoti Iyer & Leland Kusmer, 267–278. University of Connecticut.

White, Aaron Steven, Valentine Hacquard, & Jeffrey Lidz. 2018. Semantic information and the syntax of propositional attitude verbs. *Cognitive Science* 42:416–456.

White, Aaron Steven, & Kyle Rawlins. 2016. A computational model of S-selection. *Semantics and Linguistic Theory* 26:641–663.

White, Aaron Steven, & Kyle Rawlins. 2017. Question agnosticism and change of state. In *Proceedings of Sinn und Bedeutung 21*, to appear.