Assessing and Interpreting Semantic Data from the Web

Cleo Condoravdi, Lauri Karttunen, Annie Zaenen, Stanley Peters

The WWW is a trove of language usage. This facilitates access to productions by a much broader variety of language users than was previously available; but the dialect and level of proficiency of web authors is usually unknown. Because language use is influenced by norms, treating observed uses as evidence requires identifying the pertinent language norms. In addition, some language use contravenes an applicable norm; so an observation's evidential import often depends on whether or not it conforms to applicable norms. These determinations are difficult in general, with respect to syntax and semantics alike. In the case of semantics, however, it is sometimes possible to determine "normativity" by examining how certain aspects of an expression's apparent meaning interact with other dimensions of its meaning, as this talk discusses.

One can be reasonably confident an observed example of an expression uses it to communicate a particular meaning if examination of the surrounding text confirms this inference about the target expression's meaning. An example is that "I wasn't stupid to go stumbling through the junkyard in the dark and get hurt." means the writer didn't stumble through the junkyard in the dark. Another is that "I wasn't brave to take the lie detector test. I knew I was innocent." means the author did take the lie detector test.

The following table shows two schematic inferential patterns, labeled (F) and (I), that are attested by uses of affirmative and negative sentences of the forms exemplified above with evaluative adjectives such as *stupid*, *dumb*, *foolish*, *clever*, *smart*, *brave*, *lucky*, etc.

	NP was Adj to VP	NP was not Adj to VP
(F)	NP VPed & NP VPing was Adj	NP VPed & NP VPing was not Adj
(I)	NP VPed & NP VPing was Adj	NP didn't VP & NP not VPing wasn't Adj

The lie detector example instantiates the (F) pattern for negative sentences, and the stumbling example the (I) pattern.

It has previously been proposed, based on a modest sample of English speakers' intuitions, that the (F) pattern represents the only normative English usage. When examples are pointed out of the (I) pattern successfully communicating the indicated implications, linguists have orally responded that these are contrary to norm. Some explain the communicative success by invoking facts such as the implausibility that stumbling through the junkyard in the dark would not be stupid.

Other interpretations of the attested usage data are nevertheless possible, including that both (F) and (I) patterns represent normative usage by different subpopulations of English speakers; these groups might, for example, speak slightly different dialects.

Experimental methods can help one decide between these two views. Judgments

about instances of attested usage patterns can be elicited from a substantial population of English speakers by means of Amazon Mechanical Turk, and subjects can be asked follow up questions to aid in analyzing their judgments. By systematically varying interacting dimensions of meaning such as the already mentioned plausibility, one can assess with some measure of statistical significance the likelihood that a particular interpretation of patterns like (F) and (I) is correct.

We present data from an elicitation experiment and argue that statistically significant interactions in the results favor the conclusion that type (F) and type (I) usage are both normative for subpopulations of English speakers. We briefly discuss implications of this conclusion for linguistic analysis of the construction under investigation as semantically factive or implicative. And we comment on the value of systematic elicitation to aid in interpreting observations of naturalistic language use.

References

Barker, Chris. 2002. The dynamics of vagueness. *Linguistics and Philosophy* **25**:1–36.

Norrick, Neal R. 1978. *Factive adjectives and the theory of factivity.* Tübingen: Niemeyer.