Discovering the factivity of *know* from its distribution
Rachel Dudley[1], Meredith Rowe[2], Valentine Hacquard[1], Jeffrey Lidz[1]
1 University of Maryland, College Park   2 Harvard University

Think and *know* are both attitude verbs that express beliefs, but the two verbs differ in factivity: *think* can be used to report beliefs that are either true or false, but *know* can only be used to report true beliefs that the speaker takes for granted [1]. How do children discover that *know* is a factive presupposition trigger but *think* isn't? Are there reliable cues to factivity made available in linguistic input to young children and, if so, what are they? We use corpus methods to examine children's input with *know* and *think* to determine which distributional cues might signal their (non-)factivity. We find that traditional semantic diagnostic contexts, which we often use to identify presupposition triggers as theorists, are less readily available in the input: (i) *know* is rarely used in contexts where it's complement is presupposed and (ii) *think*, just like *know*, is typically used when the speaker is committed to the truth of the complement. In addition, we find that there are differences in the syntactic and pragmatic distributions of the two verbs, which could be informative about their (non-)factivity, especially when considered in concert: (iii) *know* embeds both interrogative and declarative complements, while *think* only embeds declarative complements; and (iv) *know* is used primarily to ask and answer questions while *think* is used to make weak assertions.

Know sentences like (1)—or uses of them—trigger the presupposition that Mary is home and assert that John believes Mary to be home. Thus, in uttering (1) but not (2), the speaker should take it for granted that Mary is home, and (1) cannot be uttered felicitously if the speaker does not take Mary to be home.

(1) John knows that Mary is home.   [necessarily expresses true beliefs]
(2) John thinks that Mary is home.   [can express true or false beliefs]

Nevertheless, as observed in the literature, *know* does not reliably exhibit this behavior. Sometimes *know* is used when the addressee has no reason to take the truth of the complement for granted: discourse initial uses of *know* are felicitous (3) and over half of uses of *know* in speech between adults are "informative" (Spenader 2003). Secondly, Simons (2007) has shown how the complement of *know* can provide content with addresses the Question Under Discussion (4). And lastly, the truth of the complement does not always project out of family-of-sentences contexts with *know*, as in the numerous examples provided by Beaver (2010).

(3) Did you know that John won the lottery (uttered discourse initially)
(4) Q: Where was Louise yesterday? A: I know from Henry that she was in Princeton

What evidence of this distinction between the verbs is present in speech to children? Is it possible that children could directly observe that *know* presupposes the truth of its complement while *think* does not or that *think* can describe false beliefs while *know* cannot? Or are the counterexamples to the generalization that are found in adult speech present in speech to children as well?

To answer this questions, we examined all tokens of *know* and *think* in child ambient speech in the Gleason corpus of CHILDES (N=24 families, mean age of child = 3.5 years) [2,3]. All tokens (*know*: 1231, *think*: 1156) were coded for syntactic features that would be indicative of the contexts which the verbs are used in: person of subject, type of complement and type of clause, as well as discourse information that would suggest whether or not the complements of

the verbs express information that is in the common ground. Our results show that sentences of *x knows p* such as (1) are rare in the input, occurring in only 15% of all *know* tokens. In contrast, sentences of *x thinks p* such as (2) are frequent in the input, occurring in 85% of all *think* tokens, and approximately 2/3 of them occur with 1st person subjects in the present tense (e.g., *I think Mary will come to the party*). Out of these *x verbs p* tokens, *p* expresses information that can be taken for granted at similar rates for the two verbs: 60% of these tokens for *know* and 40% of these tokens for *think.* Together, these data suggest that (i) *know* is rarely used in contexts where its presupposition can arise, and it is not always used to describe information that can be taken for granted, (ii) while *think* is rarely used to describe false beliefs.

While these corpus data suggest that direct cues may not be the clearest evidence of *know*'s factivity and *think*'s non-factivity, there are alternative cues that are made available in the input. There are syntactic correlates of *know*'s factivity: verbs that embed interrogatives in addition to declarative complements have "factive" readings [4,5]. And *know* most naturally lends itself to different kinds of speech acts from *think* (5-6) [6].

     (5) Do you know where Mary went?   [intending to ask: Where did Mary go?]
     (6) I think she's in a meeting.   [intending to say: She's in a meeting.]

Upon hearing an utterance like (5) in the appropriate context, the listener should understand that the speaker intends to ask them a question, and they can assume that—all else equal—it is a question that they have the answer to. While these cues are more indirectly related to *know*'s factivity than direct cues might be, they can be useful in conjunction Since children have an understanding of indirect speech acts at this age [7] they might use syntactic cues like question embedding to differentiate factive verbs from others and then use their knowledge of indirect speech acts to uncover the verbs' (non-)factivity, among other aspects of their underlying semantics [8,9]. More explicitly, if the child often hears *Do you VERB what time it is?*, understanding the speaker to be asking for the time, she might conclude that the speaker expects her have the true answer to their question and thus that verbs which occur in this context relate individuals to true propositions. And our data suggest that the cues need to arrive at this conclusion are available in the input: (iii) unlike *think*, *know* frequently embeds interrogative complements (52% of *know* tokens), and (iv) *know* occurs most frequently in indirect requests for information (e.g., (4); 39% of *know* tokens) and the primary direct speech acts are answers to questions (e.g., *I don't* know; 12% of *know* tokens) and knowledge questions (e.g., *Do you know?*; 11% of *know* tokens), whereas *think* is used to make indirect assertions most often (e.g., *I think it's time for bed*; 67% of *think* tokens).

These corpus results give us a fine-grained picture of children's linguistic experience with the factive *know* (in contrast to related verbs). This kind of data can illuminate which kinds of distributional cues to factivity are reliably available in the input and, in conjunction with further experimental work, this kind of data could lead to a greater understanding of how factive presuppositions are represented. At this stage, our results suggest that traditional contexts which are diagnostic of factivity are rare in the input, that syntactic cues and pragmatic cues to factivity are readily available to the learner, and may be used together to discover the (non-)factivity of *know* and *think*.

Selected references:

[1] Stalnaker (1974). Pragmatic presuppositions. *Semantics and philosophy*.

[2] Gleason (1980). The acquisition of social speech and politeness formulae. *Language: Social psychological perspectives*.

[3] MacWhinney (2000). *The CHILDES project: The database*.

[4] Egré (2008). Question-embedding and factivity. *Grazer Philosophische Studien* 77.

[5] Hintikka (1975). *The intensions of intensionality*.

[6] Searle (1975). Indirect speech acts. *Syntax and Semantics, Vol 3: Speech Acts*.

[7] Spekman & Roth (1985). Preschool children's comprehension and production of directive forms. *Journal of Psycholinguistic Research*, 14.

[8] Lewis (2013). *Pragmatic Enrichment in Language Processing and Development*.

[9] Hacquard (2014). Bootstrapping attitudes. *Semantics and Linguistics Theory* 24.