Journal of Experimental Psychology:
Learning, Memory, and Cognition
2012, Vol. 38, No. 5, 1450–1460

0278-7393/12/$12.00 DOI: 10.1037/a0027850

RESEARCH REPORT

Speaker Knowledge Influences the Comprehension of Pragmatic Inferences

Leon Bergen
Massachusetts Institute of Technology

Daniel J. Grodner
Swarthmore College

Inferring what speakers mean from what they say requires consideration of what they know. For instance, depending on the speaker's level of expertise, uttering *Some squirrels hibernate* can imply that not all squirrels hibernate, or it might imply the weaker proposition that the speaker does not know whether all squirrels hibernate. The present study examines the extent to which speaker knowledge influences implied meanings as well as the timing of any such influence. Using a self-paced presentation, participants read sentences containing *some* in contexts where a speaker should know whether *all* was true, or where the speaker merely might know whether *all* was true. This knowledge manipulation was found to have immediate and reliable effects on the type of inference that was drawn. In contrast, knowledge played no role when the same meanings were conveyed literally. This work thus demonstrates that perceivers consider the speaker's knowledge state incrementally to establish the speaker's communicative goals.

Keywords: perspective taking, pragmatics, language comprehension, self-paced reading

Sensitivity to the mental states of others is critical for interpreting their overt behavior (Perner, 1999; Saxe, 2008). This is particularly apparent when it comes to understanding linguistic behaviors (Clark, 1996). The expressions that speakers say explicitly encode only a part of the meaning that they convey. To recover the remaining, implicit meaning, perceivers must consider the speaker's knowledge and communicative goals. An illustration is given in (1).

(1) Some of the investments lost money.

Though (1) would be literally true if all of the investments lost money, it is often the case that the speaker intends to communicate that only a proper subset lost money. Since Grice (1967/1989), the dominant view has been that such nonliteral meanings arise from a process of social reasoning. Specifically, Grice laid out a set of maxims that speakers are expected to follow in cooperative conversation. Addressees can make use of these maxims to determine the communicative goals of the speaker. For example, the subset inference associated with (1) arises from a maxim prescribing that speakers make the strongest statement compatible with their knowledge. When (1) is uttered, the addressee reasons that the speaker could not have truthfully said anything stronger. Hence the addressee infers that the speaker was not in a position to utter the stronger statement that all the investments lost money. This reasoning typically leads to the inference that not all of the investments lost money.

The not-all inference is a species of conversational inference called a *scalar implicature*. Scalar implicatures are pervasive; they can arise whenever a speaker uses an expression that is weaker (less informative) than a salient alternative (Hirschberg, 1985; Horn, 1972, 1989). It is worth emphasizing the social aspect of computing scalar implicatures. Under a strictly Gricean account, the addressee must reason about the speaker's perspective and what else she might have said from that perspective. This requires consideration of the speaker's goals (such as providing the addressee with relevant information) and, possibly, the speaker's knowledge.

Though perspective taking is essential to the standard account of scalar implicatures, there has been little experimental evidence bearing on how or when speaker perspective is applied in their comprehension. This would be useful in part because there has been disagreement about how sensitive scalar implicatures are to what the speaker knows. If w is the sentence uttered by the speaker and s is a salient alternative that is stronger than w, there are a number of possible implicatures. Some investigators have proposed that uttering w implicates the negation of s ($\neg s$; Horn, 1972) or that the speaker knows s to be false ($K\neg s$; Gazdar, 1979). These accounts predict a strong implicature—that the stronger statement is actually false. These accounts provide no mechanism by which the speaker's knowledge might alter the implicature. Other investigators have argued that uttering w implicates merely that the speaker does not know the stronger statement is true ($\neg Ks$; Horn, 1989; Soames, 1982; van Rooij & Schulz, 2004). This is weaker than positing that s is false. The weak implicature corresponding to (1) is that

This article was published Online First April 30, 2012.

Leon Bergen, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology; Daniel J. Grodner, Department of Psychology, Swarthmore College.

Correspondence concerning this article should be addressed to Daniel J. Grodner, Department of Psychology, Swarthmore College, 500 College Avenue, Swarthmore, PA 19081. E-mail: dgrodne1@swarthmore.edu

the speaker does not know that all of the investments lost money, whereas the strong implicature is that not all of the investments lost money. The weak implicature does not explain how sentences like (1) can be interpreted as implying that *s* is false. In order to generate the strong implicature these accounts also include mechanisms that can strengthen a weak implicature. For instance, if the speaker is assumed to know whether the stronger statement is true or not ($Ks \vee K\neg s$), the weak implicature can be strengthened to generate a strong implicature interpretation [$\neg Ks \wedge (Ks \vee K\neg s) \Rightarrow K\neg s$] (see Geurts, 2010; Russell, 2006; Sauerland, 2004, for more examples of strengthening mechanisms). Thus, for existing weak implicature accounts, but not for strong implicature accounts, the ultimate interpretation of the perceiver depends on what he knows about the knowledge state of the speaker.

The present study investigates how, and at what point in interpretation, speaker knowledge affects the comprehension of scalar implicatures. Previous work has examined whether scalar implicature processing is context sensitive. Some theorists have argued that the strong implicature is generated automatically and mandatorily upon encountering a weak expression regardless of the context (Levinson, 2000). However, recent evidence rebuts this view. For instance, Breheny, Katsos, and Williams (2006) had individuals read sentences containing a weak scalar trigger (*some*) in contexts that supported either a strong implicature or a literal interpretation. A sample stimulus from their Experiment 3 is given in Table 1. The trigger sentences were followed by a sentence containing the anaphoric expression *the rest*. This refers to the complement set evoked by a strong scalar implicature (the relatives that were not to be hosted by John).

The scalar trigger should have been read more slowly if the strong implicature had been generated. This is because the scalar quantifier would convey more information (*some and not all*) than if no implicature were generated (*some and possibly all*). Interpreting more informative linguistic elements increases comprehension difficulty and slows reading times (see Levy, 2008, in press, for a review; see also Bott & Noveck, 2004; De Neys & Schaeken, 2007, for evidence that scalar implicature interpretations are costly.) In contrast, the strong implicature interpretation should have facilitated processing of the complement sentence containing *the rest*. This is because the complement set to which the anaphor refers would have been established by the earlier trigger. Without the strong implicature, the complement set must be computed on the fly in this sentence. Consistent with these predictions, the implicature supporting contexts elicited longer reading times (RTs) over the region containing a scalar trigger (see also Katsos, 2007; Panizza, Chierchia, & Clifton, 2009; Schwarz, Clifton, & Frazier, in press) and faster processing over the subsequent anaphor. This demonstrates that the strong implicature is not obligatorily generated whenever a weak scalar expression is encountered (see also Bott & Noveck, 2004).

Table 1
Sample Stimulus From Breheny, Katsos, and Williams (2006) Experiment 3

Sentence type	Example
Context	
Implicature supporting	Mary asked John whether he intended to host all his relatives in his tiny apartment.
Literal supporting	Mary was surprised to see John cleaning his apartment and she asked why.
Trigger	John replied that he intended to host some of his relatives.
Complement	The rest would stay in a nearby hotel.

Note. Adapted from "Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences, by R. Breheny, N. Katsos, and J. Williams, 2006, *Cognition*, *100,* p. 453. Copyright 2005 by Elsevier.

The Breheny et al. (2006) results demonstrate that context affects online implicature generation but do not speak to the role of speaker knowledge. They showed that the strong implicature is suspended in certain contexts by manipulating the salience of the stronger alternative expression, not the knowledge state of the speaker. There are reasons to believe that speaker knowledge would have different effects from other kinds of context. First, a number of recent proposals suggest that scalar implicatures are computed as part of the conventional content (lexical or grammatical) of a sentence (Chierchia, 2004, 2006; Chierchia, Fox, & Spector, in press; Levinson, 2000). These accounts do not include a grammatical mechanism for consideration of the knowledge state of the speaker. They thus predict either a strong implicature or no implicature at all.[1]

Second, keeping track of the speaker's knowledge might pose a computational challenge for the language comprehension mechanism. The information known by a speaker is heterogeneous, open ended, not directly observable, and can be at odds with the knowledge of a perceiver. These sorts of observations have led some investigators to argue that perceivers are fundamentally egocentric and do not incorporate speaker knowledge immediately during online interpretation because it would be too computationally burdensome (Barr, 2008; Garrod & Pickering, 2004; Keysar, Lin, & Barr, 2003). Other investigators have theorized that this kind of perspective taking is immediately incorporated into natural comprehension in a continuous probabilistic fashion as other types of context would be (Brennan & Hanna, 2009; Brown-Schmidt, Gunlogson, & Tanenhaus, 2008; Grodner & Sedivy, 2011; Heller, Grodner, & Tanenhaus, 2008). On this view, the complexity of computing perspective-specific information may be mitigated because it requires a type of social pragmatic reasoning at which humans appear to excel even in infancy (Csibra, 2010; Liszkowski, Carpenter, & Tomasello, 2008). If the present study yields evidence that speaker knowledge affects the incremental computation of scalar implicatures, it would militate against the view that comprehenders are inherently egocentric as well as the view that grammatical mechanisms alone can capture scalar implicatures.

[1] Strictly speaking, these theories are not inconsistent with the presence of a weak implicature; however, they provide no grammatical mechanism to explain how one would arise. Only Chierchia et al. (in press) predicted a weak implicature, but to do so they included a nongrammatical Gricean mechanism that operates like the one proposed by Sauerland (2006).

The Present Experiment

The present study adapts the method of Breheny et al. (2006) to examine whether implicature generation is sensitive to the speaker's knowledge state. Participants read first-person sentences containing a scalar trigger of the form *Some P Q*. These were preceded either by a context for which the speaker was likely to know whether the stronger statement was true (full-knowledge) or by a context for which it was merely possible that the speaker knew whether the stronger statement was true (partial-knowledge). The trigger sentences were followed by a sentence that began with an anaphor referring to the complement set (the *P* that were not *Q*). An example is given in Table 2.

If perceivers incorporate speaker knowledge during online comprehension, then they will generate a strong implicature in the full-knowledge context but a weak implicature in the partial-knowledge context. In this case, the scalar quantifier will convey more information in the full-knowledge condition (meaning *some and not all* as opposed to merely *some and possibly all*). The cost of integrating this extra information should cause elevated reading times over the scalar quantifier. In contrast, when the subsequent complement sentence is processed, reading times should reveal the reverse pattern of processing difficulty. In the full-knowledge condition, computing the strong implicature would have evoked a complement set during the trigger sentence. For the sample stimulus, this corresponds to a set of investments that did not lose money. Referring back to this set in the complement sentence should be relatively straightforward. By contrast, in the partial-knowledge condition no such implicature should arise, and the complement set will not be invoked. The complement set must therefore be instantiated on the fly while processing the anaphor and predicate of the complement sentence. The additional difficulty of establishing reference should result in elevated reading times. Thus we would anticipate the complement continuation to be processed more easily in the full-knowledge condition.

An additional control condition included a focus particle in the trigger sentence (*Only some* . . .). This has the effect of asserting the not-all interpretation as part of the literal content of the utterance (Rooth, 1992). As a result, contextual information about the speaker's knowledge should not alter interpretation for either the trigger sentence or the complement sentence. Perceivers will have to process the same amount of information at the quantifier in each condition. Thus the trigger sentence should be equally difficult regardless of the knowledge context. In addition, a complement set will be constructed after the focused trigger in both contexts. This should cause the anaphoric reference in the complement sentence to be equally easy to assign for each knowledge context.

Method

Participants

Forty-two native English-speaking Swarthmore College students were paid $8 to participate.

Materials

Twenty-four stimulus passages consisting of three sentences told from a first-person perspective were prepared. Each item was constructed by crossing two factors: knowledge-context (full or partial) and trigger-type (scalar or focused). In addition to the four critical conditions, two additional conditions were run. These were identical to the scalar–trigger conditions except that the continuation sentence was replaced with a sentence that asserted that the strong alternative was true (e.g., *In fact, they all did* . . .). This contradicts either a strong or weak implicature. These conditions were intended to explore whether the difficulty of implicature cancelation might differ as a function of knowledge context. No processing differences were observed over any region of these sentences (ts $<$ 1.1; maximal difference = 8 ms). Because this null result is open to multiple interpretations, we do not discuss it further. Six counterbalanced presentation lists were prepared using a Latin square design. Stimuli were pseudorandomly mixed with 65 fillers consisting of 2- to 4-sentence first-person passages.

The first sentence of full knowledge passages depicted the speaker as having expertise in the main topic of the passage and therefore as in a position to know whether the stronger alternative was true. For partial-knowledge conditions, this sentence did not resolve the extent of the speaker's knowledge. To ensure that the contexts were biased in this way, 20 additional participants completed an offline survey. They were presented with the context sentence and were asked to judge on a scale from 1 to 7 if the speaker would know whether the strong alternative was true (e.g., "How likely is it that the speaker knows whether all of the real estate investments lost money?"). The items were assigned to one of two lists and were presented in a single pseudorandom order.

Table 2
Sample Stimulus Item With a Complement Continuation

Sentence type	Example
Context	
Full-knowledge	At my client's request, I meticulously compiled the investment report.
Partial-knowledge	At my client's request, I skimmed the investment report.
Trigger	
Scalar	Some of the real_estate investments lost money.
Focused	Only some of the real_estate investments lost money.
Continuation	
Complement	The rest were successful despite the recent economic downturn.

Note. Each stimulus item consisted of a context sentence, followed by a trigger sentence, followed by a continuation sentence.

Each subject judged only the full- or partial-knowledge version of each item. Ratings were reliably higher for full-knowledge (5.6 ± .12) than for partial-knowledge (3.6 ± .15).[2]

The second sentence of the critical passages began with either a scalar quantifier, *Some of*, or a focused quantifier, or *Only some of*. This was followed by a noun phrase and a predicate. The referent of the noun phrase was not explicitly mentioned in the context sentence but could be easily accommodated based on the content of the context (e.g., *the real estate investments* after *investment portfolio*).

The final sentence of the complement sentences began with either *The rest* or *The others* (12 items apiece). This anaphor on its own does not unambiguously refer to the complement set (e.g., "Only some of the real estate investments lost money. The rest of my portfolio was holding steady."). In order to fully disambiguate toward complement set reference, the anaphor was followed by a predicate that contrasted with the predicate of the trigger sentence (e.g., *lost money* vs. *were successful*). This was to facilitate integration with the complement set. A processing advantage should emerge by the point at which the predicate is encountered if a complement set has already been established. The complete set of stimuli is provided in Appendix A.

Procedure

Sentences were presented using a noncumulative, self-paced, word-by-word display on a Macintosh computer running Linger v.2.94 (Rohde, 2003). Each trial began with dashes standing in for non-white-space characters. Participants pressed the space bar to replace each series of dashes with the word it concealed. This caused the previous word to disappear. Time between button presses was recorded. Following each passage, participants answered a yes-or-no comprehension question.

Results

Participants responded accurately to 91% of comprehension questions. Two participants had accuracy rates below 80%. Their data were eliminated from RT analyses. Extreme individual RTs (greater than 1,500 ms or less than 100 ms) were trimmed (0.3% of all data). The remaining RTs were log-transformed and submitted to linear mixed-effect regression analyses using lme4 (Bates, 2008) in the statistical language R (R Development Core Team, 2008). Participants and items were modeled as crossed random factors. Speaker knowledge (full vs. partial), trigger-type (scalar vs. focused), and their interaction were modeled as fixed effects using analysis of variance (ANOVA)-style sum coding. The random effects specifications for these models are provided in Appendix B. Analysis regions and untransformed RTs are provided in Figures 1, 2, and 3.

Trigger Sentences

At the focus particle, there was no effect of speaker knowledge ($t = 0.78$). Over the critical quantifier region, there was no effect of speaker knowledge ($t = 1.03$), a main effect of trigger-type ($\beta = .035$, $t = 4.98$, p Markov Chain Monte Carlo [MCMC] = .0001), and a reliable interaction ($\beta = .016$, $t = 2.37$, pMCMC = .0164). To investigate individual contrasts, separate models were fit to the scalar and focused conditions. Consistent with the predictions of the knowledge sensitivity account, full-knowledge slowed processing for the scalar conditions ($\beta = .023$, $t = 2.72$, pMCMC = .0080), but speaker knowledge did not affect the focused conditions ($t = -1.07$).

Over the next two words, there was a main effect of speaker knowledge ($\beta = .015$, $t = 2.49$; $p < .05$), but there was no effect of trigger type, nor was there an interaction (ts < 0.5). Separate comparisons were again conducted for the scalar and focused items. After a scalar trigger, the full-knowledge conditions continued to be read more slowly ($\beta = .017$, $t = 2.0$, $p < .05$), but no difference emerged after a focused trigger ($t = 1.3$). There were no reliable effects or interactions over subsequent regions (ts < 1.3).

Complement Sentences

Similar analyses were conducted over each region of the complement sentence. Over the anaphor, none of the conditions differed (ts < 1.4). The critical region was the predicate where the anaphoric referent was disambiguated to the complement set. There were no main effects of knowledge or trigger type (ts < .65), but there was a reliable interaction ($\beta = .018$, $t = 2.09$; $p < .05$). The data from the focused and scalar conditions were again fit to separate models. Consistent with the knowledge sensitivity account, the full-knowledge condition was processed significantly faster than the partial-knowledge condition after a scalar trigger ($\beta = .022$, $t = 2.56$, pMCMC = 0.0146), but there was no effect of speaker knowledge after a focused trigger ($t = 1.2$). There were no reliable effects or interactions in subsequent regions (ts < 1.2).

Correlational Analyses

Speaker knowledge had reliable and opposing effects at the quantifier region of the trigger sentences and the predicate region of the complement sentences. If these effects are both caused by sensitivity to speaker knowledge, they should be inversely correlated. Factors that promote the computation of the strong implicature when processing the quantifier should later facilitate complement set anaphora. To investigate this, a correlational analysis was conducted between RTs over the quantifier and predicate regions. Note that reaction times within any single trial of a reading experiment are highly correlated. For example, in focused trials, where the role of speaker knowledge is minimized, RTs over the quantifier were an excellent predictor of RTs at the predicate of the subsequent sentence, $r = .57$; $F(1, 317) = 156$, $p < .0001$. This powerful relationship is likely to overwhelm other within trial effects. To correct for the effects of merely being in the same trial, a hierarchical regression was conducted. The linear model estimated from the focused trigger trials was used to predict RTs at the predicate for the scalar trigger trials. These predictions were subtracted from the observed values, and the residuals were entered into the analysis. Consistent with knowledge sensitivity, there was a reliable inverse relationship between RTs at the quantifier and

[2] For 20 items this difference was reliable (Fs > 4, ps < .05), but all items followed the pattern. The context sentences for the remaining four items were altered slightly to enhance the difference between conditions for the reading study.

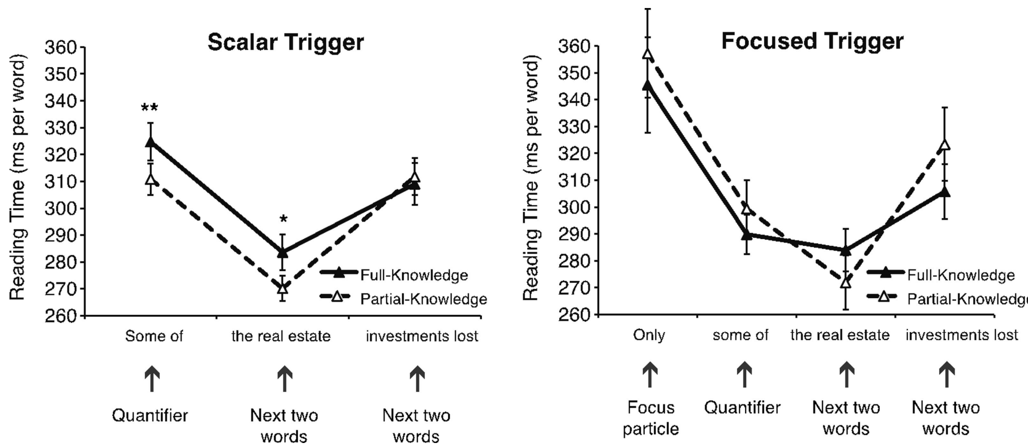

Figure 1. Mean reading times for trigger sentences used in the present study. Error bars depict one standard error of the item means. * Indicates a reliable contrast at $p < .05$; ** Indicates a reliable contrast at $p < .01$.

residual RTs at the predicate for scalar trials, $r = -.39$; $F(1, 318) = 56.7$, $p < .0001$.

Discussion

The evidence indicates that speaker knowledge reliably affected the likelihood of generating the strong implicature. This conclusion is supported by two opposing effects. In the trigger sentence, full-knowledge caused greater processing difficulty at the scalar quantifier, where the extra information associated with a strong implicature would have to be integrated. In the complement sentence, full-knowledge facilitated processing when referring back to the complement set evoked by the strong implicature. These two effects were inversely correlated at the individual trial level. Thus factors associated with the penalty at the quantifier contributed to the processing advantage at the predicate. Importantly, there were no effects of speaker knowledge in the trigger sentence or in the complement sentence, for the focused conditions. Under focus, the *not-all* interpretation and associated complement set were obligatory rather than merely implicated. As a result, speaker knowledge did not affect whether the scalar implicature was generated or whether the complement set was available.

A potential concern with the above explanation is that manipulating speaker knowledge required altering the lexical content of the context sentences. Thus lexical differences unrelated to speaker knowledge per se may have led to differences in processing that were reflected in the subsequent sentence. However, this alternative explanation does not explain the lack of context effects for the focused trigger sentences, the interaction between context and trigger type, or the inverse correlation between the effects of knowledge at the quantifier and the predicate regions. Furthermore, *Only* was read slightly faster (albeit nonsignificantly) in the full-knowledge condition so it cannot be the case that full-knowledge contexts uniformly caused slower RTs at the onset of the next sentence. Finally, for 18 of our items, at least the last two words of the context sentences were identical. When these items were analyzed separately, the statistical patterns were unchanged (interaction at the quantifier: $t = 2.5$, pMCMC $< .01$; knowledge

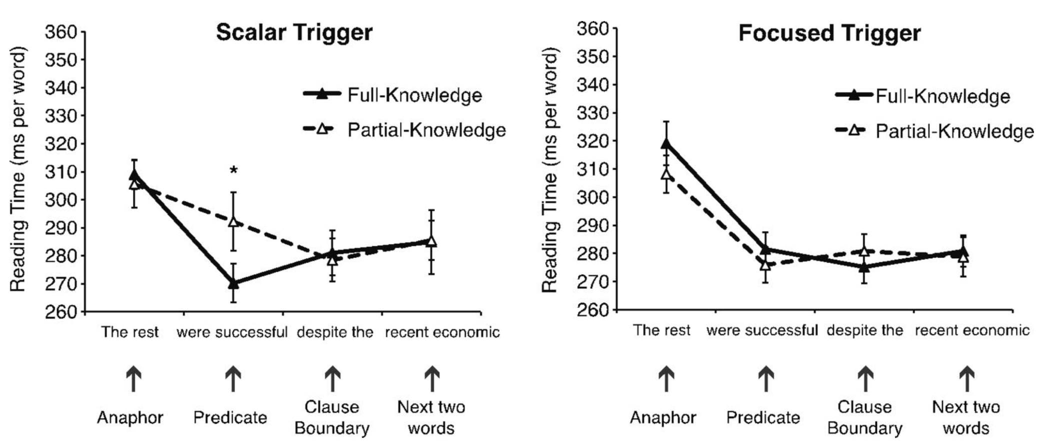

Figure 2. Mean reading times for complement sentences used in the present study. Error bars depict one standard error of the item means. * Indicates a reliable contrast at $p < .05$.

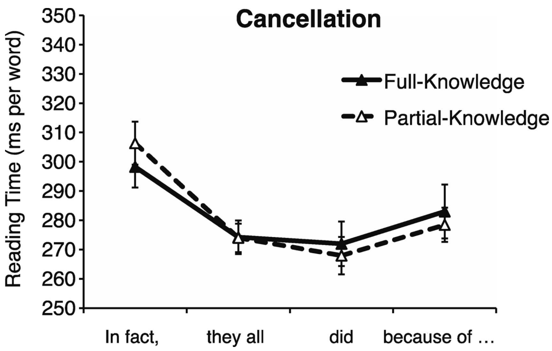

Figure 3. Mean reading times for cancelation sentences used in the present study. Error bars depict one standard error of the item means.

effect for scalar items: $t = 2.3$, pMCMC $<.05$; knowledge effect for focused items: $t = -1.8$, *ns*).[3]

It is important to acknowledge that the identity of the trigger sentence's subject was not yet determined at the scalar quantifier where the effects of speaker knowledge were first observed. In order to explain the present results, the knowledge context must have contributed to perceiver expectations about likely subjects. The full-knowledge conditions were constructed so that the speaker was strongly constrained to have expertise in the salient topic of the passage (e.g., the speaker "carefully inventoried my wine collection" or "meticulously compiled the investment report"). It would be a non sequitur to begin speaking about a completely unrelated topic in the trigger sentence. As a result, when the speaker utters *Some of* in the full-knowledge context, it is highly likely that she or he has expertise in the topic domain and, hence, knows whether *All of* would hold true.

The present results are also relevant to the question of whether computing the implicature is delayed relative to the literal content of an utterance (Grodner, Klein, Carbary, & Tanenhaus, 2010; Huang & Snedeker, 2009, 2011). Using a visual world paradigm, Huang and Snedeker (2009, 2011) found evidence that eye movements to referential targets that could be identified based on literal meaning emerged 200 ms after the literal trigger in the speech input. Fixations to targets that could be identified by a strong implicature were not observed until 800 to 1,200 ms after the trigger. Grodner et al. (2010) argued that the delay was not due to a late arriving inference but rather due to difficulty using the implicated meaning in an identification task when visual properties of the stimuli, referential labels in the fillers, and diffuse communicative goals of the speaker provided weak contextual support for the inference. The current study revealed a difference between full- and partial-knowledge immediately upon encountering the scalar trigger. The difference processing *some* was reliable ($M = 332$ ms; Knowledge $\times$ Trigger-Type interaction: $\beta = .027$, $t = 2.54$, pMCMC $= .0116$; main effect of knowledge for scalar: $\beta = .039$, $t = 2.97$, pMCMC $= .05$). Thus there is evidence of the strong implicature within a few hundred milliseconds of the trigger in a reading task when provided with strong contextual support as in the full-knowledge conditions.

We have demonstrated knowledge sensitivity in processing a paradigmatic case of scalar implicature: the existential quantifier *some*. This quantifier and the associated implicature appear to be universal across languages (Horn, 1989). It is also among the most common words in human language. The frequent co-occurrence of the quantifier and the associated implicature make it more likely that the strong implicature could become a conventional aspect of the existential quantifier's meaning. In turn, this decreases the probability that idiosyncratic elements of the context (such as speaker knowledge) could affect its interpretation. The robust effects of speaker knowledge here strongly suggest that other scalar implicatures should be sensitive to speaker knowledge. Indeed, using a visual world, eye-tracking paradigm, Breheny, Katsos, and Ferguson (2010) provided preliminary evidence that speaker knowledge also incrementally affects ad hoc scalar implicatures, which are not triggered by a particular lexical item.

Conclusion

The present study demonstrates that speaker knowledge affects incremental implicature computation. This runs counter to proposals that claim the strong implicature is generated obligatorily or which provide no mechanism for incorporating speaker knowledge. These results also argue against the notion that there is special difficulty using perspective information in online processing. The findings are also prima facie at odds with claims that reasoning about others' knowledge is inherently difficult or delayed (Birch & Bloom, 2007; Epley & Caruso, 2008; Lin, Keysar, & Epley, 2010). It is notable that such difficulty is typically found when investigating phenomena, such as false belief, where the experimental participant's egocentric knowledge contains highly salient information that is explicitly in conflict with the other's knowledge. For the present phenomenon this was not the case: The addressee should have had no prior commitment as to the truth of the stronger alternative statement. Our results are thus consistent with other recent findings that show perceivers' egocentric biases are diminished when another's perspective is made sufficiently salient relative to one's own (Brown-Schmidt et al., 2008; Ferguson & Breheny, 2011; Hanna, Tanenhaus, & Trueswell, 2003; Heller, Grodner, & Tanenhaus, 2009).

[3] Moreover, steps were taken to buffer regions of interest from differences in the context sentence. The knowledge contexts were of similar lengths (partial: 13.04 words, full: 13.74 words, $t < 1.46$), and the trigger sentence always began on a new line.

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* New York, NY: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi:10.1016/j.jml.2007.12.005

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition, 109, 1,* 18–40. doi:10.1016/j.cognition.2008.07.005

Bates, D. (2008). *Linear mixed model implementation in lme4.* Unpublished manuscript, University of Wisconsin, Madison.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18,* 382–386. doi:10.1111/j.1467-9280.2007.01909.x

Bott, L., & Noveck, I. A. (2004). Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language, 51,* 437–457. doi:10.1016/j.jml.2004.05.006

Breheny, R., Katsos, N., & Ferguson, H. J. (2010). *Taking the epistemic step: Towards a model of on-line access to conversational implicatures.*

Poster presented at the 23rd Annual CUNY Conference on Human Sentence Processing, New York, NY.

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition, 100,* 434–463. doi:10.1016/j.cognition.2005.07.003

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science, 1,* 274–291. doi:10.1111/j.1756-8765.2009.01019.x

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition, 107, 3,* 1122–1134. doi:10.1016/j.cognition.2007.11.005

Chierchia, G. (2004). Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond* (pp. 39–103). New York, NY: Oxford University Press.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry, 37,* 535–590. doi:10.1162/ling.2006.37.4.535

Chierchia, G., Fox, D., & Spector, B. (in press). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. Maienborn, K. von Heusinger, & P. Portner (Eds.), *Handbook of semantics.* Berlin, Germany: Mouton de Gruyter.

Clark, H. H. (1996). *Using language.* Cambridge, England: Cambridge University Press.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language, 25,* 141–168. doi:10.1111/j.1468-0017.2009.01384.x

De Neys, W., & Schaeken, W. (2007). When people are more logical under cognitive load: Dual task impact on scalar implicature. *Experimental Psychology, 54,* 128–133. doi:10.1027/1618-3169.54.2.128

Epley, N., & Caruso, E. M. (2008). Perspective taking: Misstepping into others' shoes. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds.), *The handbook of imagination and mental simulation* (pp. 295–309). New York, NY: Psychology Press.

Ferguson, H. J., & Breheny, R. (2011). Eye movements reveal the time-course of anticipating behaviour based on complex, conflicting desires. *Cognition. 119,* 179–196. doi:10.1016/j.cognition.2011.01.005

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8,* 8–11. doi:10.1016/j.tics.2003.10.016

Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form.* New York, NY: Academic Press.

Geurts, B. (2010). *Quantity implicatures.* New York, NY: Cambridge University Press. doi:10.1017/CBO9780511975158

Grice, H. P. (1989). Logic and conversation. In P. Grice, *Studies in the way of words* (pp. 22–40). Cambridge, MA: Harvard University Press. (Original address given 1967)

Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). Rapid interpretation of pragmatic "some." *Cognition, 116,* 42–55. doi:10.1016/j.cognition.2010.03.014

Grodner, D. J., & Sedivy, J. C. (2011). The effect of speaker-specific information on pragmatic inferences. In E. Gibson & N. Pearlmutter (Eds.), *The processing and acquisition of reference* (pp. 239–272). Cambridge, MA: MIT Press.

Hanna, J., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory & Language, 49,* 43–61.

Heller, D., Grodner, D., & Tanenhaus, M. K. (2008). The role of perspective in identifying domains of reference. *Cognition, 108,* 831–836. doi:10.1016/j.cognition.2008.04.008

Heller, D., Grodner, D., & Tanenhaus, M. K. (2009). Taking perspective: Evidence for real-time integration of information about common ground. In U. Sauerland & K. Yatsushiro (Eds.), *Semantics and pragmatics: From experiment to theory* (pp. 228–248). New York, NY: Palgrave Macmillan.

Hirschberg, J. (1985). *A theory of scalar implicature* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia, PA.

Horn, L. R. (1972). *On the semantic properties of logical operators in English* (Unpublished doctoral dissertation). University of California, Los Angeles.

Horn, L. R. (1989). *A natural history of negation.* Chicago, IL: University of Chicago Press.

Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology, 58,* 376–415. doi:10.1016/j.cogpsych.2008.09.001

Huang, Y. T., & Snedeker, J. (2011). Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes, 26,* 1161–1172.

Katsos, N. (2007). *Structural and contextual constraints on processing scalar implicatures.* Unpublished manuscript, Cambridge University, Cambridge, England.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition, 89,* 25–41. doi:10.1016/S0010-0277(03)00064-7

Levinson, S. (2000). *Presumptive meanings: The theory of generalized conversational implicature.* Cambridge, MA: MIT Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition, 106,* 1126–1177. doi:10.1016/j.cognition.2007.05.006

Levy, R. (in press). *Probabilistic models in the study of language.* Cambridge, MA: MIT Press.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology, 46,* 551–556. doi:10.1016/j.jesp.2009.12.019

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition, 108,* 732–739. doi:10.1016/j.cognition.2008.06.013

Panizza, D., Chierchia, G., & Clifton, C., Jr. (2009). On the role of entailment pattern and scalar implicatures in the processing of numerals. *Journal of Memory and Language, 61,* 503–518. doi:10.1016/j.jml.2009.07.005

Perner, J. (1999). Theory of mind. In M. Bennett (Ed.), *Developmental psychology: Achievements and prospects* (pp. 205–230). Hove, England: Psychology Press.

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Rohde, D. (2003). Linger, v 2.94. Available from http://tedlab.mit.edu/~dr/linger/

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1,* 75–116. doi:10.1007/BF02342617

Russell, B. (2006). Against grammatical computation of scalar implicatures. *Journal of Semantics, 23,* 361–382. doi:10.1093/jos/ffl008

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and Philosophy, 27,* 367–391. doi:10.1023/B:LING.0000023378.71748.db

Sauerland, U. (2006).

Saxe, R. (2008). Theory of mind—neural basis. In W. P. Banks (Ed.), *Encyclopedia of Consciousness.* Oxford, England: Academic Press.

Schwarz, F., Clifton, C., Jr., & Frazier, L. (in press). Focus, downward entailing contexts, and scalar implicatures. In J. Anderssen, K. Moulton, F. Schwarz, & C. Ussery (Eds.), *Semantics and processing*. Amherst, MA: GLSA Publications.

Soames, S. (1982). How presuppositions are inherited: A solution to the projection problem. *Linguistic Inquiry, 13,* 483–545.

van Rooij, R., & Schulz, K. (2004). Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information, 13,* 491–519. doi:10.1007/s10849-004-2118-6

Appendix A

Experimental Materials

Experimental stimuli are given below. See main text for details. Items were displayed word by word. Words connected with hyphens or underscores were presented together.

Full: I carefully inspected the new shipment of jewelry.
Partial: I helped unload the new shipment of jewelry.
Trigger: (Some/Only some) of the gold watches were fakes.
Complement: The rest were real, but the company is still planning to sue.
Cancelation: In fact, they all were, so the company is planning to sue.

Full: Before the hurricane landed, I checked every house in town.
Partial: Before the hurricane landed, I volunteered to help out in town.
Trigger: (Some/Only some) of the residents had evacuated.
Complement: The rest stayed at home and foolishly risked their lives.
Cancelation: In fact, they all did and, as a result, they survived the storm.

Full: This morning, I took attendance at an important meeting with the manager.
Partial: This morning, I heard about at an important meeting with the manager.
Trigger: (Some/Only some) of the company's accountants were there.
Complement: The rest were missing because they had to audit the company's finances before the end of the quarter.
Cancelation: In fact, they all were in order to communicate how budget cutbacks were crippling their division.

Full: I was the chief medic overseeing the bus accident on the side of the highway.
Partial: I stopped to see if I could help with the bus accident on the side of the highway.
Trigger: (Some/Only some) of the passengers were still alive.
Complement: The others were killed on impact and there was no opportunity to save them.
Cancelation: In fact, they all were and none were critically injured.

Full: At my client's request, I meticulously compiled the investment report.
Partial: At my client's request, I skimmed the investment report.
Trigger: (Some/Only some) of the real_estate investments lost money.
Complement: The others were successful in spite of the recent economic downturn.
Cancelation: In fact, they all did because of the recent economic downturn.

Full: As the office's main technician, I had to check each computer for the dangerous new virus.
Partial: As the office's main technician, I had been alerted to the dangerous new virus.
Trigger: (Some/Only some) of our computers were infected.
Complement: The rest were clean because their owners had been very cautious.
Cancelation: In fact, they all were and the virus nearly destroyed the whole system.

Full: Earlier today, I was leading a small group of tourists around the sights downtown.
Partial: Earlier today, I passed by a group of tourists looking at sights downtown.
Trigger: (Some/Only some) of tourists got soaked by the rain storm.
Complement: The rest were dry because they had remembered their umbrellas.
Cancelation: In fact, they all did because they had forgotten their umbrellas.

Full: After the furniture sale, I catalogued all the merchandise in the store.
Partial: After the furniture sale, I closed down and locked up the store.
Trigger: (Some/Only some) of the leather couches had been sold.
Complement: The rest were going to be stored in the warehouse until the following season.
Cancelation: In fact, they all had since they were stylish and cheap.

Full: Last week, I tasted every dish at a family potluck.
Partial: Last week, I attended a family potluck.
Trigger: (Some/Only some) of the dishes were spicy.
Complement: The rest were mild and I found them to be bland.
Cancelation: In fact, they all were but fortunately I love spicy food.

(*Appendices continue*)

Full: When I entered Disney World, I asked about the status of each of the rides.
Partial: When I visited Disney World, I quickly walked around the amusement park.
Trigger: (Some/Only some) of my favorite rides were still running.
Complement: The rest were shut_down since they were no longer popular.
Cancelation: In fact, they all were since they were still popular.

Full: After my house was burglarized, I carefully inventoried my wine collection.
Partial: After my house was burglarized, I briefly checked the basement.
Trigger: (Some/Only some) of the bottles of Chardonnay were missing.
Complement: The rest were safe but I was still extremely upset.
Cancelation: In fact, they all were even though I had secured them.

Full: While working in the veterinary clinic, I closely examined the mouth of large bulldog.
Partial: While working in the veterinary clinic, I caught a glimpse of a large bulldog.
Trigger: (Some/Only some) of the dog's teeth were missing.
Complement: The rest were intact so it should still be able to eat solid food.
Cancelation: In fact, they all were because its owners completely neglected its oral hygiene.

Full: In the school parking lot, I carefully inspected an old bus.
Partial: In the school parking lot, I passed by an old bus.
Trigger: (Some/Only some) of its tires were flat.
Complement: The others were fine so it would not cost too much to fix it.
Cancelation: In fact, they all were so the cost to repair it would be enormous.

Full: To check on the progress of my research, I meticulously recorded the results of the experiments.
Partial: To check on the progress of my research, I quickly scanned the experiments in the lab.
Trigger: (Some/Only some) of my predictions were correct.
Complement: The rest were wrong so my theory must be mistaken.
Cancelation: In fact, they all were so I should be able to publish the results.

Full: When they returned, I reviewed each of the receipts for my family's trip to Europe.
Partial: When they returned, my family told me about their trip to Europe.
Trigger: (Some/Only some) of their hotels were expensive.
Complement: The rest were cheap despite my family's love for luxurious holidays.
Cancelation: In fact, they all were because my family loves luxurious holidays.

Full: After the babysitter left, I carefully examined my liquor collection.
Partial: After the babysitter left, I went to my liquor collection and made a drink.
Trigger: (Some/Only some) of my new bottles of vodka were opened.
Complement: The others were untouched, but I was still concerned and decided to call her parents.
Cancelation: In fact, they all were but I decided not to call her parents because it was so hard to find a babysitter.

Full: I examined the damage after I dropped a bowling ball down the stairs.
Partial: I heard that my friend dropped a bowling ball down the stairs.
Trigger: (Some/Only some) of the steps were damaged.
Complement: The others were fine so the repairs shouldn't be too expensive.
Cancelation: In fact, they all were so they will require extensive repairs.

Full: At a friend's suggestion, I completely worked through an entire math textbook.
Partial: At a friend's suggestion, I read a review of a new math textbook.
Trigger: (Some/Only some) of its problems were difficult.
Complement: The others were straightforward and I feel like I learned a lot.
Cancelation: In fact, they all were but it received a positive review anyway.

Full: To keep the laboratory well stocked, I write down exactly how much of the chemicals are used every night.
Partial: To keep the laboratory well stocked, I occasionally look at which chemicals are being used.
Trigger: Today, (some/only some) of the liquid nitrogen containers needed to be refilled.
Complement: The others were normal but I decided to order more anyway.
Cancelation: In fact, they all did which confused me because we hadn't used a lot this morning.

(*Appendices continue*)

Full: The PTA picked me to organize every award for the assembly at my daughter's school.
Partial: My daughter told me about the awards assembly at her school.
Trigger: (Some/Only some) of the honors students received prizes.
Complement: The others were not invited to attend because the teachers didn't want anyone to feel left out.
Cancelation: In fact, they all did because the teachers didn't want any of them to feel left out.

Full: As the new librarian, it was my responsibility to catalog every book in the reference section.
Partial: The librarian told me about how annoying it was to reorganize the reference section.
Trigger: (Some/Only some) of the dictionaries were labeled incorrectly.
Complement: The others were labeled appropriately though a few of them had been shelved in the wrong place.
Cancelation: In fact, they all were which made his job much more difficult.

Full: I am a huge fan of my old high_school football team and attended every game last season.
Partial: I ran into a friend who started telling how my old high_school football team did last season.
Trigger: (Some/Only some) of their losses were close.
Complement: The others were blowouts which made the games boring to watch.
Cancelation: In fact, they all were which made the games stressful to watch.

Full: To prepare for my Spanish test, I spent hours studying the new vocabulary items.
Partial: I had little time to prepare for my Spanish test so I quickly glanced at the new vocabulary items.
Trigger: (Some/Only some) of the words sounded like they do in English.
Complement: The others were totally unfamiliar which made the test somewhat challenging.
Cancelation: In fact, they all did which made the test somewhat easier.

Full: While eating breakfast, I had lots of time to pore over the stock prices from yesterday.
Partial: While eating breakfast, I had little time to skim the stock prices from yesterday.
Trigger: (Some/Only some) of my stocks went up.
Complement: The others were relatively steady but my portfolio was still down for the year.
Cancelation: In fact, they all did but my portfolio was still down for the year.

Appendix B

Random Effects Specifications for Mixed Effect Analyses

Mixed effects models that do not include contributing random effects can be anticonservative (Baayen, 2008), while models including noncontributing random effects can overfit. The appropriate random effects structure was determined by iteratively comparing models with increasingly more complex random effects using a likelihood ratio test. The model chosen was the most complex model that provided a significantly better fit than all simpler models (Baayen, Davidson, & Bates, 2008; discussion at R-Language mailing list from February 20, 2011). Random effects for the analyses below are described in Appendix B. For models without random slopes, significance values for fixed effects were obtained using the function pvals.fnc (Baayen, 2008). This function uses Markov Chain Monte Carlo (MCMC) simulation to approximate the posterior distribution over possible values of the model parameters and reports the proportion of the distribution that crossed zero. MCMC sampling is not currently available for models with random slopes. Significance values for models with random slopes were determined by using the associated *t* statistic with degrees of freedom equal to the number of observations less the number of degrees of freedom in the fitted model (see Levy, in press, Chapter 8). Below are the maximal random effects structures justified by the data for the models used in the present study. 1 represents random intercepts, K represents random effects of speaker knowledge, and T represents random effects of trigger type. K * T indicates all interactions and main effects involving K and T (see Tables B1, B2, and B3).

(*Appendices continue*)

Table B1
Trigger Sentences

Data type	Focus: (Only)	| Quantifier: | S/some of	| Region 2: the | real-estate	| Region 3: | investments lost	| Region 4: | money
Full data					
Participants	n/a	1	1 + K * T	1	1 + T
Items	n/a	1	1 + T	1	1
Scalar trigger data					
Participants	n/a	1	1	1	1
Items	n/a	1	1	1	1
Focused trigger data					
Participants	1	1	1 + K	1	1
Items	1	1	1 + K	1	1

Table B2
Complement Sentences

Data type	Anaphor: The rest	| Predicate: | were successful	| Clause break: | in spite of	| Next two words: | the recent
Full data				
Participants	1	1 + K * T	1	1
Items	1	1 + K + T	1	1
Scalar trigger data				
Participants	1	1 + K	1 + K	1
Items	1	1	1	1
Focused trigger data				
Participants	1	1 + K	1	1
Items	1	1	1	1

Table B3
Cancelation Sentences

Effect type	Region 1: In fact	| Region 2: | they all	| Region 3: | did	| Region 4: | because of
Random effects				
Participants	1	1 + K	1	1
Items	1	1	1	1 + K

Received August 19, 2011
Revision received February 21, 2012
Accepted February 22, 2012 ■