

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/102009>

Please be advised that this information was generated on 2021-09-19 and may be subject to change.

Speech Act Recognition in Conversation: Experimental Evidence

Rosa S. Gisladdottir (rosa.gisladdottir@mpi.nl)

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
International Max Planck Research School for Language Sciences, Nijmegen, The Netherlands

Dorothee J. Chwilla (d.chwilla@donders.ru.nl)

Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Herbert Schriefers (h.schriefers@donders.ru.nl)

Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Stephen C. Levinson (stephen.levinson@mpi.nl)

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Abstract

Recognizing the speech acts in our interlocutors' utterances is a crucial prerequisite for conversation. However, it is not a trivial task given that the form and content of utterances is frequently underspecified for this level of meaning. In the present study we investigate participants' competence in categorizing speech acts in such action-underspecific sentences and explore the time-course of speech act inferencing using a self-paced reading paradigm. The results demonstrate that participants are able to categorize the speech acts with very high accuracy, based on limited context and without any prosodic information. Furthermore, the results show that the exact same sentence is processed differently depending on the speech act it performs, with reading times starting to differ already at the first word. These results indicate that participants are very good at "getting" the speech acts, opening up a new arena for experimental research on action recognition in conversation.

Keywords: Action; Speech acts; Implicature; Pre-Offers; Conversation Analysis; Self-paced reading.

Introduction

Knowing a language doesn't just require syntax or semantics, but the ability to extract speech acts from our interlocutors' utterances. This is crucial in conversation, since all actions – be they non-verbal or verbal – have implications for how we should respond (Schegloff, 2007): A greeting calls for another greeting, an offer is followed by an acceptance or declination. Scholars in Conversation Analysis were the first to reveal the systematicity of courses of action in turn-taking (e.g. Sacks, Schegloff, & Jefferson, 1974; Schegloff, 2007), moving away from the single act as the fundamental unit of analysis – the perspective of Speech Act theory (Austin, 1976; Searle, 1969) – to the role of sequential context. One of the main observations from this literature is that turns tend to come in pairs of actions and in such *adjacency pairs* the first (pair) part sets up powerful constraints on what type of action can follow (Schegloff, 2007). Moreover, not all actions have equal status. Reflecting an orientation towards *preference structure*, dispreferred actions, such as rejections to invitations, tend to

be delivered with inter-turn gaps, turn-initial delay, hedges or other discourse markers (*uhm*, well) (Schegloff, 2007), while preferred actions (e.g. acceptances) do not. This suggests that conversationalists not only monitor speech for actions but also orient to sequential constraints.

Given that action is the *sine qua non* of conversation, how do we map speech acts onto our interlocutors' utterances, bridging the gap between sentence meaning and action? In some cases this is a simple matter. In the utterance *Please close the door*, for instance, the imperative mood and the adverb *please* function as "special markers" or "illocutionary force indicating devices" (Levinson, 1983; Clark, 1979; Schegloff, 2007) that clearly indicate this is a request. In most utterances, however, the absence of such dedicated vocabulary leaves the propositional content underspecified for the speech act level of meaning. As an example, an assertion like *I have a credit card* can deliver different speech acts, depending on context. When responding to a question from our interlocutor (e.g. *How are you going to pay for the ticket?*), *I have a credit card* functions as an information-giving answer. If it follows an offer of payment (e.g. *I can lend you money for the ticket*), it is used to indirectly decline it. In this case it could be characterized as an indirect speech act, in which "one illocutionary act is performed indirectly by way of performing another" (Searle, 1975). In yet another interchange, if our interlocutor has expressed the need or desire for some means of payment (e.g. *I don't have any money to pay for the ticket*), the same statement of ownership can function as a prelude to an offer, called a Pre-Offer in Conversation Analysis (Schegloff, 1988; Schegloff, 2007). In all three cases, the form and semantic content is underspecified for the action such that the full import of the utterance *I have a credit card* can only be ascertained relative to the context, in this case the prior speech act in the conversation.

There is some psychological evidence that people do extract speech act information online. Using a recognition probe task and lexical decision task, Holtgraves (2008a) addressed whether the comprehension of a sentence like

Don't forget to go to your dentist (an “implicit speech act”) entails automatic activation of the speech act performed (reminding). He found that the recognition of such speech acts is automatic, both in written and spoken utterances. A further study (Holtgraves, 2008b) suggests that people recognize and retain in long-term memory the actions that people perform with their utterances. In line with Speech Act Theory and Conversation Analysis, Holtgraves (2008a) argues that “in conversation there is an action dimension, a dimension that does not exist for isolated sentences or texts. Speakers are usually constructing utterances with the intention to perform certain actions, and with the intention of having the recipient recognize those actions” (p. 640).

Clearly, action recognition crosscuts research on topics such as communicative intention and implicature in Pragmatics (Grice, 1975; Levinson, 1983; Sperber & Wilson, 2004), the study of indirect speech acts (e.g. Gibbs, 1979; Clark, 1979; Clark, 1996; Coulson & Lovett, 2010) and discourse processing (e.g. Graesser, Singer, & Trabasso, 1994) in Psycholinguistics, as well as a more general discussion of action and theory of mind in the cognitive sciences. There is limited experimental research, however, on speech act recognition in spoken dialogue. The experimental approach used in Holtgraves (2008a, 2008b) involves artificial tasks (lexical decision and recognition probe) and does not unravel the time-course of action recognition. The puzzle remains: how is it that we can extract speech acts from utterances so efficiently, as evidenced by extraordinarily fast turn transitions (Sacks et al., 1974; Stivers et al., 2009; Levinson, in press)? To address this we are currently planning an Event-Related Potential (ERP) study on action recognition in auditorily presented dialogues to track the time course of action comprehension in real time. The experiment presented in this paper was designed to assess the feasibility of such a study.

The Experiment

The aim of the present experiment was to investigate participants' competence in identifying speech acts in action-underspecific sentences and explore the time-course of speech act inferencing. To do this we presented target sentences using the self-paced reading paradigm and asked participants to categorize the speech acts and rate how sure they were in the categorization. In the domain of Pragmatics, self-paced reading has been used to investigate the processing of phenomena such as scalar implicatures (Breheny, Katsos, & Williams, 2006), with longer reading times (in comparison to a control) interpreted as indicating the generation of an inference. The self-paced reading paradigm allows us to obtain information on the word-by-word processing of action-underspecific utterances, thereby exploring the time-course of speech act inferencing.

The stimuli in our study consist of a context sentence which is presented auditorily, followed by a target sentence designed to be interpreted as an Answer, Pre-Offer or Declination depending on the context (see Table 1). These

actions are commonly found in conversation and their form and function has been described in the conversation analytic literature.

Table 1: Examples of stimuli in Dutch and translations.

Condition	Context	Target Sentence
Answer	Hoe ga je voor het ticket betalen? <i>How are you going to pay for the ticket?</i>	Ik heb een creditcard. <i>I have a credit card.</i>
Declination	Ik kan je wat geld lenen voor het ticket. <i>I can lend you money for the ticket.</i>	Ik heb een creditcard. <i>I have a credit card.</i>
Pre-Offer	Ik heb geen geld om het ticket te betalen. <i>I don't have any money to pay for the ticket.</i>	Ik heb een creditcard. <i>I have a credit card.</i>

The Answers in our study complete an adjacency pair by responding to a *wh*-question in the first turn. This condition serves as a benchmark for inferencing in the reading time analysis since the gap between literal (sentence) meaning and the action intended is the smallest. Moreover, since the other actions in the study can superficially be viewed as Answers, because they respond to the prior turn, this condition provides a check on whether participants go beyond a simple characterization of the sentences and identify the correct speech act.

The Declinations satisfy an adjacency pair by responding to a proposal (an offer or invitation) in the first turn, but require a somewhat complex inference to infer the action. Conversation analysts have noted that, at least in English, such indirect responses “need not be polite, nor unclear or obfuscatory. For certain activities, in specific sequential locations, responding indirectly may be the most efficient form of communication,” (Walker, Drew, & Local, 2011, p. 17).

The third action is the Pre-Offer, which is a type of pre-sequence. Pre-sequences are preliminary to, or project, the main course of action - in this case an offer (Schegloff, 2007; Schegloff, 1988), demonstrated in the following example:

Bookstore, 2.1: 107 (modified from Schegloff, 2007, p. 35)

- 1 A: I'm gonna buy a thermometer though because I
 2 B: but
 3 A: think she's got a temperature.
 4 C: *Pre-Offer* we have a thermometer.
 5 A: *Go-ahead* yih do?
 6 C: *Offer* wanta use it?
 7 A: *Acceptance* yeah.

Only if the response to the Pre-Offer is positive (line 5) is the offer put forward (line 6). This strategy allows conversationalists to check whether an offer would be welcome or not, preventing them from embarrassment that would arise if an offer were to be rejected.

Crucially, the Pre-Offers differ from the Answers and Declinations in that they do not complete an adjacency pair but rather open up or project a continuation of the sequence (with the response and possibly a subsequent offer). By including Pre-Offers in our study we can not only investigate the impact of sequential context on processing, but also explore whether the distinction between projection (Pre-Offers) and a backward directed inference (Declinations) is borne out in reading times.

Given that the same utterance can be used as an Answer, Declination or Pre-offer depending on the sequential context, in this study we investigate: 1) Can participants reliably categorize action-underspecific speech acts? 2) Does the time-course of speech act inferring differ for these actions as reflected in self-paced reading times? Due to the exploratory nature of the study and lack of research in this area – in particular on Pre-Offers – we do not make specific predictions regarding reading times. However, we speculate that the reading time pattern of Pre-Offers and Declinations may differ relative to Answers, based on the structural properties described above.

Methods

Participants 39 native speakers of Dutch were recruited from the student population in Nijmegen, The Netherlands. Participants were paid 8 euros for participating.

Materials and Design The stimulus materials were 126 target sentences, presented visually one word at a time (self-paced reading), which were preceded by 378 auditory context-setting utterances that biased the interpretation of the target as an Answer, Pre-offer or Declination. To maintain a balance of variety and control in the stimulus materials, half of the target sentences started with the pronoun “I” (Dutch *ik*) and the verb “have” (*heb*), e.g. “I have a credit card”. The other half was more varied (including simple utterances like “I am going to the market” and “My brother is a mechanic”). We varied word-length to make the stimuli as natural as possible, but constructed the target sentences such that the final word is critical for understanding the propositional content of the utterance (irrespective of speech act level meaning). In line with reported characteristics of indirect replies (Walker et al., 2011), the target sentences do not involve ellipsis or pronominalization.

To maintain consistency in the way the Declinations and Pre-offers are connected to their contexts, we ensured that there is at least one clear *implicated premise* and an *implicated conclusion* for each sentence-pair: when presented with an utterance that is indirect, the hearer needs to access an implicated premise and combine it with the

proposition expressed to derive the implicated conclusion (Blakemore, 1992).¹

In order to get a measure of the semantic relatedness between context and target sentence in each condition, Latent Semantic Analysis (LSA) values (Landauer, Foltz, & Laham, 1998) were computed for the English translation of each sentence pair using document-to-document mode with “General reading up to 1st year of college” as the semantic space. The average LSA values for each condition were: Answers 0.13, Declinations 0.32, and Pre-Offers 0.42 (the higher the value, the more semantic similarity).

The stimuli were translated from English into Dutch and checked by two native speakers of Dutch. The sentences were recorded by four native speakers. The recordings of the target sentences were not used in this experiment, since they are presented visually in self-paced reading.

The stimuli were pseudo-randomized and balanced across three lists, such that participants saw each target sentence only once, in one context. After each trial (sentence pair), participants were given a comprehension and rating task. They were first asked to indicate what the second speaker was doing with his response and were given the options of Answering, Offering and Declining (D. *antwoorden, aanbieden, weigeren*). Since Pre-Offer is not a colloquial term, the broader term of offering was chosen. Participants were then asked to rate how sure they were in their categorization decision on a rating scale from 1 (very uncertain) to 7 (very certain). The purpose of the rating task was to assess the feasibility of using the items in future studies.

Procedure Participants were given instructions that included one example of each action. They were instructed to imagine that they were listening to a conversation between friends or colleagues, and to read the sentences as quickly as possible, but not too quickly as they would have to “judge the underlying meaning” of the sentences. They were then seated in a chair in front of a monitor in a soundproof experimental booth. On each trial the context sentence was played while a small picture of a loudspeaker was presented at the middle of the screen. 500 msec after the end of the spoken sentence participants were presented with the target sentence in a moving window self-paced reading format (Just, Carpenter, & Woolley, 1982). A series of lines appeared on the screen representing each word in the target sentence. When participants clicked on the mouse the first word appeared and upon subsequent button presses a new word was shown, while the previous word was again replaced by a line. When participants clicked the mouse after the last word had been shown, they were presented with the action categorization question, immediately followed by the certainty rating. There were 126 experimental trials, preceded by a brief practice session.

¹ In the dialogue (A): *I can lend you money for the ticket.* – (B): *I have a credit card*, the implicated premise is that a credit card can pay for things, including tickets. The implicated conclusion is that speaker B does not need A’s help with paying for the ticket.

Results

Accuracy Overall accuracy (number of correct responses in the action categorization question divided by the total number of responses) was very high, 95.8 percent. Accuracy percentages (summarized in Table 2) were very similar across conditions. A repeated measures ANOVA revealed that accuracy was not affected by action [$F(2, 76) = .07, p = .93$].

Table 2: Accuracy and mean certainty ratings.

	Accuracy	Certainty Ratings	
		Mean	SD
Answer	96.0%	6.60	0.36
Declination	96.0%	6.50	0.39
Pre-Offer	95.6%	6.35	0.51

Ratings Participants rated how certain they were in answering the action categorization question on a scale from 1 (very uncertain) to 7 (very certain). The overall mean certainty rating was 6.48. Mean certainty ratings for each condition are summarized in Table 2. We conducted a repeated-measures ANOVA on the mean ratings and found an effect of action [$F(1.42, 53.99) = 11.20, p < .01$ (Greenhouse-Geisser)]. Pairwise comparisons using Sidak adjustment for multiple comparisons revealed that Pre-Offers ($M=6.35, SD=.51$) had lower ratings than Answers ($M=6.60, SD=.36$), [$p < .01$], and Pre-Offers were also rated lower than Declinations ($M=6.50, SD=.39$), [$p < .01$]. The comparison between Answers and Declinations was not significant, [$p = .15$].

Reading Times The time between button presses was recorded as the reading time for each word. Extreme values below 100 msec were excluded, as well as values above 1200 msec for non-final words and above 7000 msec for final words (in total 12 outliers). Since online speech comprehension and the subsequent off-line categorization task tap different types of information, error trials were not excluded from the reading time analysis.

Mean reading times for the first word, the verb and the final word of the target sentences were used for the analysis, in addition to the mean reading time of the entire sentence and mean reading time per word (sentence reading time divided by number of words). A repeated-measures ANOVA was carried out to examine the effect of action (Answer, Pre-Offer, Declination) on reading times and post-hoc comparisons were performed using the Sidak adjustment. The conditions started to differ already at the first word, with mean reading time being affected by the action manipulation [$F(1.69, 64.26) = 5.24, p = .01$ (Greenhouse-Geisser)]. Post-hoc comparisons revealed that first word reading times in Pre-Offers ($M=259, SD=65$) were longer than in Answers ($M=251, SD=56$), [$p = .03$], and Pre-Offers tended to be longer than Declinations ($M=252, SD=57$), [$p = .06$]. The comparison between Answers and Declinations was not significant [$p = .98$].

Table 3: Mean reading times in msec.

		Answer	Declination	Pre-Offer
First Word	Mean RT	251	252	259
	SD	56	57	65
Verb	Mean RT	260	267	265
	SD	60	69	70
Final Word	Mean RT	564	622	593
	SD	384	447	412
Word ²	Mean RT	339	354	352
	SD	135	146	147
Entire Sentence	Mean RT	1459	1528	1501
	SD	584	652	603

At the verb, there was also an effect of action on mean reading time [$F(2, 76) = 3.41, p = .04$]. Pairwise comparisons revealed that verb RTs in Declinations ($M=267, SD=69$) were longer than in Answers ($M=260, SD=60$), [$p = .048$]. The comparison between Pre-Offers and Answers [$p = .25$] was not significant, nor between Pre-Offers and Declinations [$p = .80$].

At the final word, mean reading times also differed between actions, [$F(1.63, 61.83) = 4.28, p = .03$ (Greenhouse-Geisser)]. Pairwise comparisons on the final RTs revealed that Declinations ($M=622, SD=447$) were marginally significantly longer than Answers ($M=564, SD=384$) [$p = .052$], while there were no differences between Answers and Pre-Offers [$p = .17$], nor Pre-Offers and Declinations [$p = .41$].

An ANOVA on mean RTs per word revealed an effect of action [$F(2, 76) = 5.73, p = .01$]. Declinations ($M=354, SD=146$) took longer than Answers ($M=339, SD=135$) [$p = .02$], and Pre-Offers ($M=352, SD=147$) took longer than Answers ($M=339, SD=135$), [$p = .04$]. There was no difference between Pre-Offers and Declinations [$p = .95$].

Finally, although the RTs for the entire sentence differed descriptively, these differences were not reliable [$F(2, 76) = 2.73, p = .09$].

Discussion

The present experiment demonstrates that participants categorize the speech acts of sentences whose form and semantic content is underspecified for action with very high accuracy (95.8%). They are able to do so based on limited context (the prior speech act) and without any prosodic information in the target sentence. Importantly, the accuracy was the same for all three actions (Answer, Declination, Pre-Offer). If participants had processed the target sentences superficially, ignoring the speech act content, they could have categorized Declinations and Pre-Offers as Answers.

² RT of the entire sentence divided by number of words.

This is the case since the Dutch term for answering (*antwoorden*) also means *to respond* and all three speech acts can superficially be seen as responses. This should have resulted in lower accuracy for Pre-Offers and Declinations vis-à-vis Answers. The high accuracy rate across actions shows that participants go beyond a simple characterization of the target sentences as responses and “get” the correct action. Participants were also very confident in categorizing all actions and rated the certainty of their categorizations on average 6.48 (out of 7). These results provide further support that participants orient to the action content of sentences.³

The reading time results demonstrate that the exact same sentence is processed differently depending on the speech act it performs. In all conditions the reading time increased throughout the sentence, but Declinations and Pre-Offers had different trajectories relative to Answers, which had shortest RTs on all measures. Reading times differed already at the first word, with first word RTs in Pre-Offers being longer than in both Answers and Declinations. It should be pointed out that descriptively the difference between the means is small and standard deviation large. Event-related brain potentials might be a more sensitive measure to reveal processing differences at early positions in the sentence. At the verb and the final word however, RTs were longest in the Declination condition.

What could explain the different reading times across conditions? One possible source of difference is the *amount* of inferencing required. In text processing, reading times are predicted to be the longest for words in the text that generate “many online inferences” (Graesser, Swamer, & Baggett, 1996). The fact that Declinations have the longest RTs at the final word may be because *more* inferences (e.g. Gricean implicatures) are needed to relate the sentence to the prior context.

Another source for differences may be the *kind* of inferencing required. Given the exploratory nature of this study we did not make predictions regarding reading time results. However, based on differences in sequence organization, we speculated that Declinations and Pre-Offers would exhibit different reading time patterns relative to Answers. While recognizing the speech act in a Declination requires computation of how the utterance can be understood as the second pair part to the prior proposal, identifying a Pre-Offer involves knowing that an offer will follow in the sequence. Because Declinations close an adjacency pair, they do not heavily constrain the relevant next action. Pre-Offers, on the other hand, call for either go-

ahead (e.g. *you do?*) and a subsequent offer, or a blocking response (*that’s ok*). Pre-Offers, therefore, invite stronger predictive inferences about the next speech act. The distinction between an inference based on a backward bridge to the prior turn in an adjacency pair and an inference based on forward projection of a sequence is akin to the difference between causal antecedent and causal consequence inferences in text processing (e.g. Magliano, Baggett, Johnson, & Graesser, 1993). The reading time differences between Pre-Offers and Declinations provide some indication that the distinction between forward projection and a backward directed inference plays a role in the online processing of speech acts. Why the projective nature of Pre-Offers would call for more processing at the first word, however, is less clear. More research is needed to investigate whether this finding holds for spoken language processing as well.

Considering that action comprehension is crucial for everyday conversation, it seems likely that people can predict upcoming speech acts, making the fast transitions between turns (Sacks et al., 1974; Stivers et al., 2009; Levinson, in press) possible. Neuroimaging studies suggest that people use their knowledge of the wider discourse context to predict specific upcoming words and that prediction is not the result of relatively low-level, word-based priming mechanisms, “but involves a more sophisticated message-level mechanism that can take into account the actual nuances of the preceding discourse,” (Otten & Van Berkum, 2008). Whether and how sequential context and the implicit knowledge of the organization of actions guides the interpretation of utterances is a topic for further investigation. We will explore this in future research using event-related brain potentials.

An alternative explanation for the reading time results is that the experimental manipulation does not address speech act recognition per se, but some other confounding variable such as semantic priming from the context. Latent Semantic Analysis can be used to determine semantic relatedness of two texts and LSA similarity relations have been found to correspond well with the pattern of results in priming studies (Landauer et al., 1998). If semantic priming from the context is the main factor governing the reading times one would expect the condition with the lowest LSA values (least amount of priming) to have the longest mean reading times. The opposite is true: Answers had the lowest average LSA value but the shortest reading times on all measures. This suggests that the differences in reading times across conditions in our study were not due to lexico-semantic relationships between the content words of the context and the target sentences.

Conclusion

In this study on speech act comprehension we investigated the processing of sentences that perform different speech acts depending on prior context. In each case an assertion is used as a vehicle for some other action, and it is “part of competent membership in the society/culture and being a

³ Note that the certainty ratings were lower for Pre-offers (6.35) than Answers (6.60) and Declinations (6.50). Since “pre-offer” is not a colloquial term, participants were instructed to categorize Pre-offers as “offering”. The ratings may reflect participants’ awareness that the speech acts they were instructed to label as “offering” are preparatory to, or project, offers and in most cases do not constitute offers in themselves. However, for the purpose of this study it was sufficient that participants could identify speech acts as belonging to one of the defined categories.

competent interactant to analyze assertions of this sort for what (else) they may be doing at this moment, at this juncture of the interaction, in this specific sequential context” (Schegloff, 2007, p. 35). Our study tapped into this competence by addressing two primary questions: how reliably participants can categorize action-underspecific speech acts, and whether the time-course of speech act inferencing differs for the actions as reflected in self-paced reading times.

Participants in our study categorized the speech acts of action-underspecified utterances with very high accuracy, based on limited context and without any prosodic information. Furthermore, the exact same sentence was processed differently depending on the speech act it performed, with reading times starting to differ already at the first word. These findings open up a new arena for experimental research on speech act recognition in conversation.

As a crucial component of social behavior, communication involves actions. Being a competent member of society must require a cognitive architecture that is oriented to speech acts. However, given that the form and content of utterances is frequently underspecified for this level of meaning, assigning speech acts to our interlocutors is not a trivial matter. Having demonstrated that participants orient to the action content of sentences and can categorize speech acts with high accuracy, the next experimental step is to shed light on this ability in spoken dialogues – the foundation of doing things with words.

References

- Austin, J. L. (1976). *How to do things with words*. Oxford: Oxford University Press.
- Blakemore, D. (1992). *Understanding utterances: An introduction to pragmatics*. Oxford: Blackwell.
- 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(3), 434–463.
- Clark, H. H. (1979). Responding to indirect speech acts. *Cognitive Psychology*, 11(4), 430–477.
- Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Coulson, S., & Lovett, C. (2010). Comprehension of nonconventional indirect requests: An event-related brain potential study. *Italian Journal of Linguistics*, 22(1), 107–124.
- Gibbs, R. W. (1979). Contextual effects in understanding indirect requests. *Discourse Processes*, 2(1), 1–10.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), 371–395.
- Graesser, A. C., Swamer, S. S., & Baggett, W. B. (1996). New models of deep comprehension. In B. Britton & A. C. Graesser (Eds.), *Models of understanding text*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, 3: Speech Acts*. New York: Academic Press.
- Holtgraves, T. (2008a). Automatic intention recognition in conversation processing. *Journal of Memory and Language*, 58(3), 627–645.
- Holtgraves, T. (2008b). Conversation, speech acts, and memory. *Memory & Cognition*, 36(2), 361–374.
- Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2), 228–238.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3), 259–284.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, S. C. (in press). Action formation and ascription. In T. Stivers & J. Sidnell (Eds.), *Handbook of conversation analysis*. Malden, MA: Wiley-Blackwell.
- Magliano, J. P., Baggett, W. B., Johnson, B. K., & Graesser, A. C. (1993). The time course of generating causal antecedent and causal consequence inferences. *Discourse Processes*, 16(1-2), 35–53.
- Otten, M., & Van Berkum, J. J. A. (2008). Discourse-based word anticipation during language processing: Prediction or priming? *Discourse Processes*, 45(6), 464–496.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735.
- Schegloff, E. A. (1988). Presequences and indirection: Applying speech act theory to ordinary conversation. *Journal of Pragmatics*, 12(1), 55–62.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Searle, J. R. (1975). Indirect speech acts. In P. Cole & J. Morgan (Eds.), *Syntax and semantics, 3: Speech Acts*. New York: Academic Press.
- Sperber, D., & Wilson, D. (2004). Relevance theory. In G. Ward & L. Horn (Eds.), *Handbook of pragmatics*. Oxford: Blackwell.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K.-E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), 10587–10592.
- Walker, T., Drew, P., & Local, J. (2011). Responding indirectly. *Journal of Pragmatics*, 43(9), 2434–2451.