Journal of Memory and Language 84 (2015) 49–74

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Interpretation of informational questions modulated by joint knowledge and intonational contours

Sarah Brown-Schmidt [a,*], Scott H. Fraundorf [b]

[a] University of Illinois at Urbana-Champaign, United States
[b] University of Pittsburgh, United States

ARTICLE INFO

Article history:
Received 6 May 2014
revision received 13 May 2015
Available online 4 June 2015

Keywords:
Perspective-taking
Common ground
Eye-tracking
Prosody
Interactive
Executive function

ABSTRACT

We examine processes by which dialogue partners form and use representations of joint knowledge, or *common ground*, during on-line language processing. Eye-tracked participants interpreted wh-questions that inquired about task-relevant objects during interactive conversation. Some objects were known to both speaker and listener, and thus in *common ground*, whereas others were only known to the listener, and thus in *privileged ground*. Questions were produced with a typical, falling intonation (Experiment 1) or with either falling or rising intonation (Experiments 2–3). Unlike the falling contour, the rising contour can indicate a request for clarification about previously mentioned information. Participants interpreted falling-contour questions as asking about privileged-ground objects. By contrast, rising questions elicited more consideration of common-ground objects. Directly comparing questions that were produced during live conversation vs. questions that were pre-recorded revealed that this sensitivity to common vs. privileged ground emerged only during live conversation. Finally, individual difference analyses in all three experiments did not support the claim that individuals fail to take perspective when executive function is limited. Taken together, these findings provide evidence for the on-line integration of perspective and intonation during conversational language processing. The lack of perspective effects in non-interactive settings speaks to the inherently interactive nature of conversational processes.

Introduction

Everyday language use is shaped by representations of which information is and is not jointly shared with a conversational partner. Consider, for example, that asking an informational question, such as *What's in the box?*, presumes that the speaker does not know what is in the box but that the addressee does know. According to classic theories of conversational language, representations of joint knowledge, or *common ground*, serve as the fundamental context for language use (Clark, 1992, 1996). Studies of language production in task-based conversation largely support this claim, showing that speakers adjust the specificity of their referential expressions based on the knowledge the addressee holds jointly with the speaker (Clark & Wilkes-Gibbs, 1986; Isaacs & Clark, 1987; Wilkes-Gibbs & Clark, 1992).

Studies of the use of common ground during the on-line comprehension of language, however, tell a more mixed story. Some research shows dramatic failures in the ability of listeners to appreciate the perspective of the speaker in real time (Converse, Lin, Keysar, & Epley, 2008; Keysar, Lin, & Barr, 2003), but other research finds that listeners do readily adjust language processing based on the speaker's

* Corresponding author at: 603 E Daniel St., Champaign, IL 61820, United States.
E-mail address: sarahbrownschmidt@gmail.com (S. Brown-Schmidt).

http://dx.doi.org/10.1016/j.jml.2015.05.002
0749-596X/

perspective (Hanna, Tanenhaus, & Trueswell, 2003; Heller, Grodner, & Tanenhaus, 2008). Thus, a central goal for language processing research is understanding the factors that lead individuals to fail or succeed at appreciating the perspective of another person as they process speech.

Failures to appreciate the perspective of one's conversational partner are taken as evidence in support of egocentric-first accounts of language processing. In one version of this view, perspective representations are not routinely incorporated into language comprehension processes. This view is supported by findings that when asked for, e.g., "the tape," addressees frequently reached for a roll of tape that was hidden in a bag out of the speaker's view despite the fact that a cassette tape was in full view of both speaker and addressee (Keysar et al., 2003; also see Keysar, Barr, Balin, & Brauner, 2000). In a related egocentric view, it is proposed that listeners *anticipate* that speakers will reference common ground entities, but they fail to use the distinction between common and privileged ground during *on-line* interpretation of words. Evidence consistent with the anticipation-only view includes findings that listeners hearing pre-recorded instructions like "Click on the bucket" looked at a common ground target (a bucket) more than a privileged ground competitor (a buckle) *prior* to the word "bucket." However, *during* the word "bucket," this common ground preference did not increase over baseline (Barr, 2008; also see Barr & Seyfeddinipur, 2010).

These egocentric accounts contrast with constraint-based accounts of perspective-taking, in which common ground is thought to be one of many partial constraints that guide language processing (Hanna et al., 2003). On this view, these apparent failures to take perspective into account reflect situations in which competing cues push language comprehension toward the perspective-inappropriate response (see Brown-Schmidt & Hanna, 2011). For example, interpreting "tape" as a reference to a privileged-ground roll of tape, rather than a common-ground cassette tape (Keysar et al., 2003), may be largely driven by the dominance of the roll-of-tape meaning of the word "tape" over the cassette-tape meaning. Perspective effects may be particularly weak in situations when the evidence for common ground is weaker. The strength of evidence for common ground can vary (Clark & Schaefer, 1989) such that the degree to which something is considered joint knowledge varies in a gradient fashion (Brown-Schmidt, 2012). Evidence for joint knowledge may be weaker, for instance, when a listener is not an interactive conversational participant (Brown-Schmidt, 2009a; Schober & Clark, 1989): On the classic view of common ground, an interactive process between conversational partners establishes what is joint knowledge (Clark & Wilkes-Gibbs, 1986). According to Brown-Schmidt (2009a), common ground is thus more likely to guide language processing in truly interactive settings where conversational partners can provide evidence of understanding (e.g., by saying, "OK, got it").

The central goal of the present research is to understand why common ground sometimes fails, and sometimes succeeds, at guiding on-line language processing. We focus on the processes by which common ground is formed and used during interactive conversation. We address two key predictions that are derived from the classic view of how common ground is formed and from constraint-based views of how common ground guides processing. First, the degree to which perspective guides on-line processing should be modulated by the interactive conversational processes that establish common ground. Second, the form of an utterance, including its syntactic (e.g., interrogative vs. declarative) and prosodic form (e.g., falling vs. rising), should be integrated with perspective during on-line processing. In what follows, we present some background on these issues, and we then describe the results of three experiments designed to test these predictions.

Formation of common ground: findings and open questions

The processes of interactive dialogue are thought to provide, through a conversational procedure known as *grounding* (Clark & Brennan, 1991), the necessary means to establish common ground. Analyses of task-based conversation show that active conversational participants develop and refine shared referential terms for hard-to-describe game-pieces, but participants who only overhear the conversation show limited learning of these conversationally-established terms (Schober & Clark, 1989). This difference has been suggested to stem from the fact that active partners, unlike overhearers, can use feedback, including head nods and back-channel remarks such as "OK," to interactively establish, or *ground*, these terms as understood. Further, interlocutors are sensitive to this difference and expect overhearers not to have learned these terms (Wilkes-Gibbs & Clark, 1992; also see Gorman, Gegg-Harrison, Marsh, & Tanenhaus, 2013; Lockridge & Brennan, 2002).

Analyses of the on-line processing of entrained terms, however, provide mixed evidence for use of partner-specific common ground (Horton & Slaten, 2012; Kronmüller & Barr, 2007; Metzing & Brennan, 2003; Shintel & Keysar, 2007), but the use of different methods makes comparisons across those studies difficult. The fact that common ground effects are not always seen on all measures emphasizes the importance of experimental manipulations that attempt to identify why common ground sometimes does, and sometimes does not, guide processing. One relevant factor might be the degree to which the interlocutors interactively established common ground. Indeed, a study that directly compared interactive and non-interactive test paradigms showed that only in live, interactive conversation did addressees show partner-specific expectations for the use of these terms (Brown-Schmidt, 2009a[1]).

These findings suggest that learning conversationally-established referring terms, and the process of attributing

[1] One study evaluated whether addressees were more sensitive to shared labels than were overhearers and concluded there was no addressee benefit during processing (Barr & Keysar, 2002, Exp. 3). However, in that study, there was no true interaction; all participants listened to pre-recorded instructions and did not actively establish common ground. Instead, "addressees" were simply told they were listening to a live partner with whom they could not interact whereas overhearers were told they were listening to recordings.