

Partner-Specific Adaptation in Dialog

Susan E. Brennan,^a Joy E. Hanna^b

^a*Department of Psychology, Stony Brook University*

^b*Department of Psychology, Oberlin College*

Received 9 April 2008; received in revised form 2 January 2009; accepted 14 January 2009

Abstract

No one denies that people adapt what they say and how they interpret what is said to them, depending on their interactive partners. What is controversial is when and how they do so. Several psycholinguistics research programs have found what appear to be failures to adapt to partners in the early moments of processing and have used this evidence to argue for modularity in the language processing architecture, claiming that the system cannot take into account a partner's distinct needs or knowledge early in processing. We review the evidence for both early and delayed partner-specific adaptations, and we identify some challenges and difficulties with interpreting this evidence. We then discuss new analyses from a previously published referential communication experiment (Metzing & Brennan, 2003) demonstrating that partner-specific effects need not occur late in processing. In contrast to Pickering and Garrod (2004) and Keysar, Barr, and Horton (1998b), we conclude that there is no good evidence that early processing has to be “egocentric,” “dumb,” or encapsulated from social knowledge or common ground, but that under some circumstances, such as when one partner has made an attribution about another's knowledge or needs, processing can be nimble enough to adapt quite early to a perspective different from one's own.

Keywords: Joint action; Audience design; Referential communication; Entrainment; Collaborative cognition

1. Introduction

Spoken dialog is a form of joint action in which interacting individuals coordinate their behavior and processing moment by moment and adapt their linguistic choices and non-verbal behavior to each other. Often this results in convergence—of word choice, conceptual perspective, syntactic form, dialect, pronunciation, speaking rate, posture, and other

Correspondence should be sent to Susan E. Brennan, Department of Psychology, State University of New York at Stony Brook, Stony Brook, NY 11794-2500, USA. E-mail: susan.brennan@sunysb.edu

behavior by both individuals. It may also result in behavior that is not convergent but complementary, as when adjustments take the specific needs, knowledge, or perspective of the partner into account. Adjustments to a partner, whether convergent or complementary, have been termed “audience design” when done by speakers; of course, addressees may adjust as well, interpreting an utterance differently depending on who produced it.

The interesting question is when and how such adaptation emerges (Schober & Brennan, 2003), that is, what underlying cognitive processes and representations give rise to partner-specific processing and behavior. Some have claimed that partner specificity can emerge rapidly and automatically in the language processing system, and that common ground established with a partner (as well as other kinds of contextual and social constraints) can be taken into account in the earliest moments of processing (Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003; Metzging & Brennan, 2003; Nadig & Sedivy, 2002). This view finds partner-specific information to be just like any other information in memory (see also Horton & Gerrig, 2005a,b; Polichak & Gerrig, 1998). Others have argued for two-stage models in which the only rapid, early processing is “egocentric,” and in which partner-specific adjustments emerge relatively later, as more effortful adjustments or repairs (Bard et al., 2000; Brown & Dell, 1987; Ferreira & Dell, 2000; Horton & Keysar, 1996; Keysar, Barr, Balin, & Brauner, 2000; Keysar, Barr, Balin, & Paek, 1998a; Keysar et al., 1998b; Kronmüller & Barr, 2007). This view assumes that partner specificity requires complex inferences about the partner’s needs, knowledge, or perspective and proposes that maintaining and updating a model of the partner is computationally expensive, so is done only when necessary (Pickering & Garrod, 2004).

Here, we evaluate issues and evidence behind both sets of claims. First, we consider some methodological issues in the search for the locus of partner-specific processing, because differences in handling these issues have the potential to lead to quite different outcomes and conclusions. Then we revisit a data set (originally collected by Metzging & Brennan [2003] in order to address whether interpretation of referring expressions is partner specific), conduct additional analyses inspired by Kronmüller and Barr (2007), and discuss the implications in light of claims about an early modular stage of egocentric processing (Keysar et al., 1998b). Finally, we consider that taking the perspective of a partner into account early in processing need not involve extensive inferences or processing resources (even when the partner’s perspective is distinct from one’s own), but it is computationally feasible when interlocutors can make simple pragmatic attributions that amount to representing partner’s knowledge, needs, or perspectives as “one-bit” models (c.f. Galati & Brennan, 2006). The picture that emerges is one in which the early moments of language processing can be flexible, nimble, and responsive to such attributions, rather than reflexive, egocentric, and “dumb.”

2. Language processing in dialog

Psycholinguistics research has been shaped by two distinct traditions, the “language-as-product” and “language-as-action” traditions (e.g., Clark, 1992; Tanenhaus & Trueswell,

2005). The product tradition has focused on discovering the set of core processes that involve recovering linguistic structure independent of context, as inspired by the competence/performance distinction laid out by Chomsky (1965, 1980); studies in this tradition have focused on production by single speakers or comprehension by single addressees, ignoring any effects of being engaged in dialog. In contrast, the action tradition has focused on language use in more realistic contexts such as those involving referential communication between pairs of people (e.g., Brennan & Clark, 1996; Clark, 1992; Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1989, 1991, 1992; Glucksberg, Krauss, & Weisberg, 1966; Krauss, 1987; Schober & Clark, 1989). Recently, the unobtrusive analysis of eye gaze during spontaneous language use in what has come to be known as the “visual worlds” paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) has been used to bridge the two traditions (Tanenhaus & Trueswell, 2005).

Studies of language processing in dialog have the potential to yield different sorts of discoveries than studies of language processing in monolog (Clark, 1992, 1997; Pickering & Garrod, 2004; Schober & Brennan, 2003), although there is fundamental disagreement as to how such differences are manifested. According to Pickering and Garrod’s (2004) “interactive alignment” view, dialog has its impact on language processing because both production and comprehension systems are engaged within the same mind at once, with parity between production and comprehension leading automatically to convergent representations at all levels of language processing. Pickering and Garrod explicitly propose a two-stage model, where only those kinds of adaptation in which one partner’s behavior converges with another’s can be achieved automatically; other kinds of adjustments to a partner’s needs or perspective (as distinct from one’s own) must be achieved effortfully and late:

...we argue that interlocutors do not need to monitor and develop full common ground as a regular, constant part of routine conversation, as it would be unnecessary and far too costly. Establishment of full common ground is, we argue, a specialized and nonautomatic process that is used primarily in times of difficulty (when radical misalignment becomes apparent). We now argue that speakers and listeners do not routinely take common ground into account during initial processing. (Pickering & Garrod, p. 179)

This view echoes two-stage proposals by Keysar, Barr, and colleagues (e.g., Horton & Keysar, 1996; Keysar et al., 1998a, 2000; Kronmüller & Barr, 2007). These proposals have argued for modularity in processing partner-specific information; that is, unless two partners’ perspectives happen to be aligned (which fortuitously enough, they often happen to be), then information about one partner’s knowledge *cannot* be incorporated into the other partner’s planning or interpretation early on, due to the reflexive influence of priming. Two-stage proposals predict that egocentric behavior will be the early default whenever two partners’ perspectives differ, and that audience design can occur only later, as an “inferential” process that involves monitoring and repair.

According to a contrasting view, language use is inherently collaborative, and so interpersonal coordination takes center stage; the incremental unfolding of utterances in

dialog is proposed to be shaped not only by information about a partner's social identity and the common ground accumulated with that partner (Clark & Marshall, 1981) but also by moment-by-moment feedback or evidence of understanding (Brennan, 1990, 1991, 2005). Recent accounts have argued that such partner-specific information constrains processing just like any other kind of contextual information in memory, and when available, can be used from the earliest moments of utterance planning or interpretation (Hanna & Tanenhaus, 2004; Hanna et al., 2003; Metzger & Brennan, 2003; Nadig & Sedivy, 2002). On this probabilistic view, occasional "egocentric" behavior and misunderstandings do not prove that language processing is egocentric, but merely that it is fallible, influenced or impaired at times by competition, distraction, overload, interference, or ambiguity.

Before weighing new evidence, we will highlight some issues associated with studies of dialog. Some studies have been criticized for reaching beyond the data, with conclusions that are either clouded by confounds or built on evidence that fails to address the timeline and resources involved with processing partner-specific cues. Others have been criticized for failing to take seriously the fact that dialog involves coordination between distinct minds. Specific issues include (1) how to distinguish interlocutors' perspectives experimentally, (2) how to make appropriate tradeoffs between ecological validity and experimental control, (3) what kind of linguistic processing is presumed to be under the influence of partner-specific cues, and (4) what kind of evidence about the availability of information is necessary to convincingly support conclusions that a process is modular.

2.1. Distinguishing speakers' perspectives from addressees' perspectives

A cogent critique (first advanced by Brown & Dell, 1987 and elaborated by Dell & Brown, 1991) is that what may *appear* to be adaptation to an addressee may simply reflect what is easiest for a speaker. Such egocentric "adjustments" should be considered to be *for the speaker* as opposed to truly *for the addressee* (for discussion, see Dell & Brown, 1991; Galati & Brennan, 2006; Keysar, 1997; Kraljic & Brennan, 2005; Lockridge & Brennan, 2002). For example, in spoken language production, repeated or predictable tokens of a word (given information) in utterances such as *a stitch in time saves nine* have shorter durations and are pronounced less clearly than the word's first or unpredictable mention (new information) in utterances like *the next number is nine* (Bard et al., 2000; Fowler & Housum, 1987; Lieberman, 1963; McAllister, Potts, Mason, & Marchant, 1994; Samuel & Troicki, 1998). Although some have presumed that this adjustment is made for communicative purposes (e.g., driven by the addressee's needs; e.g., Nootboom, 1991; Samuel & Troicki, 1998), others have argued that speakers do this egocentrically (Bard et al., 2000), and that ordinarily, since speakers' and addressees' context coincide, what is easy for the speaker is also easy for the addressee. The first challenge in designing an experiment that tests for partner-specific effects in processing, then, is that the task must put people into a situation in which their knowledge, needs, or perspectives are distinguishable from their partners' (Keysar, 1997).

2.2. *Tradeoffs*

Another challenge for experimenters is that validity should not be sacrificed for control; that is, the task must not be so unlike spontaneous dialog that the participants find themselves playing quite a different game. This is a potential concern when an experiment that aims to be about dialog resorts to a monologic task, prerecorded utterances, or confederates who do not behave naturally, or when it places participants in situations where they must cope with coincidences that bias them toward egocentricity (e.g., Keysar et al., 1998a, 2000; see Gerrig, Brennan, & Ohaeri, 2000 for discussion) or task requirements that depart from the natural perceptual co-presence (visual and/or auditory) that characterizes ordinary conversation (e.g., Horton & Keysar, 1996). Meeting this challenge is not easy because of the need for sufficient control. For instance, Brown and Dell's (1987) study failed to find evidence that speakers took an addressee's needs into account regarding whether to mention atypical instruments while retelling stories; their speakers mentioned atypical instruments more often than typical instruments, regardless of whether the addressee had an illustration of the story showing the instrument (and so already knew about it). However, the addressee was a confederate who had heard the story many times already and actually knew it better than the speakers did. In contrast, a study by Lockridge and Brennan (2002), using the same methods and materials, found that speakers mentioned atypical instruments more often than typical ones when retelling stories to naive addressees who did not have an illustration. This suggests that studies that using confederates in the addressee role may be risky, as speakers may not show adjustments to addressees who have (and so signal that they have) no actual needs (Kuhlen & Brennan, 2008; Lockridge & Brennan, 2002). Other kinds of common tradeoffs made with the goal of achieving experimental control include using imaginary addressees or prerecorded utterances, placing partners under time pressure or increased memory load, or using tasks that require them to cope with unusual contexts or ambiguous coincidences. Of course, determining which aspects of spontaneous language use can be safely simulated in an experimental setting is not a simple matter, and success in making such tradeoffs depends on what the essential nature of dialog is assumed to be.

2.3. *What aspects of linguistic representation and processing to measure*

A third challenge is that some kinds of potential adaptations during the planning of spoken utterances appear to be more subject to a partner's influence than others (Bard & Aylett, 2005; Bard et al., 2000; Kraljic & Brennan, 2005). For instance, Bard et al. (2000) found that speakers shortened referring expressions upon re-referring to the same objects, even though the second time they mentioned the objects, they were speaking to different addressees than the first time (neither addressee had heard the expressions before). This was taken as evidence for egocentricity. At the same time, speakers appeared to adjust their use of definite versus indefinite referring expressions to the needs of the addressees. On this evidence, Bard et al. (2000) proposed a variant of a two-stage model, a "dual process model," in which fast-acting processes (e.g., articulation) automatically default to being egocentric (and are encapsulated from partner-specific knowledge), while other, more inferential

processes (e.g., planning of definite expressions) proceed in parallel; any partner-specific adjustments must emerge from the latter processes. Leaving aside the issue of whether dual processes are necessary to account for this pattern of results (which we will take up again shortly), we note that this study lacked a control condition in which speakers re-used referring expressions to addressees who had heard them before; studies that included such a control have shown evidence for audience design even in articulation (Galati & Brennan, 2006; Gregory, Healy, & Jurafsky, 2001, 2002).

2.4. *Timing and availability*

A final challenge is that a speaker cannot adjust to an addressee's needs unless information about those needs is available in a timely fashion (Horton & Gerrig, 2002). Several studies' conclusions have entertained a degree of modularity despite evidence for what might be better described as a kind of coarse adjustment to a partner's needs. For instance, omission of optional function words such as the complementizer *that* in sentences such as *I knew (that) you... were going to be late* seems to be driven by whether subsequent words are activated in the mind of the speaker, even though such words sometimes reduce ambiguity for the listener (*I knew you... when you were a child*) (Ferreira & Dell, 2000). But when such sentences were spoken to live addressees rather than to a tape recorder, speakers were marginally more likely to insert the complementizer, regardless of whether there was any actual ambiguity to be avoided. Moreover, in Brown and Dell (1987), speakers were marginally more likely to mention instruments when the addressee lacked an illustration of the story than when they had an illustration. And in Kraljic and Brennan (2005) study, when utterances included prepositional phrases with attachment ambiguities (e.g., *put the dog in the basket on the star*), speakers tended to pause longer before the goal phrase, regardless of whether the utterance would actually have been ambiguous given the pragmatics of the situation. In all of these situations, the speakers' choices ended up reducing ambiguity when ambiguity existed. That these choices were not precisely tuned to addressees' actual needs suggests that the speakers may not have had the time, knowledge, or motivation to assess information about those needs, rather than that the system is necessarily encapsulated from using it (for more discussion, see Horton & Gerrig, 2002; Kraljic & Brennan, 2005).

Different degrees of attention to these challenges and success at handling them may account for some of the variability in findings and conclusions by studies of partner effects in dialog. As a case in point, the next section presents the debate initiated by Brennan and Clark (1996) and Barr and Keysar (2002), followed by an attempt to reconcile contrasting findings from two subsequent investigations by Metzger and Brennan (2003) and Kronmüller and Barr (2007).

3. **Lexical entrainment and conceptual pacts: A new look at some old data**

We focus now on a specific kind of adaptation between speakers and addressees—the perspective-taking that underlies referential communication—and survey some

contradictory evidence concerning the time course with which partner-specific information impacts the design of referring expressions. Objects, even common ones, can be referred to in many ways; a speaker's choice of a particular expression can reflect a particular perspective on the referent or propose that attention be focused on its most relevant aspect. People in conversation tend to use the same terms when they refer repeatedly to the same objects; this phenomenon has been labeled *entrainment* (Brennan & Clark, 1996; Garrod & Anderson, 1987). As discussed in Brennan and Clark (1996), entrainment could emerge from an ongoing interaction in at least three ways: (a) if interlocutors re-use the most recent referring expression that has been successful (consistent with "output-input coordination," Garrod & Anderson, 1987); (b) if interlocutors follow the strongest *precedent*, which predicts that the fastest, most automatic mapping of referring expression to referent should be the one that reflects the most available conceptualization, even if it is egocentric (consistent with proposals by Barr & Keysar, 2002; Kronmüller & Barr, 2007; Pickering & Garrod, 2004); and (c) if a pair of interlocutors establishes a *conceptual pact*, or flexible and temporary agreement to conceptualize a referent according to a particular perspective (consistent with Brennan & Clark, 1996 and Metzing & Brennan, 2003). While all three proposals address adaptive behavior in some way, only the last involves adaptation tailored to a specific partner.

A series of referential communication experiments in which pairs of naive partners matched pictures of objects compared these three proposals (Brennan & Clark, 1996). In Experiment 3, after speakers had established precedents during spontaneous conversation with a particular partner (e.g., using *penny loafer* to distinguish one shoe from several), the speakers either continued to interact with the same addressee or switched to a new one. They continued to use the over-informative terms they had entrained-upon (e.g., *penny loafer* to refer to the only shoe in a set of objects) when they continued with the same addressee, but they tended to be only as informative as necessary and switch to the unadorned basic level term (e.g., *shoe*) with the new addressee. Although the conclusion was that conceptual pacts were partner specific, Brennan and Clark acknowledged that it was not clear whether the persistence in using the over-informative term emerged from an episodic partner-specific association in memory (an expression to referent mapping that could be considered to be a rudimentary partner model) or else from feedback provided by same versus new partners about what terms were acceptable ("a conceptual pact, then, need not be represented explicitly but may emerge from the conceptual coordination of two people interacting," Brennan & Clark, p. 1490).

This theme was taken up in studies by Barr and Keysar (2002); naive addressees were first exposed to perspectives by interacting with an experimental confederate who produced scripted referring expressions, and then the addressees heard the same expressions spoken by either the confederate or else by a prerecorded voice played through an earphone. The logic was that, if entrainment was based on partner-specific conceptual pacts, addressees should be inhibited in looking at and reaching for a referent object when a familiar expression was produced by a new speaker. There was no such inhibition. In another of their experiments, addressees who had already entrained on terms (e.g., *sportscar* for a picture of a car) with the confederate experienced equal competition from lexical cohort objects

(e.g., *carnation* for a picture of a flower) when they heard the referring expression *car...*, regardless of whether it was produced by the confederate or by the prerecorded new voice. Based on these null findings, Barr and Keysar concluded that entrainment emerges from simple precedent (as a generic adjustment), and that such precedents are represented independently from any partner-specific information.

A different conclusion was reached by Metzging and Brennan (2003), who found evidence of partner specificity by having live confederate speakers “break conceptual pacts” previously entrained-upon with naive addressees. The interaction between confederate speakers and naive addressees was as spontaneous and contingent as possible, with only the critical instructions scripted (undetectably to the addressees). The addressees were told that the experiment was about how well they could follow instructions given by different people. After repeatedly matching objects such as *the shiny cylinder* with one speaker, critical trials continued with either the same speaker or a new one, who either continued with the entrained-upon expression or with a new referring expression that was equally descriptive (*the silver pipe*). Addressees’ initial looks to target objects were delayed by 286 ms on average when the familiar speaker uttered an entirely new expression but not when the new speaker uttered the same new expression. That is, when the familiar partner (inexplicably) broke a conceptual pact, addressees seemed to experience interference in mapping the new expression to the old object, perhaps searching around for an object they might have missed (although no new objects had been introduced into the display). The conclusion was that such jointly achieved perspectives are both partner specific from the early moments of processing, as well as quite flexible, since addressees were so quick to abandon precedents when interacting with a new partner. This result is incompatible with Pickering and Garrod’s (2004) alignment theory, in which precedent (but not speaker’s identity) should matter.

Finally, Kronmüller and Barr (2007) reexamined this paradigm, proposing that a two-stage model might still be warranted if the partner specific effect of abandoning a precedent occurred after an effect of precedent. To this end, they argued that Metzging and Brennan’s (2003) analysis of first looks to the target was not fine grained enough, as it did not quantify what people looked at before the first target looks, and so they set out to replicate breaking a conceptual pact in two experiments (one with additional cognitive load and one without). Several changes were made to the method, including doing away with the interaction between confederate speakers and naive subjects. Subjects were given a credible story that they were hearing prerecorded sessions among previous subjects. The prediction made by Kronmüller and Barr (2007) is that a new expression would be “preempted” from being mapped to a familiar object that was previously associated with another expression (or *precedent*), regardless of who the speaker was. Their first experiment showed more looks to the target than to other objects (a “target advantage”) that began in the early 300–600 ms time window along with strong early effects of precedent shortly thereafter (reliable in the 600–900 and 900–1200 ms intervals), and it failed to show any partner effects (main or interaction) until the 1500–1800 (marginal) and 1800–2100 (reliable) intervals. Their second experiment, with fewer objects in the displays, found earlier precedent and partner effects, but no partner effects when speakers had to cope with additional task loads. On the basis of this evidence, Kronmüller and Barr concluded that there are “distinct processing systems”

for precedent-specific information (old vs. new expressions) and for partner-specific (inferential) information.

This conclusion seems unsatisfying for several reasons. First, in Kronmüller and Barr's method, new (previously unmentioned) objects were inexplicably introduced into the arrays not long before the instructions with the new referring expressions were heard. It is entirely possible that the new objects induced a different pattern of looking that had nothing whatsoever to do with the spoken expressions; alternatively, the new, previously unmentioned objects may have led to an exaggerated early effect of precedent. In studies of visual attention, novel stimuli have been shown to draw attention (Johnston, Hawley, Plewe, Elliott, & DeWitt, 1990), especially in the situation when a new object appears among familiar objects (Yang & Zelinsky, 2006). Second, according to Brennan & Clark (1996), a conceptual pact is a flexible agreement between speakers to take a perspective on an object that is then marked by their re-using the same term to refer to it; a pact is established through contingent interaction between the speakers. Kronmüller and Barr's prerecorded instructions afforded not even a pretense of interaction between subjects and speakers, which may have had the result of attenuating, delaying, or eliminating any speaker-specific effects. So it is unclear as to whether Kronmüller and Barr's conclusions can be compared to Metzinger and Brennan's, because on the one hand, Kronmüller and Barr removed any potential for interaction and thus reduced the salience or utility of taking speakers' perspectives, and on the other, introduced a sharp contrast between old and new objects that could have swamped any partner-specific effects.

Inspired by Kronmüller and Barr's (2007) enterprise and by the desire to better understand the extent to which their findings compare with Metzinger and Brennan's, we revisited the latter data set, calculated and graphed target advantages from the onset of the critical referring expressions, and tested the reliability of unfolding differences due to precedent and partner identity (see supplemental materials at <http://www.cogsci.rpi.edu/CSJarchive/Supplemental/index.html>). Following Kronmüller and Barr (2007), we calculated the "target advantage" at each frame (every 33 ms) as the proportion of looks at the target minus the proportion of looks at other objects in the display (so looks to the fixation cross or to empty parts of the display were not included in this calculation). An actual target advantage consists of a greater likelihood of looking at the target than at other objects in the display; possible values range from -1 to 1 , with 0 meaning that observers are equally likely to look at target as at nontargets (and a negative number meaning a target disadvantage). Fig. 1 shows the timeline by which a target advantage emerges for each condition. On the x -axis, the zero point is at the onset of the referring expression. In all but one case (where the expression consisted of a bare noun), referring expressions consisted of an adjective and a noun, with approximately 400 ms for the adjective and 400 for the noun, with no differences by condition (see Metzinger & Brennan, 2003, p. 208; Fig. 2). So the mean offset of the target expression is at ~ 800 ms in Fig. 1.

The finding is that in all three conditions that do not involve breaking conceptual pacts, the rises in proportions of looks representing a target advantage begin to emerge at around 700–800 ms; in the broken precedent condition, there is no rise until ~ 1200 ms. This is consistent with the finding Metzinger and Brennan (2003) originally reported of delayed first

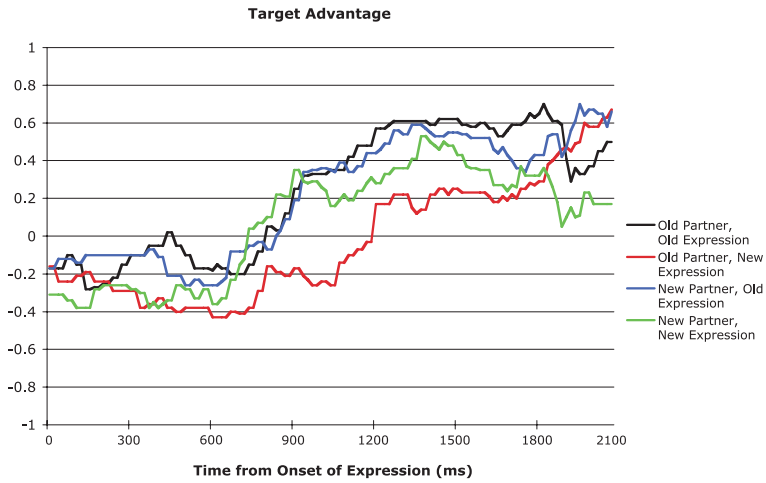


Fig. 1. Target advantage (looks to target minus looks to other objects in the display), beginning at the onset of the critical referring expression. The offset of referring expressions is at approximately 800 ms for all conditions. The red line represents the condition in which the speaker breaks a conceptual pact.

looks to the target when conceptual pacts were broken. Fig. 1 also shows that very briefly early on (at about 450 ms), looks to the target are equally likely as looks to other objects, but only for the condition in which re-referring is most natural (where a familiar speaker uses an entrained-upon expression).

Also following Kronmüller and Barr (2007), we tested for reliable effects in each 300 ms. interval. Table 1 displays the results of these tests for Partner (Familiar or New), Expression (Familiar or New), Partner × Expression, as well as a planned contrast of New Expressions spoken by either New or Familiar Partners (the latter representing a broken conceptual pact). The earliest reliable effect is one of precedent (300–600 ms), which then disappears in the next interval, to emerge again later (1200–1500 ms). There is a strong partner

Table 1
Summary of effects across intervals

Interval	Target Advantage?	Precedent	Partner	Precedent × Partner	Contrast of New Expression, New vs. Familiar Partner (Broken Pact)
0–300	No	n.s.	n.s.	n.s.	n.s.
300–600	No	$p = .015$	n.s.	n.s.	n.s.
600–900	Target advantage begins, except for broken pacts	n.s.	n.s.	n.s.	n.s. ($p = .136$)
900–1200	Yes, except for broken pacts	n.s. ($p = .11$)	$p = .005$	$p = .034$	$p = .001$
1200–1500	Yes, in all conditions	$p = .026$	n.s.	n.s.	n.s.
1500–1800	Yes	n.s.	n.s.	n.s.	n.s.
1800–2100	Yes	n.s.	n.s.	n.s.	n.s.

effect in the 900–1200 ms interval that is driven by a reliable partner-by-expression interaction. What do we make of this pattern?

Because the precedent effect in the 300–600 ms interval actually shows a nontarget advantage for both new and familiar expressions, it is not convincing to argue that new expressions are being preempted from being mapped to familiar objects, since not even familiar expressions are yet mapped to familiar objects. As a result, we are inclined to conclude that this apparently early effect of new versus familiar expressions is not (yet) due to binding a familiar expression to a familiar referent, but simply due to elevated looking around when a new expression is heard. Such looking around would have been rampant with the recent introduction of previously unmentioned objects, as in Kronmüller and Barr's (2007) procedure.

Moreover, in the next few 300 ms intervals, there appears to be no good evidence that precedent has its effect before partner. Early on, while the referring expression is still midway through articulation in the 300–600 ms interval after its onset, the slope of the New-Expression-New-Partner line begins to diverge from the broken pact condition (New-Expression-Old-Partner), as early as 450 ms (see Fig. 1). Although this divergence is not reliable until 900–1200 ms, it is apparently strong enough to swamp the evolving precedent effect until that reappears later (at 1200–1500 ms).

It is worth noting that in the data of Kronmüller and Barr (2007, Fig. 2, p. 443), the target advantage began substantially earlier than in Metzing & Brennan's data, at about 450 ms for familiar expressions and about 600 ms for new expressions. Their partner-specific effect (New-Expression-New-Partner vs. New-Expression-Familiar Partner) did not even begin to emerge until 1200 ms. This striking difference in time course between the two experimental situations, which we take up in the next section, suggests that they may have represented quite different language games for their participants.

4. Discussion

Let's return now, in light of the material above, to the question of modularity. Modularity is a very strong claim—if a process is encapsulated within a module, it is impervious to influence from outside or parallel processes. It *cannot* take account of external information until the module is exited. Even if the claim is mitigated for a cascading system, where results can be fed forward for further processing before computation within the first module is finalized, this still means that the information operating in the first stage of processing does so independently. What pattern of data would convincingly support a two-stage model of language processing, where the initial stage is automatic and egocentrically based? To support a modularity claim, there should be not only consistent evidence about the effect of this proposed initial stage but also reasonably consistent evidence of the early timing with which it occurs.

The timing of the use of partner-specific information ranges widely in the literature. The Metzing and Brennan (2003) data show a speaker identity effect by 900–1200 ms after the onset of the referring expression, while recent data from Brown-Schmidt (2008) replicating Metzing & Brennan with computer-based displays found this effect substantially earlier,

at 180–300 ms. These findings are in contrast to Kronmüller and Barr (2007), who did not find a partner-specific effect until 1200–1500 ms in Experiment 1, and either earlier or later than that depending on the presence of cognitive load in Experiment 2. Some visual worlds studies have found effects of common ground as early as 200–500 ms after the onset of the point of disambiguation (Experiment 1, Hanna et al., 2003) and others have found effects of a speaker's pragmatic perspective immediately (Hanna & Tanenhaus, 2004) and of a speaker's mismatching perspective slightly later, within about 700 ms of the onset of the point of disambiguation (Experiment 2, Hanna et al., 2003).

How can these timing differences be reconciled, and which model provides the most parsimonious explanation for them? Two-stage models do not provide a ready theoretical, architectural, or computational explanation for why a supposedly automatic egocentric stage would last 200 ms in one experiment and 1800 ms in another. On the other hand, constraint-based models of comprehension (Jurafsky, 1996; MacDonald, 1994; Tanenhaus & Trueswell, 1995) can account for such differences quite naturally. Constraint-based models propose that all relevant and salient sources of information are simultaneously integrated, as they become available, in order to provide probabilistic evidence for competing interpretations. These models predict that the influence of one source of information will be modulated by the presence and strength of other constraints, and this type of computational system naturally gives rise to the prediction that in some tasks, speaker-based information will show faster or stronger effects than in others. In particular, different sources of information can make use of different computational resources, and information that is primarily top down in nature (like speaker identity) might sometimes take longer to interact with information that is data driven from the bottom up (like lexical semantics), even though the architecture of the processing system is not imposing any staged delays.

Finally, positing a two-stage, modular architecture just to account for occasionally egocentric behavior becomes even less convincing in light of new evidence from experiments in social neuroscience. Recent electrophysiological work by van Berkum, van den Brink, Tesink, Kos, and Hagoort (2008) using scalp-recorded event-related potentials (ERPs) demonstrated that evidence about another (in this case, noninteracting) speaker's social perspective can be taken into account from the very earliest moments of processing. In this experiment, listeners heard statements that did or did not match stereotypical inferences about the speaker; for example, they heard statements that were odd for a young speaker, but not for an adult speaker, such as *every evening I drink some wine before I go to sleep*, statements that were odd for a female speaker, but not a male speaker, such as *just before the counter I dropped my aftershave on the floor*, and statements that were odd for a male but not a female, such as *I recently had a check-up at the gynecologist in the hospital*. The mismatches between the perspectives implicit in the speakers' voices and those implicit in the utterances were processed incidentally (listeners were not told to monitor for such mismatches) and evoked N400 waves (a standard measure of semantic anomaly). These effects of incongruity were reliable and immediate; although the magnitude of the N400 was smaller than for lexical semantic anomalies, it showed up in brain potentials just as early, beginning 300–500 ms after the onset of the inconsistent word (van Berkum et al., 2008). Remarkably, this immediate N400 effect of speakers' social perspective was cued entirely

by the prerecorded voices of the speakers presented in blocks (there was no interaction or visual co-presence needed to support this perspective-taking effect). That the effect of speaker's identity in the van Berkum et al. study was relatively small underscores the importance of ensuring that the hunt for partner-specific effects turn over every stone before concluding that they do not exist (see Kraljic & Brennan, 2005), as well as avoiding biases that could obscure such effects.

5. Conclusions

To the extent that language is *for* communication, language use in dialog contexts is a fundamental kind of joint action that may be used to coordinate many other joint activities such as collaborative motor actions, joint visual search, and recall or problem solving in groups; we expect that better understanding how interlocutors adapt their communicative behavior and processing to each other will contribute to a more nuanced understanding of how cognition works in collaborative contexts more generally. Pickering and Garrod (2004) appealed to *priming* as an explanation for interactive alignment. But we argue that priming (whether between or within the minds of interlocutors) is not an *explanation* for adaptive behavior, but is simply the currency with which memory-based processes are purchased more generally. Communicative processes are opportunistic and fallible rather than deterministic, and the reason that they succeed so frequently is that interlocutors are willing to distribute their effort jointly (Clark & Wilkes-Gibbs, 1986).

We conclude that the patterns of evidence, both old and new, seem to support an architecture for probabilistic, constraint-based processing rather than two-stage processing. Concerning entrainment in referential communication, there is no compelling evidence to support encapsulated processes that delay the use of partner-specific information while giving precedence to egocentric information (nor the other way around, to information in common ground); in other words, all information in memory, whether it concerns one's own or a partner's perspective (or anything else), can function as a constraint to probabilistically guide processing, as long as it is activated and available (Horton & Gerrig, 2002, 2005a; Jurafsky, 1996; MacDonald, 1994; Metzing & Brennan, 2003; Sedivy, 2005; Spivey-Knowlton, Trueswell, & Tanenhaus, 1993).

Although we have argued that partner-specific adaptation must be explainable by general principles of memory and cognition rather than by appealing to special modules, it is certainly possible that the social importance of certain cues could lead interlocutors to use them with more facility, attentiveness, or motivation than less relevant cues. And if the general activity of collaborating with a partner on joint tasks has a meaningful and common neural basis (e.g., leading to increased activation in the medial prefrontal cortex as suggested by Sebanz, Rebbechi, Knoblich, Prinz, & Frith, 2007), this too may hold implications for the status of socially important cues, both verbal and nonverbal. Such cues may be mediated by the same neural circuits that support other forms of joint action.

One of the arguments that true (partner-specific) audience design cannot happen early or automatically has been that such adjustments require complex inferences about the partner's

needs, knowledge, or perspective, and so maintaining and updating a model of the partner could simply be computationally too expensive. But in fact, taking the perspective of a partner into account early in processing does not have to be computationally costly because maintaining and updating an elaborate partner model need not be implicated (as pointed out by Galati & Brennan, 2006). It seems to be no coincidence that the studies most likely to show early partner-specific effects can often be summed up as simple either/or attributions about a partner's knowledge or needs, where the information about the partner can be available or attributed unambiguously in a timely fashion. That is, in many kinds of situations, speakers and addressees who hold somewhat different perspectives nevertheless quickly and accurately adapt to each other with "one-bit models" (Galati & Brennan, 2006) along the lines of *my partner can see what I'm doing*, or not (Brennan, 2005; Nadig & Sedivy, 2002); *my partner can reach the object she's talking about*, or not (Hanna & Tanenhaus, 2004); *my partner has a picture of what we're discussing*, or not (Lockridge & Brennan, 2002); *my partner and I have spoken about this before*, or not (Galati & Brennan, 2006; Metzing & Brennan, 2003); *my partner is currently gazing at this object*, or not (Hanna & Brennan, 2007); *my partner is a child*, as opposed to an adult; or *my partner is a native speaker of English*, as opposed to a nonnative speaker (Bortfeld & Brennan, 1997). The simple constraints inherent in such one-bit partner models can explain why some kinds of "smart" adjustments seem to be apparently immediate and relatively effortless (Galati & Brennan, 2006), even when a partner's knowledge, needs, and perspective differ from one's own.

Note that we do not view the episodic cues that underlie lexical entrainment to be inflexible or even necessarily stable, as interlocutors continually adjust to contextual changes. Interlocutors do not expect one another to rigidly adhere to conceptual pacts or "established expression-referent mappings," but, we argue, consider such episodic mappings to be temporary and for current purposes. Interlocutors remain flexible when a partner uses a new or revised expression, making appropriate attributions (e.g., that a new object is referred to, that the figure-ground relationship of an old object has changed, that the speaker wishes to add information or focus attention on a relevant feature, or that the object needs to be distinguished from something else). Of course, it may take time to initially make such an attribution, but once it is available, there is no architectural reason why it cannot be used rapidly. The attributions that speakers or addressees make about each other are essentially self-generated cues that in some cases originate with a simple inference about a partner's knowledge or needs. Once the inference or attribution has been made, a relevant cue may be available for audience design even if conflicts with an individual's own perspective. In this way, audience design can be both automatic and smart.

We have argued that in studies of dialog as coordinated cognition, it can be difficult to achieve sufficient control without obliterating the very processes that make dialog what it is (especially when different researchers have different assumptions about what those processes are; for discussion, see Kuhlen & Brennan, 2008). Several basic questions remain concerning the cognitive and neural underpinnings of dialog as joint action. One question is whether *representational parity* (in Pickering & Garrod, 2004) or *common coding* (in Sebanz & Knoblich, 2009) should after all be labeled as egocentric, especially when the self's frame of reference can be shaped by observing the actions of others (see, e.g.,

Knoblich & Sebanz, 2008; Sebanz & Knoblich, in press). At the same time, a second question is whether this association between the perspectives of self and other are necessarily automatically and tightly coupled (to what extent, and under what conditions, can people keep their own and others' perspectives apart, in order to represent a triadic relationship between self, other, and object? See Sebanz & Knoblich [2009]). Addressing such questions is an interdisciplinary adventure that requires synthesizing methods and perspectives from cognitive science, social psychology, and neuroscience.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants # ISI-0527585, ITR-0325188, ITR-0082602, and IIS-0713287. We thank Darron Vanaria for his assistance. This project was inspired by our late colleague and collaborator, Charles Metzger (http://www.psychology.sunysb.edu/sbrennan-/CM_tribute.html), without whom it would not have been possible; this paper is dedicated to him.

References

- Bard, E. G., Anderson, A. H., Sotillo, C. F., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*, 1–22.
- Bard, E. G., & Aylett, M. P. (2005). Referential form, word duration, and modeling the listener in spoken dialogue. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 173–191). Cambridge, MA: MIT Press.
- Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. *Journal of Memory and Language*, *46*, 391–418.
- Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers. *Discourse Processes*, *23*, 119–147.
- Brennan, S. E. (1990). *Speaking and providing evidence for mutual understanding*. Unpublished doctoral dissertation. Stanford, CA: Stanford University.
- Brennan, S. E. (1991). Conversation with and through computers. *User Modeling and User-Adapted Interaction*, *1*, 67–86.
- Brennan, S. E. (2005). How conversation is shaped by visual and spoken evidence. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 95–129). Cambridge, MA: MIT Press.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482–1493.
- Brown, P. M., & Dell, G. S. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, *19*, 441–472.
- Brown-Schmidt, S. (2008). *Time-course of processing conceptual pacts in conversation reveals early partner-specific effects*. Poster presented at the Twenty-First Annual CUNY Conference on Human Sentence Processing. Chapel Hill, NC.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1980). Rules and representations. *The Behavioral and Brain Sciences*, *3*(1), 1–62.
- Clark, H. (1992). *Arenas of language use*. Chicago: University of Chicago Press.
- Clark, H. H. (1997). Dogmas of Understanding. *Discourse Processes*, *23*, 567–598.

- Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge, England: Cambridge University Press.
- Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.
- Dell, G. S., & Brown, P. M. (1991). Mechanisms for listener-adaptations in language production: Limiting the role of the ‘Model of the Listener’. In D. Napoli & J. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore Festschrift for Lila Gleitman* (pp. 105–129). Hillsdale, NJ: Erlbaum.
- Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40, 296–340.
- Fowler, C., & Housum, J. (1987). Talkers’ signaling of “new” and “old” words in speech and listeners’ perception and use of the distinction. *Journal of Memory and Language*, 37, 24–40.
- Fussell, S. E., & Krauss, R. M. (1989). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology*, 19, 509.
- Fussell, S. E., & Krauss, R. M. (1991). Accuracy and bias in estimates of others’ knowledge. *European Journal of Social Psychology*, 21, 445–454.
- Fussell, S. E., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers’ assumptions about what others know. *Journal of Personality and Social Psychology*, 62, 378–391.
- Galati, A., & Brennan, S. E. (2006). Given-new attenuation effects in spoken discourse: For the speaker, or for the addressee? *Abstracts of the Psychonomic Society, 47th Annual Meeting* (p. 15), Dallas, TX.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, 181–218.
- Gerrig, R. H., Brennan, S. E., & Ohaeri, J. O. (2000). What can we conclude from speakers behaving badly? *Discourse Processes*, 29, 173–178.
- Glucksberg, S., Krauss, R., & Weisberg, R. (1966). Referential communication in nursery school children: Method and some preliminary findings. *Journal of Experimental Child Psychology*, 3, 333–342.
- Gregory, M. L., Healy, A. F., & Jurafsky, D. (2001). *The role of the hearer in durational shortening*. 14th Annual CUNY Conference on Human Sentence Processing, University of Pennsylvania, Philadelphia, PA.
- Gregory, M. L., Healy, A. F., & Jurafsky, D. (2002). Common ground in production: Effects of mutual knowledge on word duration. Unpublished manuscript.
- Hanna, J. E., & Brennan, S. E. (2007). Speakers’ eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language*, 57, 596–615.
- Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. *Cognitive Science*, 28, 105–115.
- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49, 43–61.
- Horton, W. S., & Gerrig, R. J. (2002). Speakers’ experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, 47, 589–606.
- Horton, W. S., & Gerrig, R. J. (2005a). Conversational common ground and memory processes in language production. *Discourse Processes*, 40, 1–35.
- Horton, W. S., & Gerrig, R. J. (2005b). The impact of memory demands on audience design during language production. *Cognition*, 96, 127–142.
- Horton, W., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59, 91–117.
- Johnston, W. A., Hawley, K. J., Plewe, S. H., Elliott, J. M. G., & DeWitt, M. J. (1990). Attention capture by novel stimuli. *Journal of Experimental Psychology: General*, 119, 397–411.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic disambiguation. *Cognitive Science*, 20, 137–194.
- Keysar, B. (1997). Unconfounding common ground. *Discourse Processes*, 24, 253–270.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11, 32–38.

- Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998a). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language*, *39*, 1–20.
- Keysar, B., Barr, D. J., & Horton, W. S. (1998b). The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Science*, *7*, 46–50.
- Knoblich, G., & Sebanz, N. (2008). Evolving intentions for social interaction: From entrainment to joint action. *Philosophical Transactions of the Royal Society, B*, *363*, 2021–2031.
- Kraljic, T., & Brennan, S. E. (2005). Using prosody and optional words to disambiguate utterances: For the speaker or for the addressee? *Cognitive Psychology*, *50*, 194–231.
- Krauss, R. M. (1987). The role of the listener: Addressee influences on message formulation. *Journal of Language and Social Psychology*, *6*, 81–98.
- Kronmüller, E., & Barr, D. J. (2007). Perspective-free pragmatics: Broken precedents and the recovery-from-preemption hypothesis. *Journal of Memory and Language*, *56*, 436–455.
- Kuhlen, A., & Brennan, S. E. (2008). Addressees shape speaking: When confederates may be hazardous to your data. *Abstracts of the Psychonomic Society, 49th Annual Meeting* (p. 6), Chicago, IL.
- Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*, 172–187.
- Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin and Review*, *9*, 550–557.
- MacDonald, M. C. (1994). Probabilistic constraints and syntactic ambiguity resolution. *Language and Cognitive Processes*, *9*, 157–201.
- McAllister, J., Potts, A., Mason, K., & Marchant, G. (1994). Word duration in monologue and dialogue speech. *Language and Speech*, *37*, 393–405.
- Metzing, C., & Brennan, S. E. (2003). When conceptual pacts are broken: Partner-specific effects in the comprehension of referring expressions. *Journal of Memory and Language*, *49*, 201–213.
- Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective taking constraints in children's on-line reference resolution. *Psychological Science*, *13*, 329–336.
- Nooteboom, S. G. (1991). Perceptual goals of speech production. *Proceedings of the XII International Congress of Phonetic Sciences, Vol. 1*, (pp. 107–110). Aix-en-Provence, France, August 19–24.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–226.
- Polichak, J. W., & Gerrig, R. J. (1998). Common ground and everyday language use: Comments on Horton and Keysar (1996). *Cognition*, *66*, 183–189.
- Samuel, S. G., & Troicki, M. (1998). Articulation quality is inversely related to redundancy when children or adults have verbal control. *Journal of Memory and Language*, *39*, 175–194.
- Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 123–164). Hillsdale, NJ: Lawrence Erlbaum.
- Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*, 211–232.
- Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, *1*(2), 353–367.
- Sebanz, N., & Knoblich, G. (in press). From mirroring to joint action. In I. Wachsmuth, M. Lenzen, & G. Knoblich (Eds.), *Embodied communication*. (pp. 129–150). Oxford, England: Oxford University Press.
- Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., & Frith, C. (2007). Is it really my turn? An event-related fMRI study of task sharing. *Social Neuroscience*, *2*, 81–95.
- Sedivy, J. C. (2005). Evaluating explanations for referential context effects: Evidence for Gricean mechanisms in online language interpretation. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 345–364). Cambridge, MA: MIT Press.

- Spivey-Knowlton, M. J., Trueswell, J. C., & Tanenhaus, M. K. (1993). Context effects in syntactic ambiguity resolution: Discourse and semantic influences in parsing reduced relative clauses. *Canadian Journal of Experimental Psychology*, 37, 276–309.
- Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634.
- Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Handbook of perception and cognition, Vol. 11: Speech language and communication* (pp. 217–262). San Diego, CA: Academic Press.
- Tanenhaus, M., & Trueswell, J. (2005). Eye movements as a tool for bridging the language-as-product and language-as-action traditions. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-action traditions* (pp. 3–37). Cambridge, MA: MIT Press.
- Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, 20, 580–591.
- Yang, H., & Zelinsky, G. (2006). Evidence for guidance in categorical visual search. *Journal of Vision*, 6(6), 449a.