tors rarely if ever have *identical* models. Instead, they have partially aligned models that may differ in many – sometimes important – respects. And because evidence for situation models is mediated through language, it seems highly unlikely that they can act directly upon one another (contra P&G's Fig. 2).

One interesting result of the distinction between alignment of linguistic and situation models is that alignment of linguistic representations may sometimes lead to misaligned situation models. Garrod and Clark (1993) found that young children had a tendency to use the same words to describe a maze – that is, showed lexical alignment – even when their situation models were quite different. Similarly, in the case of the doctor-and-patient scenario, one speaker's use of the term *chronic* may well reinforce the other's use of the same term, leading to more misunderstanding than if a different term were used. In both examples, full alignment at the linguistic level misleads interlocutors into believing that they also have alignment at the level of situation models.

ACKNOWLEDGMENTS

This research was supported by a British Academy Postdoctoral Fellowship and by ESRC grant R000239363.

Two steps forward, one step back: Partner-specific effects in a psychology of dialogue

Susan E. Brennan and Charles A. Metzing

Department of Psychology, State University of New York, Stony Brook, NY 11794-2500. **susan.brennan@sunysb.edu cmetzing@ic.sunysb.edu**
http://www.psychology.sunysb.edu/psychology/personnel/Susan.htm

Abstract: Pickering & Garrod's (P&G's) call to study language processing in dialogue context is an appealing one. Their interactive alignment model is ambitious, aiming to explain the converging behavior of dialogue partners via both intra- and interpersonal priming. However, they ignore the flexible, partner-specific processing demonstrated by some recent dialogue studies. We discuss implications of these data.

In human language processing, the whole is greater than the sum of the parts; therefore, those who study the language processing system in dialogue contexts are poised to make different sorts of discoveries than those who study the parts working alone. Pickering & Garrod (P&G) present a convincing argument that psycholinguists should pay attention to dialogue. In fields such as artificial intelligence and human-computer interaction, where the goal is often to build a fully working dialogue partner, many will find this a worthy enterprise as well. After presenting evidence for phonological, lexical, and syntactic convergence between dialogue partners and for representations shared between comprehension and production, P&G make a strong claim that is far *less* convincing: "normal conversation does not routinely require modeling the interlocutor's mind" (sect. 4.4, para. 4). They support this position with evidence from studies that fail to meet the very standards they seek to advance, while ignoring evidence that complicates matters for their interactive alignment model. Thus, their position on the importance of studying language in dialogue does not go far enough.

This position assumes that interlocutors achieve aligned mental representations without having to track anything specific about each other's knowledge because both have evolved with the same cognitive architecture; what is easiest for speakers is easiest for addressees (Brown & Dell 1987). It further assumes that there is no need to track common ground, as interlocutors each use their own memory of the conversation as a proxy. By this argument, what appears to be partner-specific or "audience design" is actually inflexible and unavoidable, at least in the earliest moments of processing. P&G propose a two-stage model (similar to that of Horton & Keysar 1996), arguing that interlocutors "do not routinely take common ground into account during initial processing . . . full common ground is only used when simpler mechanisms are ineffective" (sect. 4.1, last para.). This (circular) view relegates any aspect of production or interpretation that displays flexibility or sensitivity to an interlocutor's needs (as distinct from one's own) to the status of a relatively late adjustment, managed as a kind of repair or pragmatic garden path.

Granted, it is difficult to design a good experiment on audience design. A good experiment must distinguish one interlocutor's perspective from another's, avoid confounding individual perspectives with common ground (Keysar 1997), and allow interlocutors to interact naturally or contingently (Schober & Brennan 2003). But we are surprised that studies succeeding in all this (and finding partner-specific effects early in processing, e.g., Hanna et al. 2003; Nadig & Sedivy 2002) are dismissed by P&G: "their task was repetitive and involved a small number of items, and listeners were given explicit information about the discrepancies in knowledge" (target article, sect. 4.2, para. 3). Then follows a very broad claim: "Under such circumstances, it is not surprising that listeners develop strategies that may invoke full common ground. During natural dialogue, we predict that such strategies will not normally be used."

Paradoxically, evidence to support this position comes mainly from studies that did not allow any potential for interaction. These include Brown and Dell (1987), Ferreira and Dell (2000), Horton and Keysar (1996), and others in which partners did not interact naturally or provide contingent feedback. Sometimes this matters; for example, Brown and Dell (1987) concluded that speakers did not take addressees' specific needs into account when retelling stories; but their addressees had no needs (they were confederates who knew the stories better than the speakers did). When we ran a similar study using spontaneously interacting speakers and addressees (Lockridge & Brennan 2002), speakers' early syntactic choices indeed showed sensitivity to addressees' needs.

There is additional good evidence of rapid, partner-specific effects from the comprehension side. Hanna and Tanenhaus (2004) asked addressees to follow a (confederate) speaker's directions in a cooking task (e.g., *Hand me the cake mix*); the addressees' eye fixations showed that they restricted candidate referents for ambiguous expressions (e.g., when two cake mixes were present) depending on what the speaker was holding and what she could not reach; they did this from the earliest moments of processing.

And we have demonstrated that addressees interpret the same utterance differently when it is spoken by different speakers with whom the addressees have different dialogue histories (Metzing & Brennan 2001; 2003). In our experiment, addressees were instructed by (confederate) speakers to reposition objects among a relatively large set; they did this several times, evolving shared perspectives and terms for critical objects (e.g., *the shiny cylinder*). Then the speaker left the room and either returned or else a new confederate speaker entered. In the final trial, the new or old speaker used either the familiar term or a new, equally good term (e.g., *the silver pipe*) for the same critical object (amid many other references that did not use different terms). Addressees gazed immediately at the object when either speaker used the old term. However, when the old speaker used a new term (inexplicably breaking a conceptual pact), addressees experienced interference, delaying gazing at the target object. There was no such delay when the new speaker used the new term (in fact, resolving this was just as fast as the old term spoken by the new speaker). This partner-specific interference suggests that the pragmatic force of breaking a conceptual pact has impact immediately, rather than just as a late adjustment or repair.

Such immediate effects provide evidence of impressive agility and potential for partner-specific processing in the language processing system, which the interactive alignment proposal fails to address. Pragmatic and partner-specific knowledge is implemented by basic mechanisms of memory and does not rely on special processes or exhaustive partner models. Audience design – truly partner-specific processing – can occur immediately and effortlessly as well as more slowly and deliberately, depending on

https://doi.org/10.1017/S0140525X04240055 Published online by Cambridge University Press

how activated relevant information is. P&G's strong separation between "implicit common ground" (automatic but excluding any partner specificity) and "full common ground" (requiring reprocessing) is unconvincing.

Yes, the potential for interaction matters! But the interactive alignment model has farther to go before it can accommodate the flexible and adaptive processing that these data support.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under grants No. 0082602 and No. 9980013. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Priming and alignment: Mechanism or consequence?

Sarah Brown-Schmidt and Michael K. Tanenhaus

Department of Brain and Cognitive Sciences, University of Rochester, River Campus, Rochester, NY 14627. **sschmidt@bcs.rochester.edu**
mtan@bcs.Rochester.edu
http://www.bcs.rochester.edu/people/sschmidt/
http://www.bcs.rochester.edu/people/mtan/mtan.html

Abstract: We agree with Pickering & Garrod's (P&G's) proposal that dialogue is an important empirical and theoretical test bed for models of language processing. However, we offer two cautionary notes. First, the enterprise will require explicit computational models. Second, such models will need to incorporate both joint and separate speaker and hearer commitments in ways that go beyond priming and alignment.

We applaud and second Pickering & Garrod's (P&G's) call to psycholinguists to include dialogue as an empirical and theoretical test bed for models of language processing. There is much to be gained by combining the tools for studying real-time processing developed within the language-as-product tradition with the more natural interactive situations typically used within the action tradition (Tanenhaus et al. 2004). And we believe that many of the basic processes in language comprehension, including spoken-word recognition, syntactic processing, and reference resolution, can be studied with the same precision in dialogue settings as in more traditional controlled experiments (e.g., Brown-Schmidt et al. 2002; in press). However, we would place a somewhat different emphasis on why it is important to study dialogue.

First, unlike P&G, who suggest that theories of language processing within the product tradition are admirably well-specified, we think that there has been a dearth of explicit mechanistic models of language processing. This is especially true within language comprehension, where, with the exception of some limited models developed within the neural network tradition, most computational models make only tenuous contact with behavioral data, and vice versa (Christiansen & Chater 2001). However, we believe that the emergence of interactive dialogue systems within computational linguistics offers an opportunity to develop explicit computational models in domains that can also be used to study human language processing, thus creating a synergistic feedback loop between modeling and experimentation. Although the field has not yet reached this stage, the opportunity is on the horizon, as computational linguists strive to implement systems that can engage in continuous generation and understanding.

The motivation for continuous understanding and generation in dialogue systems is instructive (Allen et al. 2001). During utterance generation, a speaker needs to have the capacity to monitor feedback from an addressee, both verbal and nonverbal, and plan or adjust the continuation of her utterance accordingly. Consider, for example, an utterance which begins, "Now, take thee,[1] uh, Phillips head screwdriver. . . ." If the addressee nods or says "uh huh," the speaker can continue with, "and tighten the bolt." However, if the addressee says "huh," looks perplexed, or begins to reach for the wrong tool, the speaker is likely to continue with, "the one with the blue handle, the one closest to your wrench" or say "No, *that* one" (while pointing). This example illustrates one of the potential benefits of studying interactive dialogue: It can shed different light on some basic assumptions. For example, most psycholinguists use the fact that comprehension occurs more or less continuously as a reason why they focus on response measures that are closely time-locked to the input. Yet, if pressed to answer the question of why comprehension is so incremental, they would likely appeal to the fact that memory capacity is limited. But, working memory constraints do not explain why processing is as relentlessly continuous as it is. In fact, the early theories of language comprehension, which were largely driven by working-memory assumptions, assumed that comprehension was a catch-up game, with many delays so that listeners could avoid making premature commitments. However, incremental processing is necessary for efficient use of feedback from an interlocutor.

In confronting the challenges involved in integrating comprehension and production, as well as speech recognition, intention recognition, and utterance planning, psycholinguists will be required to go beyond the boxes and arrows and consider more detailed models of what needs to be computed. Here, the study of dialogue provides a useful paradigm by allowing psycholinguists and computational linguists to study and model these processes in relatively constrained domains where it is possible to be explicit about the relevant components of the system and how they interact.

We believe that when P&G consider the problem of dialogue from this perspective, they will need to rethink two assumptions that guide their approach. The first is that interlocutors do not need to take into account differences between speaker and hearer knowledge and perspectives, except for repair strategies. We agree that many aspects of language use can be egocentric, though we think that P&G overstate the case, mistaking evidence that common ground does not fully constrain referential domains, with evidence for the stronger claim that it is ignored in initial processing (cf. Hanna et al. 2003). Moreover, natural language is rife with constructions that depend crucially on differences between assumed speaker and hearer knowledge. To take a simple example, the declarative question "You used the Phillips tool?" can be uttered only when the speaker believes the addressee is committed to the presupposed assertion that is being questioned (Gunlogson 2003). If the addressee had in fact used the tool, she might respond with an explanation; if she had not, she would need to contradict the presupposed commitment by saying something like "No, of course not, I. . . ."

These kinds of phenomena will emerge as psycholinguists consider richer dialogue situations in which the participants have reasons to make choices between alternative forms of utterances, and more complex interactions where it is important for interlocutors to track each other's attention and intentions. Finally, note that both intention and attention, which are crucial components of monitoring an interlocutor, do not necessarily appeal to high-level, resource-demanding processes. Eye gaze, for example, is a powerful source of information about attention. Moreover, the aspect of an interlocutor's knowledge that has to be monitored can often be circumscribed by goal structures. Cast in this light, perspective monitoring can be seen as one of the basic components of communication, rather than a special purpose, resource-intensive, revision process. In many of the examples that P&G focus on, the goal structures are completely defined by the task and the choice of utterances is limited. Under these conditions, it is easy to view alignment and priming as the primary mechanisms of dialogue, rather than as interesting phenomena that can be used to provide insight into the representations and processes underlying interactive conversation.

ACKNOWLEDGMENT

This work was supported by a National Institutes of Health grant, No. HD-27206, to Michael K. Tanenhaus.

https://doi.org/10.1017/S0140525X04240055 Published online by Cambridge University Press