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/S0140525X04250051 Published online by Cambridge University Press

NOTE

1. We use the word "thee" to indicate a disfluent pronunciation of the word "the" (and not in the old English usage as a pronoun).

A call for more dialogue and more details

J. Cooper Cutting

Department of Psychology, Illinois State University, Normal, IL 61790.
jccutti@ilstu.edu
http://main.psy.ilstu.edu/faculty/cutting/index.html

Abstract: Pickering & Garrod (P&G) argue that contemporary models of language use are inadequate. This has resulted largely because of an experimental focus on monologue rather than dialogue. I agree with the need for increased experimentation that focuses on the interplay between production and comprehension. However, I have some concerns about the Interactive Alignment model that the authors propose.

Pickering & Garrod (P&G) make an excellent argument for the need to increase the study of the interplay between production and comprehension in mainstream psycholinguistic research. Historically, most research on language has treated the processes of comprehension and production as largely independent systems. This separation may reflect the different starting points of the two processes. Input for comprehension is easy to manipulate and control experimentally. Input for production is thought, something that is much harder to control experimentally (Bock 1996). As a result, language comprehension research has a long history relative to that of language production research. Furthermore, theories of comprehension are rarely informed by research on production, and vice versa; however, the two processes are clearly intertwined. A complete model of language use needs to consider how production and comprehension processes work in conjunction. The Interactive Alignment model is explicitly designed to do just that. Rather than focus on how we process language as a monologue, as most mainstream psycholinguistic research does, the model focuses on the processes involved in dialogue between two (or more) language users.

Although I support the idea that much more research focusing on the interplay of production and comprehension is needed, I believe that the Interactive Alignment model that the authors propose is underspecified in several important respects.

1. What is alignment? The central feature of the model is the *alignment* of multiple levels of linguistic representations between interlocutors. This is achieved via "channels of alignment" through which "the activation of a representation in one interlocutor leads to the activation of the matching representation in the other interlocutor *directly*" (sect. 3.2, para. 2; my emphasis). In P&G's Figure 2, these channels are depicted as bidirectional links between representations within both dialogue participants. Although it is easy to imagine direct links between representations within a person (see the parity assumption below), it is much more difficult to determine what these "direct" links between different individuals correspond to. The authors propose that the mechanism of alignment is a "primitive and resource-free priming mechanism" both within and between levels of representation. Although this may be different from strict serial models of language use, it seems to be a feature of existing interactive models of language use (for recent reviews, see Dell et al. 1999; Pickering et al. 2000). Within these models, the impact of recent use (of representations) in dialogue could be modeled with residual activation of representations from earlier parts of the conversation. However, this priming mechanism does not seem to be very "direct" but is instead indirect.

2. Is all priming alike? The authors suggest that alignment of multiple levels of representation (including phonological, syntactic, lexical, semantic, and situational) is achieved by priming. However, "priming" phenomena in language may not all share the same underlying mechanisms. For example, Bock and Griffin (2000) have suggested that syntactic persistence is a reflection of learning rather than spreading of activation. It is unclear how the interactive alignment model distinguishes between activation-based and learning-based priming mechanisms.

3. Parity of representations. The authors propose that there is parity between the representation used by production and that used by comprehension. In other words, although production and comprehension processes may be different, they share the same set of linguistic representations. However, the issue of shared versus distinct representations needs to be considered for each level of representation (e.g., Balota 1990; Caramazza 1991; Levelt 1989). For example, Cutting (1998) used a word-picture priming technique in which two words were presented in the prime trial. The participant produced one word while the other was ignored. Cutting found that both produced and ignored semantically related prime words interfered with picture naming. However, only produced phonologically related primes influenced picture naming; ignored phonologically related primes had no effect. Cutting interpreted this pattern of results as support for a model in which semantic representations are shared by production and comprehension whereas phonological representations are separate. The authors refer to Hommel et al.'s (2001) Theory of Event Coding (TEC) as a model for language representations. Interestingly, Hommel et al.'s proposal also states that TEC is most appropriate for abstract distal coding. In other words, because phonological representations are the linguistic representations "nearest" to the sensory code and muscular innervation patterns, it should not be surprising that they are different for production and comprehension (cf. Martin et al. 1999; Zwitserlood 1994).

4. Routinized language. The authors propose that interlocutors use *routines* which are developed "on the fly" during dialogue (similar to idioms like "kick the bucket"). These routines result from the processes of alignment in dialogue. As speakers and listeners use particular lexical, semantic, and syntactic representations, some of these representations become bound into routines. The use of these routines may allow a speaker to "bypass" some of the early stages of processing assumed by traditional models of language production (e.g., Bock & Levelt 1994; Levelt 1989; Levelt et al. 1999). However, the process of routinization in the interactive alignment account is underdeveloped. It is unclear how these representations are bound (concurrent activation, learned connections, etc.), how they are represented (in either the short-term or the long-term; Cutting & Bock 1997), or how they actually "bypass" stages of language production (i.e., do routines bypass stages or just grease the wheels a bit?). Without further specification of these issues it is difficult to evaluate the model's predictions.

In conclusion, the interactive alignment model as currently proposed is underspecified with respect to how processes of alignment, priming, and routines work. As currently stated, it is difficult to determine specific testable predictions that would distinguish this model from currently available models of production and comprehension. However, the point of taking a dialogue perspective to these models (and the creation of new ones) is an excellent one. Although there is no shortage of monologue use of language (television, radio, print, lecture, etc.), clearly dialogue should be a central aspect of research. It is time to tear down some of the historically designed barriers between models of language comprehension and production and examine what impact a dialogue perspective has on these models.

https://doi.org/10.1017/S0140525X04250051 Published online by Cambridge University Press