semanticists precisely on developing theories of information states and their dynamics in dialogue – see, for example, work within the EU TRINDI project (Consortium 2000) and the annual SEMDIAL series of conferences on the semantics and pragmatics of dialogue. Links to all previous conferences, starting in 1997, in Munich, are available from: http://cswww.essex.ac.uk/semdial/

2. The # symbol is used here, as standard in linguistics, to mark an infelicitous utterance.

Dialogue: Can two be cheaper than one?

Sam Glucksberg

Psychology Department, Princeton University, Princeton, NJ 08544.
samg@princeton.edu **http://www.princeton.edu**

Abstract: Pickering & Garrod (P&G) argue that language processing in dialogue is in principle easier than in monologue. Although dialogue situations may provide more opportunities for facilitative priming, those priming mechanisms are also available in monologue situations. In both cases, the interactive alignment model calls strict modular accounts of language processing into serious question.

Pickering & Garrod's (P&G's) argument for the primacy of dialogue is a refreshing alternative to the isolated sentence as the unit of analysis in psycholinguistics. Others, such as Krauss, Clark, and their students and colleagues, have made similar arguments (cf. Clark 1996; Fussell & Krauss 1992), but none have made the additional argument that much of dialogue production and comprehension is automatic, requiring minimal cognitive or linguistic resources. The details of the priming processes have yet to be worked out in detail, but P&G provide enough empirical evidence to warrant optimism about an eventual fleshing out of their interactive alignment proposal. Beyond such details, their account has several intriguing implications, not the least of which is the claim that language production and comprehension are easier in interactive dialogue than in monologue. This flies in the face of the consensus that dialogue requires complex inferences about the mental state of one's interlocutor that are not required when one is, say, talking to oneself, or processing utterances under monologue conditions.

Is language processing in dialogues inevitably easier (more fluent, requires less processing, etc.) than in monologues? I do not think this question can be answered unequivocally. I can think of contexts in which dialogue language production would be quite difficult and contexts in which monologue production would be quite easy. An example of the former might be trying to hold a conversation with a sullen teenaged son or daughter on the merits of atonal music. Examples of the latter might include the think-aloud paradigm used in research on reasoning and problem-solving, or young children talking to themselves or to their dolls when no one else is around. So, as usual, the answer is: It depends.

A more useful question addresses the ways in which dialogue might be more likely than monologue to facilitate language processing. Such ways might in principle be available only in dialogue. Alternatively, it might just be a matter of likelihood, rather than a principled difference between monologue and dialogue. Consider the stages of speech production that could be facilitated in monologues and in dialogues. Following Levelt (1989), production could be facilitated at any or all of the following stages: the message (what to talk about); grammatical encoding (syntactic representation and lexical access); phonological encoding; and articulation. One of the more obvious ways to facilitate speech production is to provide something to talk about. In dialogue, the situation (including one's interlocutor) usually provides this, but of course there are (sometimes painful) situations where one does not know what to talk about next. In monologue, including the occasional classroom lecture (or a *BBS* commentary, for that matter), one can be at a loss as to what to say or write, but this is not a principled characteristic of monologues, just an unfortunate occasional one.

Grammatical and phonological encoding, on the other hand, may well be advantaged in dialogue via the mechanisms of priming and the opportunities for imitation. P&G make a convincing argument that such priming not only occurs but may well be automatic. However, such priming could in principle operate in monologue as well, although the opportunities are undoubtedly less likely. In monologue, as in dialogue, opportunities to repeat specific syntactic forms should facilitate successive productions. Similarly, opportunities to repeat specific word senses should facilitate lexical access and sense selection. What's missing from monologue, of course, are the contributions of an interlocutor, which can provide opportunities for imitation, as well as for syntactic, lexical, and phonological priming (but see Ferreira & Griffin 2003 for evidence of potential sources of priming in monologue).

Finally, P&G note that contextual constraints might operate more efficiently in dialogue than in monologue, but, again, this is probably not a principled difference. They write "With respect to lexical ambiguity, we predict that context will have a very strong role, so that effects of meaning frequency can be overridden" (sect. 5.4, para. 3). To my knowledge, this prediction has not yet been tested under dialogue conditions. However, it has been confirmed under monologue conditions, most recently by Sereno et al (2003), who found very early context effects on word recognition that overrode frequency effects.

To the extent that automatic priming can facilitate language comprehension and production, dialogue may well afford more opportunities for such priming than monologue. Nevertheless, the same mechanisms should be available for both. P&G's contribution is thus more profound than the observation that language processing in dialogue might be easier than in monologue. Instead, the contribution is in the insight that there may be mechanisms of language production and comprehension that can circumvent the need for costly cognitive processing in interactive conversation. People may not need to use a theory of mind or make inferences about common ground and mutual knowledge. They may, instead, rely on automatic priming both within and between interlocutors to produce alignment of representations, which, in turn, makes for successful and relatively effortless speech processing. The theoretical implication for theories of language processing is clear: A strictly modular, bottom-up language processing model simply will not do the job.

ACKNOWLEDGMENT

I thank Boaz Keysar for helpful comments on an earlier draft of this commentary.

Resonance within and between linguistic beings

Stephen D. Goldinger[a] and Tamiko Azuma[b]

[a]*Department of Psychology, Arizona State University, Tempe, AZ 85287-1104;* [b]*Department of Speech and Hearing Science, Arizona State University, Tempe, AZ 85287-0102.* **goldinger@asu.edu** **azuma@asu.edu**

Abstract: Pickering & Garrod (P&G) deserve appreciation for their cogent argument that dialogue merits greater scientific consideration. Current models make little contact with behaviors of dialogue, motivating the interactive alignment theory. However, the theory is not truly "mechanistic." A full account requires both representations and *processes* bringing those representations into harmony. We suggest that Grossberg's (1980) adaptive resonance theory may naturally conform to the principles of dialogue.

https://doi.org/10.1017/S0140525X04320054 Published online by Cambridge University Press

In their target article, Pickering & Garrod (P&G) present a clear and compelling case: Despite its being the predominate form of linguistic behavior, dialogue has been unduly ignored. Moreover, once dialogue is studied, it displays properties of entrainment at all levels (semantic, syntactic, phonetic, and others), which few psycholinguistic theories can address. To fill this void, the authors propose the *interactive alignment* model, wherein interlocutors share information at all linguistic levels, quickly arriving at states near equilibrium, allowing people to communicate with tremendous ease and efficiency. A by-product of alignment is that people tend to converge in dialogue, using similar expressions, words, and phrase structures. Beyond such linguistic constructs, interlocutors also tend to converge in their manner of speaking and even their physical postures (Capella 1981; Giles et al. 1991; Newtson 1994; Shockley et al. 2003).

The interactive alignment model is advanced to explain these coupling dynamics, using standard priming mechanisms. This approach is representation-driven, wherein one person's utterances prime ideas, words, and syntactic structures in his or her interlocutor. Mutual priming persists as a loosely coupled feedback loop (Van Gelder & Port 1995), so the same forms tend to occur repeatedly in conversation. At a general level, we believe this idea must be correct. However, the proffered model is underspecified, stated without explicit processes that act on the activated representations. Thus, the model nicely characterizes the problem space for a psychology of dialogue, but leaves many open questions. For example, how does the priming of extant representations predict new entities, such as routines, new jargon, or momentary changes in articulatory habit? How does speech perception lead to cascaded priming across linguistic levels? In short, what are the *processes* that explain the emergence of new psychological structures when people engage in dialogue?

In both language perception and production, models based on interactive activation have enjoyed great success (e.g., Dell 1986; Gaskell & Marslen-Wilson 1997; Levelt et al. 1999; McClelland & Elman 1986; McClelland et al. 1989). However, as P&G observe, these are all monologue models. Moreover, all generally assume fixed networks of nodes and connections, with weighting schemes that develop over extensive periods of training. Given their architecture, it is difficult to imagine how such models could communicate at multiple "levels," as suggested in Figure 2 of the target article. Of greater importance is that the core processes of interactive activation are not amenable to rapid transformation of representations or combinatorial rules. If one interlocutor can rapidly change another's internal weights, people would be subject to "catastrophic unlearning" of language.

As an alternative approach, we believe that interactive alignment may naturally "fall out" of processing in Grossberg's adaptive resonance theory (ART; Grossberg 1980; 1999; Grossberg & Myers 2000; Grossberg & Stone 1986; Grossberg et al. 1997; Vitevitch & Luce 1999). In speech perception, a basic hypothesis of ART is that conscious percepts are emergent products of resonant brain states. This is an interdependent system: Perception occurs when bottom-up and top-down knowledge sources bind into stable states. Processing begins when feature input activates *items* (feature clusters) in working memory. Items, in turn, activate *list chunks* in memory. Chunks are products of prior learning (perhaps prototypes) corresponding to feature combinations, such as phonemes, words, or common expressions.

Once items activate list chunks, a feedback cycle begins. Items feed activation upward through synaptic connections, and input-consistent chunks return activation. If items receive sufficient top-down confirmation, they continue sending activation upward. Within limits, this feedback loop (a *resonance*) is self-perpetuating, binding the respective activation patterns into a coherent whole. The bottom-up pattern that initiates interactive activation need not perfectly match its resultant feedback for resonance to occur. Cooperation between "levels" and competition within "levels" smooth out small mismatches, but large mismatches prohibit resonance. This stipulation seems necessary for dialogue, as another person's utterances (even their manner of producing phonemes) may never perfectly match one's stored representations.

Rather than assuming tiered representations, chunks in ART are *attractor states.* Of particular importance is that resonance can momentarily bind attractors in original combinations, allowing the creation of new mental structures (Carpenter & Grossberg 1987). This is the natural basis of communication, wherein novel utterances are constructed from familiar parts. In ART, however, those "parts" are not relegated to modular levels of representation: Although familiar patterns (e.g., common words) will cohere quickly in processing, all structures – both familiar and novel – must self-organize through competitive dynamics. Hence, in a dialogue, different structures, such as nascent jargon or momentary routines, can naturally emerge, as P&G describe.

As we have argued elsewhere (Goldinger & Azuma, in press), the self-organizing nature of ART makes it uniquely suited to model context-sensitive perception, as dialogue seems to require. For example, when ART recognizes a word spoken by Mary, its emergent percept will reflect the unique aspects of her utterance, allowing token-specific encoding (Goldinger 1998). Hence, the model has a natural basis to engage in dialogue, where context-specific cognition is pervasive. However, we are not suggesting that ART can currently explain the wide-ranging behaviors described by P&G. ART has successfully simulated many perceptual and memorial behaviors (among others; see Carpenter & Grossberg 1991). However, full sentence processing, even in monologue, remains a serious challenge.

As befits the study of dialogue, we envision a two-way street with equal benefits to students of dialogue and students of self-organizing networks. The properties of ART may give substance to the metaphoric theory described by P&G. By the same token, the unique challenges of dialogue may force changes in ART: To study dialogue, it would be ideal to have separate, self-contained ART simulations that communicate with each other through one or more channels, working together toward some shared goal (as in the authors' maze task). To our knowledge, no such study has ever been conducted using ART. However, a conceptually similar study by de Boer (2000) suggests that dialogues between autonomous models can evoke self-organization. De Boer created speech-perceiving "agents" equipped with articulatory synthesizers, perceptual systems that calculated the differences between signals, and associative memories for storing signals. Each agent was able to interact with others in an "imitation game." The agents' task was to imitate other agents accurately with a large repertoire of vowels; they could update their vowel repertoires based on their interactions. In many simulations, de Boer found that stable vowel systems consistently self-organized from these simple "dialogues." Moreover, these emergent vowel systems were remarkably similar to those found in human languages. This suggests that self-organizing structures may partially drive the basic forms of language. It also suggests that dialogue is a powerful force that should be studied both empirically and computationally.

ACKNOWLEDGMENTS

Support for this commentary was provided by grants R01-DC04535-03 (SDG) and 1-R03-DC4231-0182 from the National Institute of Deafness and Communicative Disorders (NIH). We thank Paul Luce, Greg Stone, and Guy Van Orden for many helpful discussions on the topics of resonance and speech units.

https://doi.org/10.1017/S0140525X04320054 Published online by Cambridge University Press