

nonalignment at the lexical level. For example, in Schober's (1993) example, two interlocutors who are facing each other use different terms to refer to similar locations (*on the left* vs. *on the right*) to maintain the same egocentric frame of reference. Likewise, Markman and Gentner (1993) show that successful use of analogy can require lexical misalignment. In Garrod and Anderson's (1987) maze game, if one player uses *second row* to refer to the second row from the top in a five-row maze, then the other player will tend to use *fourth row* to refer to the second row from the bottom. The player could lexically align by using *second row* in this way, but of course this would involve misalignment of situation models, and would therefore be misleading. The implication is that normally alignment at the situation level overrides alignment at lower levels.

5. We assume that a case, for example, where the speaker could not remember who he meant by *John* (while speaking) would be pathological.

6. Most theories accept that a few dialogue phenomena do need to be explained. For example, "binding" theory (Chomsky 1981) can be evoked to explain why *himself* is coreferential with *John* in A: *Who does John love?* B: *Himself*; though see Ginzburg (1999) for evidence against an account in such terms. Rather than think of question-answer pairs as a marginal phenomenon that needs special explanation in a monological account, we regard them as a particularly orderly aspect of dialogue.

7. Roughly, Ginzburg and Sag assume feature structures taken from Head-Driven Phrase Structure Grammar (Pollard & Sag 1994), in which context is incorporated into the representation of the fragments using the critical notion of QUDs ("questions under discussion").

8. Estimates from small group dialogues indicate that as many as 31% of turns are interrupted by the listener (Fay et al. 2000).

9. Jackendoff uses the term *conceptual structures* instead of *semantic structures*, for reasons that we shall ignore for current purposes.

10. Note that Jackendoff (2002) assumes interface rules between semantic (conceptual) structures and phonological structures (p. 127, Fig. 5.5). If this is correct, it suggests that Figure 2 should incorporate such a link as well. He also suggests that the lexicon should be regarded as part of the interface components (p. 131).

11. The tendency might even be stronger for young children than adults, at least when it is the verb that is repeated. According to the "verb island hypothesis," syntactic information is more strongly associated with individual verbs in young children than it is in adults (e.g., children are often able to use a particular construction with some verbs but not others; Tomasello 2000).

Open Peer Commentary

Is language processing different in dialogue?

Dale J. Barr^a and Boaz Keysar^b

^aDepartment of Psychology, University of California, Riverside, CA 92521;

^bDepartment of Psychology, University of Chicago, Chicago, IL 60637.

dale.barr@ucr.edu boaz@uchicago.edu

Abstract: Pickering & Garrod (P&G) claim that the automatic mechanisms that underlie language processing in dialogue are absent in monologue. We disagree with this claim, and argue that dialogue simply provides a different context in which the same basic processes operate.

Pickering & Garrod (P&G) call for closer attention to the mechanisms underlying coordination in dialogue. There are good empirical grounds for accepting many of the basic assumptions of the interactive alignment model. Specifically, the strong egocentrism of speakers and listeners that we have uncovered in our own studies (cf. Barr & Keysar [in press], for a recent review) makes much sense within a context of strong representational overlap, and the interactive alignment model provides an appealing explanation for how such overlap comes about.

A central assumption of P&G is a categorical distinction between language processing in monologue and in dialogue. Clearly, monologue is different from dialogue because in monologue there is no feedback and no opportunity for interactive repair. But P&G go further than this, asserting that in monologue, "the automatic mechanisms of alignment are not present" (sect. 3.2, para. 3) and that "there is no opportunity to call on aligned linguistic representations" (sect. 8.1, para. 5). Under this view, dialogue involves processes that are fundamentally distinct from those present in monologue. We disagree. We suggest that only the strategic mechanisms of feedback and interactive repair are absent in monologue, not the automatic mechanisms of alignment. Just as there is a "dialogic continuum" defining different kinds of interactive activities, we argue that there is a continuum of processing and that alignment will be observed even under non-interactive circumstances.

Differences in processing may simply be a matter of degree, not of kind, especially given that monologue-like episodes are common even in naturalistic conversation. Consider, for example, the first utterance of a conversation, which in many cases is a mini-monologue before a full-fledged interactive exchange develops. The speaker will need to go through the various stages to produce the first utterance (e.g., Levelt 1989), and the comprehender must parse the utterance in order to appreciate its significance. Later on in the exchange, enough of shared but nonmutual information (what P&G unfortunately term "implicit common ground")¹ may have built up to short-circuit many aspects of these processes. But there is no reason to believe that different theories would be required to explain the processing of the first utterance of a dialogue versus the hundred-and-first. Indeed, P&G claim that the alignment processes are automatic and resource-free, and processes of these sorts cannot simply be switched off. Finally, the "dialogic continuum" cited by the authors does not just provide a means for classifying whole conversations but actually represents a domain of activity that can be fully traversed even within the span of a single conversation. A conversationalist can at any point secure turn space in order to engage in an extended monologue (Sacks et al. 1974) – for example, about her trip to India – followed by close coordination with her interlocutor in order to arrange a time for a future meeting. There is no reason to expect alignment processes and bridging inferences to toggle on and off as the interactivity of the discourse changes.

Studies that find differences between the quality of comprehension of interlocutors and of overhearers (e.g., Schöber & Clark 1989; Wilkes-Gibbs & Clark 1992) might be taken as evidence to support the idea of radically different cognitive processes in interactive discourse. They need not be. In such studies interlocutors had opportunities for feedback and repair that the overhearers lacked. Because different people will misunderstand different things, those who can ask for clarification will receive feedback that is relevant to them and consequently might understand better. Therefore any difference between such noninteractive and interactive comprehension could be fully attributable to strategic, effortful feedback but not necessarily to automatic alignment. In fact, Barr and Keysar (2002) found that even when such feedback is removed, listeners who believed themselves to be overhearers automatically aligned their semantic representations with the speaker's to the same degree as listeners who believed themselves to be addressees.

In closing, far from qualitatively changing the nature of processing, it is likely that dialogue provides a radically different context in which the same processes operate. The context includes an interlocutor and mechanisms for feedback and interactive repair. For us there is no question that it is important to study conversation in vivo, but it remains to be seen whether this would reveal automatic processes that are truly unique to dialogue.

NOTE

1. In their seminal work on common ground, Clark and Marshall (1981) clearly make the case that common ground is a form of meta-knowledge that is conceptually distinct from shared knowledge. What P&G are referring to by "implicit common ground" is really just shared knowledge, not common ground, because interlocutors need not represent the fact that their representations are shared. Such usage is certain to contribute to the legacy of confusion that has plagued discussions of mutual knowledge and common ground (see Keysar 1997 and Lee 2001 for discussion).

Full alignment of some but not all representations in dialogue

Holly P. Branigan

Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom. holly@cogsci.ed.ac.uk
<http://www.psy.ed.ac.uk/Staff/ra/hbranigan/publications.html>

Abstract: I argue that alignment of linguistic representations and situation models in dialogue are qualitatively distinct. By virtue of the isomorphism between interlocutors' linguistic representations, interlocutors align their linguistic representations fully. However, evidence about situation models is indirect and mediated through language, with the result that alignment of situation models is only partial.

Pickering & Garrod (P&G) provide a plausible and very welcome account of language processing in dialogue. Their account assigns a central importance to the notion of alignment. Here I consider the nature of alignment and, in particular, whether alignment of strictly linguistic representations and alignment of situation models are qualitatively distinct.

P&G suggest that successful dialogue arises from the alignment of representations between interlocutors, and particularly alignment of situation models. Alignment is defined as interlocutors having the same representation at a particular level of structure. It is uncontroversial that language is represented in the same way in different speakers. Linguists and psycholinguists assume a common competence grammar in adult speakers (Chomsky 1965), even though they may differ as to whether this grammar is innate (e.g., Pinker 1989) or to some extent constructed through experience (e.g., Tomasello 2000). Crucially, adult speakers' internalised

knowledge of the syntax, lexicon, and morpho-phonology of a language is held to be identical, such that there is an isomorphic mapping from any one speaker's internalised representation of the language to any other's. In a dialogue, then, interlocutors necessarily make use of identical representations in producing their utterances.

Under P&G's maximally parsimonious assumption of parity of representations, interlocutors also necessarily draw upon identical representations in both producing and comprehending utterances. Note also that speakers' utterances provide direct linguistic evidence to the listener. So when a listener hears an utterance, he receives direct evidence (except in cases of mishearing or unresolved structural ambiguities) about the syntactic, lexical, and morpho-phonological representations that the speaker has employed. If a speaker produces an utterance like *I am in row two*, for example, the listener has direct evidence that she has used the words *I*, *am*, and so on (and their relevant inflectional markings), that she has used a pronoun and a verb and so on, and that she has used a noun phrase, a verb phrase, a prepositional phrase, and so on. Taken together, the combination of isomorphy of representations and direct evidence strongly supports P&G's contention (summarised in their Fig. 2) that linguistic representations used by interlocutors in dialogue act directly upon one another, and that, in a very real sense, when we talk about interlocutors having aligned linguistic representations, we mean that those representations are identical. In summary, P&G's arguments for full alignment at linguistic levels of representation seem well founded.

But are situation models aligned in the same way as linguistic representations? It is unclear that this is the case. In P&G's model, interlocutors' situation models act directly upon one another (see the authors' Fig. 2), in the same way as syntactic, lexical, and morpho-phonological representations do, and alignment of situation models is taken as critical for successful communication. But situation models differ qualitatively from strictly linguistic representations. A speaker's utterances do not give direct evidence of the situation model that the speaker holds, only indirect evidence encoded in linguistic representations, from which the listener has to infer the speaker's situation model. So whereas an utterance like *I am in row two* gives direct evidence about the speaker's syntactic, lexical, and morpho-phonological representations, it gives only indirect evidence about the speaker's situation model. The listener must construct a situation model based upon his or her interpretation of the speaker's meaning – which may or may not be correct. Of course, as P&G note, misunderstandings may come to light, and interlocutors may initiate repairs to bring about situation models that are aligned in the relevant aspects. But as they also note, some misunderstandings may not be repaired. In fact, it seems likely that interlocutors quite frequently have situation models that are misaligned in major respects. Communication will be (apparently) successful as long as the misalignment is not apparent to the interlocutors. To take P&G's example of interlocutors using *John* to refer to different people, it is quite possible for them to have a mutually satisfying dialogue concerning this person without ever realising that they are discussing different people; unless one of them says something that is inconsistent with the other's knowledge, they can successfully (for their purposes) complete a dialogue with quite radically different situation models. Equally, a doctor and a patient may have a dialogue concerning the patient's *chronic back problem* that appears to be successful, in that they are both satisfied that they understand each other well; yet their situation models may differ considerably because of unresolved (and unapparent) differences in their interpretation of *chronic*. Situation models need only be aligned sufficiently for the current communicative goal to be (apparently) met.

So it seems that alignment of situation models and alignment of linguistic representations are quite different. With linguistic representations, interlocutors genuinely employ aligned (i.e., identical) representations that act directly upon one another; whereas, because evidence for situation models is only indirect, interlocu-