knowing what the partner model would clarify? (This seems akin to Fodor's frame problem or the notorious conundrums about selective attentional filtering). Arguably, it might be simpler to model the partner as the ordinary case. P&G assume that representing common ground requires unusually complex mental representations; this may be so, but it is also possible that representations of common ground are ordinary memory representations with the right content.

In fact, what is parsimonious may well vary across individuals and conversational contexts. Our notions of what is effortful in dialogue are not well worked out (Schober 1995; 1998a); people's knowledge of each other's perspectives, interest in modeling their partners, and abilities for doing so probably vary substantially more than current theories allow. For example, people who have very poor mental rotation abilities may be unable to conceive of a partner's point of view on a scene, leading to a different sort of partner modeling than that carried out by mental rotation whizzes (Schober 1998b). And different abilities may be relevant in different conversational contexts; mental rotation ability may have little to do with the consideration of lexical alternatives in dialogues that are not about physical scenes, whereas working memory capacity differences or different conversational agendas (Russell & Schober 1999) might. P&G are to be applauded for furthering the attempts to bridge the "language as product" and "language as action" traditions, but they have also opened a can of worms; the ultimate mechanistic account of dialogue is likely to require far greater specification of individual and contextual variability.

Dialogue processing: Automatic alignment or controlled understanding?

Hadas Shintel and Howard C. Nusbaum

Department of Psychology, The University of Chicago, Chicago, IL 60637.
hadas@uchicago.edu **hcn@speech.spc.uchicago.edu**

Abstract: Pickering & Garrod's (P&G's) mechanistic account of dialogue assumes that linguistic alignment between interlocutors takes place automatically, without using cognitive resources. However, even the most basic processes of speech perception depend on resource use. The lack of invariant mapping between input patterns and interpretations in dialogue, as in speech perception, may require controlled, rather than automatic, processing.

In their target article, Pickering & Garrod (P&G) challenge current psycholinguistic theory by proposing the study of dialogue as the appropriate paradigm for understanding language processing. They argue that the main problem facing interlocutors is the alignment of their respective situation models, and they propose a mechanistic account of alignment in dialogue, based on an interactive and resource-free priming mechanism. Linguistic structures, produced on the surface by one interlocutor, putatively prime corresponding structures in the other interlocutor. While we agree that conversation, rather than decontextualized, isolated sentences, should be taken as basic form of language use and as the foundation upon which language has evolved, it seems less plausible that much of the cognitive work in conversation takes place automatically.

P&G's model of automatic alignment assumes that a mapping can be established directly between the representations uttered by one interlocutor and those activated in the head of the other interlocutor. But as they themselves note, the linguistic information is encoded in sound. Hence, for the listener the spoken utterance must be recognized antecedent to any alignment occurring. However, even findings regarding the basic processes of speech perception argue against the assumption of automatic processing. Automaticity implies a passive process in which the input is processed in an invariant, inflexible manner, regardless of the beliefs and expectations of the listener (Nusbaum & Schwab 1986). For a process to be automatized, there must be a consistent mapping between input patterns and responses, the benefit of which is a process that does not impose demands on cognitive and attentional resources (Shiffrin & Schneider 1977).

Of course the hallmark of speech is the lack of consistent, invariant mapping between acoustic patterns and linguistic categories across phonetic contexts (Liberman et al. 1967) and across talkers (Peterson & Barney 1952). As a result of variable mapping, speech perception depends on controlled, active processing requiring attention and working memory (Nusbaum & Magnuson 1997). For example, variation between talkers (as in the circumstance of a conversation among three interlocutors) requires talker normalization, the process by which listeners compensate for acoustic-phonetic variability in vocal characteristics (see, e.g., Nearey 1989). Nusbaum and Morin (1992) found that talker variability slowed speech recognition (see also Mullennix & Pisoni 1990), and this slowing was because of increased demand on working memory.

Furthermore, talker normalization is contingent on listeners' expectations regarding the interpretation of acoustic patterns – which should not happen for an automatic direct-mapping process. Magnuson and Nusbaum (1994) demonstrated that when listeners expected a pitch difference to signal a talker difference, they showed talker normalization, but if the same pitch difference was expected as a way one talker accented speech, no normalization occurred. Similarly, the expectation that a talker was male or female significantly changed the interpretation of vowel tokens (Johnson et al. 1999). The acoustic patterns of speech are processed differently depending on listeners' expectations, arguing against invariant automatic processing. Expectation effects suggest that alignment may not be a result of a direct, automatic, causal link between the activation of a representation in one interlocutor and the activation of an objectively matching representation in the other interlocutor. Rather, the process is mediated by listeners' expectations. For example, alignment at the level of articulation, where interlocutors converge on each other's speaking style, may be mediated by interlocutors' stereotypic expectations about the other interlocutor's accent and speech rate, resulting in subjective but not objective alignment (Thakerar et al. 1982).

Listeners may need to use controlled active processing because the one-to-many mapping (one pattern may have multiple interpretations) in speech represents a nondeterministic computational problem that cannot be solved, in principle, by a deterministic system such as an automatic process (cf. Nusbaum & Magnuson 1997). This problem of one-to-many relationships between linguistic patterns and interpretations occurs across all levels of linguistic analysis (Nusbaum & Henly 1989). Indeed, research has shown that the same spoken sentence can be differently interpreted in different visual referential contexts (Tanenhaus et al. 1995) and that the same indirect requests can be processed differently as a result of speaker status (Holtgraves 1994). Although P&G focused on utterance-level matches that can occur within stretches of dialogue, real conversation is less predictable and routinized – otherwise communication would be unnecessary.

This raises several deep questions facing dialogue research: How do interlocutors cope with variability in linguistic and perceptual context? How are diverse sources of information integrated to constrain production and comprehension? And how do interlocutors flexibly adapt to different conversational circumstances? Although in real conversations these questions are often subjectively resolved quickly and without apparent effort, this is seldom an accurate barometer of the demands on cognitive resources such as attention and working memory.

P&G propose that the process of automatic alignment bypasses the need for modeling the interlocutor's mental states and for distinct, conscious decision stages. We agree that a constant monitoring of common ground would be unnecessary and costly, but the dichotomy between automatic processes and conscious decision processes involving complex inferences about the interlocutor's mental state does not represent the full range of possible pro-

https://doi.org/10.1017/S0140525X04430052 Published online by Cambridge University Press

cessing alternatives. In clarifying this point, it is worth considering what is meant by automatic, as distinct from controlled, processing. Automatic processes have been defined by different criteria such as being unintentional, occurring outside of conscious awareness, not requiring cognitive resources, and being autonomous; but these criteria do not necessarily hold simultaneously (Bargh 1989). Speech perception involves controlled active processes that are not resource-free or autonomous, yet they do occur largely outside of awareness, do not require a conscious intention on the part of the listener, and are subjectively experienced as effortless. Likewise, the subjective ease and speed that subjectively characterize language use in dialogue may not reflect the complexity of the underlying processing. The variability and flexibility shown in the processes of language comprehension and production in dialogue call for dynamic adaptation rather than a passive automatic mechanism.

Top-down influences in the interactive alignment model: The power of the situation model

Tessa Warren[a] and Keith Rayner[b]

[a]*607 LRDC (Learning, Research, and Development Center) and Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260;* [b]*Department of Psychology, University of Massachusetts, Amherst, Amherst, MA 01003.*
tessa@pitt.edu **rayner@psych.umass.edu**

Abstract: Pickering & Garrod's (P&G's) model is an innovative and important step in the study of naturalistic language. However, the simplicity of its mechanisms for dialogue coordination may be overstated and the hypothesized direct priming channel between interlocutors' situation models is questionable. A complete specification of the model will require more investigation of the role of top-down inhibition among representations.

Pickering & Garrod's (P&G's) new model of linguistic interaction in dialogue is an important contribution to the study of psycholinguistics. This model breaks new ground between two traditionally disparate areas of language research and combines the mechanistic detail characteristic of sentence processing research with the emphasis on language as a cooperative process characteristic of dialogue research.

The central mechanism of P&G's interactive alignment model is a process they term "alignment," whereby dialogue participants' linguistic and discourse representations become more similar over the course of an interaction. Alignment is an automatic process that results from priming between linguistic representations. Priming can occur among different levels of representation within a single individual, as when increased lexical or semantic overlap between a prime and a target causes increased syntactic priming, or between the representations of different individuals, as when a speaker who has just comprehended a particular syntactic form produces the same form in a subsequent utterance. The resource-free characterization of the system relies on the assumption that alignment at one level of representation increases alignment at other levels of representation, and therefore essentially complete alignment can be obtained through simple and automatic processes.

Though the strength of this model is in its elegance and simplicity, it is not clear whether these characteristics will survive a more explicit specification of the system. According to the characterization of the model as a network of linked representations where priming and alignment in one representation causes increased alignment in all others, alignment at the level of the situation model is simply an epiphenomenon of alignment at lower levels. Yet P&G acknowledge that in some cases priming at one level will decrease alignment at other levels and specifically note that alignment of the situation model takes priority over alignment at other levels of representation. This complicates the system, as lower-level priming is constrained by the very alignment it drives. In fact, lower-level priming, supposedly a driving force in alignment, can be inhibited by extremely subtle nuances of local context. For example, P&G discuss Garrod and Anderson's (1987) findings that participants used different words for horizontal groupings of boxes depending on whether the grouping was modified with an ordinal adjective or not. So participants who used the phrase "the second row" later spoke of "the bottom line," even though "the bottom row" was an acceptable, unambiguous, and more lexically primed alternative. From examples like this it is clear that priming and alignment at the level of the situation model are very powerful mechanisms in the system, and that they have significant inhibitory power over the automatic, lower-level priming that is hypothesized to drive alignment at all levels of the system. It remains to be seen whether bottom-up mechanisms are robust enough to drive alignment, as P&G claim, or whether top-down inhibition directs the system.

The interactive alignment model relies heavily on the assumption that the situation model of one individual can directly prime the situation model of another individual through the same automatic mechanisms that are responsible for phonological, lexical, semantic, and syntactic priming between individuals. This assumption seems hasty. Phonological, lexical, and syntactic priming are similar in that the representation that will eventually be primed is an inalienable part of the structure of the message that causes the priming. The words, sounds, and structural patterns that make up an input string necessarily activate representations for exactly those words, sounds, and patterns. Semantic priming is not as direct, as semantic representations cannot be directly read off an input string. However, semantic representations are generally similar across individuals. For example, it is safe to assume that the word "dog" will activate the concept "cat" more strongly than the concept "book" for the wide majority of individuals. This similarity of representation may allow for what P&G represent as a direct priming link between the semantic representations of different individuals. If an individual produces the word "goat," the concepts that will be primed in her own semantic representation are likely similar to the concepts that will be primed in her interlocutor's representation, because the representations are structured in a similar way. In this indirect way, it could be said that the semantic representation of one individual can prime the representation of another.

The same thing cannot be said of situation models. Like semantic priming and unlike phonological, lexical, and syntactic priming, there is no direct priming channel between individuals through physical aspects of a message such as sounds, words, or word patterns. Again, priming must be indirect, through the activation of words or phrases that suggest a particular state or property of a situation model. But unlike semantic representations, individuals do not necessarily begin dialogue with similar situation representations. Therefore the priming link between individuals' situation models in the interactive alignment model must be of a different sort from the priming channels between other representations.

The interactive alignment model opens a new and exciting area of inquiry into language processing. However, more research into the details of situation model priming and the complex interplay of priming and inhibition between different levels of representation will be necessary in order to fully specify the operation of the model and to evaluate its ascribed simplicity. Carrying out this research will not be easy, as it will be difficult to maintain the necessary experimental control in the sorts of experiments that will be required. But as P&G optimistically point out, "Well-controlled studies . . . may require some ingenuity, but such experimental ingenuity has always been a strength of psychology" (target article, sect. 1, para. 4).

https://doi.org/10.1017/S0140525X04430052 Published online by Cambridge University Press