pp. 729–31), and can be seen as particular, here-and-now-with-these-participants instances of these.

What makes an interaction is not just the juxtaposition of bodies. What mediates and organizes the conduct of the parties is not a structureless, featureless, transparent medium. The composition of a turn at talk – whether it is made up of one or more component units; whether these are sentences or sub-sentential – its syntactic construction and choice of lexicon are shaped in part by the contingencies of turn production imposed by a turn-taking organization that will have others empowered or required or allowed to talk next, at points in the turn's development not wholly under the speaker's control. Particular courses of action implemented through turns at talk (such as request sequences, complaint sequences, storytelling sequences, news-conveying sequences, etc.) implicate certain ways of understanding what is being said that render meaningful and consequential selection between apparently equivalent expressions, the delay of a turn's start by two-tenths of a second or less, and the like. How one says what one says can depend on who the other is; and, of all the persons and categories which could be used to characterize "the other," depend on which ones have been made relevant at that moment in the talk, or can be made relevant by constructing the same "sayable" in this way or that. And so on.

A very high proportion of the matters discussed by P&G as if they were unrelated to anything but the mechanisms the authors are concerned to develop, are not interactionally random. They are part of the fabric of some organization of practices for talk-in-interaction. Many of them have been given quite detailed and systematic treatment in the literature – things like "routines" (target article, sect. 5.2.1, cf. Schegloff 1986) and "how are you" routines in particular (Jefferson 1980; Sacks 1975); things like "joint constructions" (sect. 7.1, para. 3; cf. Lerner 1991; 1996; Sacks 1992, vol. I, pp. 144–47 et passim); things like "non-sentential turns" (sect. 7.1, para. 6; cf. Sacks et al. 1974; Schegloff 1996a): things like "monitoring during overlapping speech (sect. 6, para. 6; cf. Schegloff 2000a; 2001); and so on and on.

Most striking is P&G's treatment of "repair"; the discussion rests on a terminology ("repair," "other-repair," "self-repair") which they neither explicate nor cite but the latter two of which they treat as discrete sets of things, not an organization of practices. This leads them – incorrectly, in my view – to treat the basic mechanisms of self-repair and other-repair as the same (see Table 2 of the target article) when, interactionally speaking, they are not the same in either execution or interactional import (Schegloff 1979, pp. 267–69; Schegloff et al. 1977, inter alia). I believe the analysis of talk-in-interaction along such lines has much to contribute not only to our understanding of the mechanisms addressed by P&G, but to work in the neurobiology of behavior more generally – precisely the remit of this journal. But that is another matter.

Some notes on priming, alignment, and self-monitoring

Niels O. Schiller[a,b] and Jan Peter de Ruiter[b]

[a]*Department of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, 6200 MD Maastricht, The Netherlands;* [b]*Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands.*

n.schiller@psychology.unimaas.nl **Niels.Schiller@mpi.nl**
JanPeter.deRuiter@mpi.nl
http://www.psychology.unimaas.nl/Base/Personal/NielsSchiller.htm
http://www.mpi.nl/world/persons/profession/schiller.html

Abstract: Any complete theory of speaking must take the dialogical function of language use into account. Pickering & Garrod (P&G) make some progress on this point. However, we question whether their interactive alignment model is the optimal approach. In this commentary, we specifically criticize (1) their notion of *alignment* being implemented through *priming,* and (2) their claim that *self-monitoring* can occur at all levels of linguistic representation.

The primary way of language use is *dialogue,* not monologue. We want to acknowledge the authors' effort to stress this important point, which needs to be addressed explicitly in empirical and modeling work in speech production and comprehension research. We believe that these issues are especially relevant for syntactic processing. For instance, one wonders how syntactically incomplete (dialogue) utterances can be syntactically encoded in more traditional models, if there is no overt verb present in the generated utterance. Take, for example, the following extract from the dialogue transcript in section 2 of the target article:

> 1——**B:** . . . Tell me where you are?
> [Utterances 2 and 3 omitted]
> 4——**A:** Right: {*I am*} **two along from the bottom one up:*** [our addition in curly brackets]

In this example, speaker A does not produce the appropriate verb form of "to be" (i.e., "I am") but nevertheless gives an acceptable and cooperative answer to speaker B's question. This type of ellipsis can only be correctly produced if the syntax generator has access to previously stored discourse information, allowing the speaker to omit "I am," even though the original question containing the verb occurred several utterances earlier in the discourse (see also Levelt 1989, p. 89, for a similar analysis).

Although we agree in principle with the authors' assessment that the dialogical structure of language should receive more attention in accounts of language processing, we are not convinced that adopting the interactive alignment model is the right way to do so. For instance, it is unclear to us exactly how *priming* can account for *alignment,* and, in particular, we fail to see in what way priming is more than "a behavioral effect" (see target article, sect. 2.2). We believe that "priming" *does not explain or implement interactive alignment.* Real interactive alignment necessarily involves *storing selected fragments* from previous utterances. Priming can raise the probability of certain linguistic structures being selected, but this is not sufficient for the strong and explicit type of alignment the authors want to incorporate in models of language processing. Also, syntactic priming effects are weak effects. It is hard to see how an elaborate mechanism such as interactive alignment could be realized by only raising the probability of selecting a certain syntactic construct by roughly 10% to 20% (see, e.g., Pickering & Branigan 1998).

Our second critical note concerns one of the few testable predictions from the interactive alignment model, namely, that self-monitoring by the speaker occurs at all levels of linguistic representation (see sect. 6). While other researchers (e.g., Wheeldon & Levelt 1995) have claimed that internal self-monitoring works on abstract phonological form representations, Pickering & Garrod (P&G) propose that self-monitoring can occur at any level of linguistic representation that can be aligned (i.e., semantic, syntactic, lexical, phonological, and phonetic representations) – and not only at the phonological level.

For example, the authors explicitly claim that speakers can correct gender errors, such as *le tête* instead of *la tête* ("the head") in French or *de been* instead of *het been* ("the leg") in Dutch not only after they have been articulated but even before their overt production. This is an interesting claim that needs to be investigated in the future. However, we are somewhat skeptical about this claim because to our knowledge there is no evidence that self-monitoring of gender features (or any other syntactic features) is possible. For example, Desrochers and his collaborators (Desrochers & Paivio 1990; Desrochers et al. 1989; Muller-Gass et al. 2000) found that selecting a gender label (e.g., feminine or masculine) took about 200 msec longer than selecting the indefinite article in French gender decision. Furthermore, Tucker et al. (1977) provided empirical evidence suggesting that French speakers implicitly construct a noun phrase including the article and the noun to determine a noun's gender. However, if speakers can self-monitor abstract gender information at the level of syntactic representation, as suggested by P&G, why would they go through the trouble of generating the gender-marked article of a noun to determine its gender?

https://doi.org/10.1017/S0140525X0441005X Published online by Cambridge University Press

In contrast to these findings about syntactic representations, recent evidence from our own laboratory as well as from other laboratories demonstrated that self-monitoring *does* occur at the level of phonological encoding. We have empirical data about the monitoring of phonological segments (Schiller, in press; Wheeldon & Morgan 2002), word stress (Schiller 2001; Schiller et al., in press), syllable boundaries (Jansma & Schiller 2004), and syllables (Morgan & Wheeldon 2003). However, we also have evidence that participants are unlikely to monitor a phonetic-acoustic representation of the respective utterances. Although gender decision as a task is widely used (Müller & Hagoort 2001; Schiller et al. 2003; Schmitt et al. 2001a; 2001b; Van Turennout et al. 1998), it remains to be shown whether or not this task actually taps syntactic processing, because abstract gender information may not be directly available to the speaker. Rather, gender information may be available only via its phonological realization, for example, an article or gender-marking suffix. Interestingly, effects of *gender congruency* have recently been re-interpreted as *determiner congruency* effects occurring at the phonological level and not at the gender feature level (see overview in Schiller & Caramazza 2003).

To conclude, we believe that internal self-monitoring does not occur at every single level of linguistic representation (as claimed by P&G) or at every processing level in models of speech production (e.g., Levelt et al. 1999). Rather, there is abundant evidence that internal self-monitoring works on phonological representations, which are created during phonological encoding in speaking, for example, when segments are prosodified into phonological words. It is at this level that information about segments, syllables, syllable boundaries, and word stress is available to the speaker. Although the abundance of evidence for phonological-level monitoring does not necessarily exclude other monitoring levels (e.g., at the conceptual level; see Levelt 1989), we are unaware of any evidence for self-monitoring at earlier or later levels of grammatical and form encoding.

ACKNOWLEDGMENTS

Niels O. Schiller is supported by a grant from the Netherlands Organization for Scientific Research (NWO). Jan Peter de Ruiter is supported by the European Union Project Conversational Multi-modal Interaction with Computers (COMIC), grant IST-2001-32311. The authors thank Pim Levelt for his comments on an earlier draft.

Just how aligned are interlocutors' representations?

Michael F. Schober

Department of Psychology, F330, New School for Social Research, New York, NY 10003. **schober@newschool.edu**
http://homepage.newschool.edu/~schober

Abstract: Conversational partners' representations may be less aligned than they appear even when interlocutors believe they have successfully understood each other, as data from a series of experiments on surveys about facts and behaviors suggest. Although the goal of a mechanistic psychology of dialogue is laudable, the ultimate model is likely to require far greater specification of individual and contextual variability.

When conversational partners come to agree that they are talking about the same thing, how aligned are their conceptualizations? The interactive alignment account holds that when dialogue is successful, interlocutors' linguistic representations are aligned at multiple levels. Although Pickering & Garrod (P&G) observe that alignment is sometimes only approximate and that evident misalignments can be interactively repaired, the general thrust of their approach is that successful communication requires representations to be the same.

I contend that interlocutors' using the same words can actually mask a surprising degree of undetected misalignment. Fred Conrad and I, with other colleagues, have carried out a series of laboratory and field studies examining how people interpret ordinary words in ongoing U.S. government surveys about facts and behaviors, words like "job," "bedroom," "smoking," and "cigarettes" (Conrad & Schober 2000; Schober & Conrad 1997; Schober et al. 2004; Suessbrick et al. 2000). Because the agencies that carry out these surveys have thorough definitions for the terms, answers to the survey questions provide evidence about the extent to which respondents' conceptions match the survey designers'. Our finding is that people's representations are frequently wildly misaligned with the survey designers' – and with each other's – without anyone's noticing.

For example, in one study (Suessbrick et al. 2000), survey respondents interpreted terms like "smoking" and "cigarettes" in a question such as "have you smoked at least one hundred cigarettes in your entire life?" differently enough (tobacco, cloves, marijuana? Finished or just a puff? Bought or borrowed?) that 10% of the respondents subsequently presented with a definition changed their answer to the question from yes to no or from no to yes. In a national telephone sample (Conrad & Schober 2000), more than 40% of reported purchases did not fit the survey designers' definitions, even though the questions had been widely pretested. And this is not just because the official definitions failed to match the population consensus about the meaning of terms; respondents' interpretations differed from each other's as much as they differed from the survey designers'.

Across our various studies, respondents are quite surprised at the thought that someone else might interpret the same words differently from the way they do; when given the opportunity to request clarification about the meanings of survey terms, they choose to do it a very small percentage of the time. People seem to follow a "presumption of interpretability" (Clark & Schober 1991): It should be the questioner's responsibility to forestall misinterpretation.

These data suggest a far more Quinian view of successful referring than the P&G account encompasses: Seemingly successful referring can mask conceptual misalignments that reflect deep underlying indeterminacies. The point is that people can believe they have understood each other well enough for current purposes (as proposed in Clark & Wilkes-Gibbs 1986) and yet never actually discover that their conceptions were misaligned.

An important contention in the P&G article is that seemingly complex interactional processes can be modeled largely with simple individual mechanistic processes. The proposal is that conversational partners, following a principle of parsimony, do not ordinarily model each other's mental states or make inferences about common ground, except when there is evidence that not doing so has led to obvious misunderstanding and when cognitive resources allow.

I would argue that none of the current data actually allow us to distinguish this position from an alternative: that the ordinary case is that conversational partners *do* model each other, and that they fail to do so only when they are under heavy cognitive load or when circumstances weigh heavily against doing so. Why should we assume that the ordinary case is one where the interlocutor does not need to be modeled and the speaker is under heavy cognitive load? As far as I can tell, no one knows the level of load encountered in the range of ordinary interactive situations. As Susan Brennan and I have argued (Schober & Brennan 2003), the evidence for egocentric processing is far from conclusive; the experiments purported to show egocentric processing as basic rely on null results and experimental methods that are far removed from ordinary processing situations. When such studies are carried out in more realistic settings, the findings can look rather less egocentric.

Not to overstate the case, but one could argue that modeling one's partner only when it is needed may require a level of situational monitoring that leads to a paradox: How can one know exactly when one needs to model one's partner without already

https://doi.org/10.1017/S0140525X0441005X Published online by Cambridge University Press