

often pressure on speakers to incorporate a high density of autoclitics according to constraints embodied in formal rules of grammar. When these pressures are weak, as in much of dialogue as well as certain literary styles (Skinner 1957, p. 356), verbal behavior is more fluid. Skinner likely would have endorsed the call in the last sentence of the target article for “a more flexible account of grammar capable of capturing linguistic constraints on linked sentence fragments.”

Regarding row 3, a routine in the IA account appears to be, in the VB account, a functional verbal unit (Skinner 1957, pp. 21, 116) that has been conditioned or strengthened in a specific situation. In the IA account, repetition of “the previous speaker’s utterance” appears to be important in this process (sect. 5.1 of target article). According to Skinner, “a verbal response of a given form sometimes seems to pass easily from one type of operant to another” (Skinner 1957, p. 188). Hence, a response emitted as an instance of echoic behavior may simultaneously or subsequently appear in other categories of verbal behavior (for examples, see Skinner 1957, pp. 188–89, 360–62).

Like routines in the IA account, functional verbal units in the VB account may be larger than a single word. Similar to the process of routinization in the IA account (sect. 5.2 of the target article), a process called composition in the VB account generates large verbal patterns that can come to function as units. Skinner (1957, Ch. 14, pp. 344–67) proposed that composition consists primarily of adding autoclitics (including ordering) to the raw verbal material mentioned above. “Formal evidence alone will not show whether sentences [or other large segments of verbal behavior] have been composed” (Skinner 1957, p. 346) as opposed to being emitted as units. Once a composed utterance has been reinforced several times, it may begin to function as a unit. Given the right conditions, however, a unit may break into smaller units (Skinner 1957, pp. 116–17). This seems very close to the dynamics of routines as described in the IA account.

Regarding row 4, in the VB account “[a]n important fact about verbal behavior is that speaker and listener may reside within the same skin” (Skinner 1957, p. 163). As in the IA account, in the VB account speakers monitor their own verbal behavior and edit it after, during, or even before its emission (Skinner 1957, Ch. 15, 16, pp. 369–402). Both accounts agree that there is no difference in principle between self- and other-monitoring. Both accounts also agree that when interacting with oneself, the stimuli need not be in the form of an external medium.

Regarding row 5, the techniques for strengthening one’s own weak verbal behavior are in principle the same as those for strengthening another’s verbal behavior (Skinner 1957, Ch. 17, pp. 403–17). These include manipulating stimuli and changing the level of editing.

Finally, regarding row 6, by focusing on contextual stimuli, the VB account provides a flexible account of grammar.

Given the correspondences between the two accounts, it may be impossible to distinguish them empirically. However, the VB account appears to require fewer terms, “and the terms created are derived from a few prior technical terms common to the whole field of human behavior” (Skinner 1957, p. 456).

Putting the interaction back into dialogue

Emanuel A. Schegloff

Department of Sociology, University of California at Los Angeles, Los Angeles, CA 90095-1551. schegloff@soc.ucla.edu
<http://www.sscnet.ucla.edu/soc/faculty/schegloff/>

Abstract: I share the authors’ stance on the dialogic or interactional character of language. The authors, however, have left actual interaction out of their conception of dialogue. I sketch a number of organizations of practices of talking and understanding that supply the basic arena for talk-in-interaction. It is by reference to these that mechanisms for speech production and understanding need to be understood.

I write as a conversation analyst. I have spent nearly 40 years studying the sorts of events which Pickering & Garrod (P&G) now take to be the fundamental premise of language. I am, of course, predisposed to take the same view. Indeed I have already done so in a number of publications (Schegloff 1979; 1989; 1996a, inter alia). The arguments of the target article aside, in the choice between a fundamentally monologic, “individualist” stance and a dialogic, interactional one, there are compelling reasons for preferring the latter. For now, one line will have to do.

For most humans on the planet since the species developed “language,” the overwhelmingly most common ecological niche for its use has been (1) the *turn* at talk, (2) as part of a coherent *sequence of turns*, (3) through which a *course or trajectory of action* is jointly pursued by some or all of the participants (not necessarily cooperatively, but jointly), (4) in an *episode of interaction*, (5) between two or more *persons*, (6) organized into two or more *parties*, (7) the *occasion* of interaction being composed of one or more such episodes. If that is where language as a publicly deployed resource and utility resides, it is plausible to expect that it has been designed and fashioned by its users and uses in a manner adapted to the contingencies of its “environment” – that is, by the contingencies of talk-in-interaction (of which the foregoing are but several aspects) and its virtually omnipresent bodily companions – gesture, posture, gaze deployment, facial expression, and so on. Such an expectation is not merely plausible; detailed and repeated examination of recorded episodes of naturally occurring talk-in-interaction shows it to be so – indeed, at a thoroughly implausible (and yet demonstrable) level of detail. The mechanisms of production and comprehension being addressed by P&G need to be understood in this context. It is this context that is missing from P&G’s treatment.

By “context” here I do not mean the ordinary characterizations of settings as domestic or public, intimate or formal, and others drawing on different genres of social and cultural diction (important as these may be). I mean the various organizations of practice that deal with the various generic organizational contingencies of interaction without which it cannot proceed in an orderly way: (1) The “turn-taking” problem: Who should talk next and when should they do so? How does this affect the construction and understanding of the turns themselves? (2) The “sequence-organizational” problem: How are successive turns formed up to be “coherent” with the prior turn (or *some* prior turn), and what is the nature of that coherence? (3) The “trouble” problem: How should one deal with trouble in speaking, hearing, and/or understanding the talk such that the interaction does not freeze in place, that intersubjectivity is maintained or restored, and that the turn and sequence and activity can progress to possible completion? (4) The word selection problem: How do the components that get selected as the elements of a turn get selected, and how does that selection inform and shape the understanding achieved by the turn’s recipients? (5) The overall structural organization problem: How does the overall structural organization of an occasion of interaction get structured, what are those structures, and how does placement in the overall structure inform the construction and understanding of the talk as turns, as sequences, and so on?

The organizations of practice addressed to these issues – turn organization (Goodwin 1979; Schegloff 1996a), turn-taking organization (Jefferson 1986; Sacks et al. 1974; Schegloff 1987a; 2000a; 2001), sequence organization (Schegloff 1990; 1995; forthcoming), the organization of repair (Drew 1997; Jefferson 1974; 1987; Schegloff 1979; 1987b; 1991; 1992; 1997a; 1997b; 2000b; Schegloff et al. 1977), the organization of word selection (Sacks 1972a; 1972b; 1992; Sacks & Schegloff 1979; Schegloff 1972; 1996b), overall structural organization (Schegloff 1986; Schegloff & Sacks 1973), and others – constitute, in the options that they shape and the practices made available, a spate of interaction recognizable as “conversation,” as “interview,” as “meeting,” as “lecturing,” as “giving a speech,” as “interrogation,” and so on. These are what we call “speech-exchange systems” (Sacks et al. 1974,

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

^aDepartment of Cognitive Neuroscience, Faculty of Psychology, Maastricht University, 6200 MD Maastricht, The Netherlands; ^bMax 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?