

# **On The Embodied Nature of Grammar: Embodied Being-in-the-world<sup>1</sup>**

Barbara A. Fox  
Dept of Linguistics  
University of Colorado, Boulder

## **I. Introduction**

In this post-modern age, the issue of representation, especially of mental representation, is being seriously re-examined. By mental representation I mean to refer to a method of describing cognition and action which is rooted in the belief that the basis for all cognition and action is a set of structural units of some kind manipulated by rules. Current scholars of the cognitive sciences (e.g. Dreyfus and Dreyfus, 1986; Dreyfus, 1993) are exploring the possibility that these traditional ways of understanding cognition and behavior may be less satisfying for describing fluent, expert behavior than might other methods of description.<sup>2</sup>

Although speech by a native speaker of a language is one of the clearest examples of expert (or fluent) behavior, linguists have not much taken the opportunity to re-examine assumptions about representation. If there is reason to re-think notions of representation for cognition and behavior in general (as Dreyfus and Dreyfus suggest), then the time may be ripe for discourse-

---

1

I'd like to thank Robert Jaspersen for introducing me to the work of Heidegger and for helping me to pay attention to prosody and gesture. I also want to thank Joan Bybee, Ceci Ford, and Sandy Thompson for insightful comments on earlier drafts of this paper. Thanks also to Chuck and Candy Goodwin for the use of their video and transcript.

<sup>2</sup> . This “non-representationalist” view has taken support from recent studies in connectionism, a computational model of cognition and behavior which has as its fundamental assumption the possibility that knowledge can be described without symbols or structuralist rules. Because of this assumption, connectionism is sometimes referred to as non-symbolic or subsymbolic “presentation.”

functionalists as a community to begin exploring new ways of describing grammar from a non-representationalist perspective.

I take the community of discourse-functionalists to be one appropriate locus for the discussion about representation because this community has a wealth of experience with naturally-occurring language from extremely diverse speech communities, and because this community has from its inception been open to re-examining deeply rooted understandings of grammar (e.g. Givon, 1979; Chafe, 1987; Hopper, 1987; Du Bois, 1987; Ono and Thompson, 1995).

The work of Sandy Thompson is especially noteworthy in this regard. Starting with her early work on the relationships between syntax and semantics (e.g. Hooper and Thompson, 1973) and continuing into her recent work exploring the very nature of syntax (Ono and Thompson, 1995; Thompson and Hopper, 2000), Thompson has challenged existing assumptions of grammatical organization. Her current work serves as an emphatic appeal for more situated and interactionally-sensitive models of grammatical organization:

...syntax...must be understood to be something rather different from what many linguists, even those of a 'functional' persuasion, have thought it to be. We concur with Jespersen (1924) and Langacker (1987: 57) in conceiving of 'grammar'...'dynamically, as a constantly evolving set of cognitive routines that are shaped, maintained, and modified by language use'. (Ono and Thompson, 1995: 215-216)

This paper is offered as one kind of answer to Thompson's appeal and thus locates itself very much within the domain of research initiated by Thompson and her colleagues. It suggests that what appears to be grammatical organization is created by a vast array of transparent and embodied practices which are non-structuralist and non-representational, in our traditional understandings of those terms. The paper takes as a dramatically different departure point the work of Heidegger, as that has been interpreted for a modern and English-speaking audience by Dreyfus (Dreyfus, 1991). The goal of the paper is to provide some evidence and argumentation that such a re-visioning of grammatical organization is worthy of further attention.

In keeping with the theme of the volume, the paper focuses on two utterances which make use of what we can call "clause combining" and discusses some of the practices for achieving combined clauses.

## II. Background

Deeply embedded in Western culture and thinking--reinforced by Greek and then European philosophy, as well as by Christian<sup>3</sup> theology--is the conviction that thinking subjects and the objects they act on and with are entirely separate, the subject acting in the world by means of the mediation of mental representation. That is, the world is "out there," but a representation of it is in the mind of every person, and it is through that mental representation that each person knows the world and can act on the world. Moreover, in the mind-body dualism which is invoked, it is the body which loses, the mind being the seat of all that is positive, good, etc. The body is not even given the status of a necessary evil--just an evil.<sup>4</sup>

For most philosophers, language is the ultimate mental representation. It is no surprise, then, that the overtly acknowledged enterprise of formal linguistics is to provide models of that mental representation. (What can manipulate symbols if not the mind?) It is perhaps also little surprise, that we in functional linguistics also tend, perhaps unconsciously, to accept language as mental representation, living in the mind. Much of the research on grammar in discourse, including my own, takes a view, which, while perhaps leaving unsaid the nature of the representation, nonetheless takes as its starting point the mental/cognitive nature of language and makes no mention of the body.

Perhaps because of this World View, amplified by the history of our tradition --philology, the study of written texts of dead languages--we have settled, quite unconsciously for many of us, I think, on a view of language which takes as primary **the unpronounceable word and its collocational properties**. Whether we view the word and its collocations as mental representations or as an abstract system which lives nowhere (or hovers above a community), we are accepting a view of language as separated from moment-by-moment actually lived human lives, which (like it or not) include bodies in "socially inscribed spaces" (M.H. Goodwin, 1995) and in interactionally achieved sequential locations. Like Chomsky's ideal speaker and hearer,

---

<sup>3</sup> But for a treatment of the complexity of the body in Judaism, see Eilberg-Schwartz (1992).

<sup>4</sup> Candy Goodwin cites Scheper-Hughes in a similar vein:

The body in social anthropology emerges as a passive, inert, dead weight attached to a lively, responsive, nomadic mind, the true agent of culture. (Scheper-Hughes, 1994: 231; cited in Goodwin, 1995: 14)

who never in fact actually "speak" or "hear," the speakers and hearers, or writers and readers, of much work in functional syntax are disembodied minds, out of a social context, using unpronounceable words (what I will call "citation forms") and their collocations (what we call syntax). They do not have hands or feet, tongues or lower backs; they do not sit or drink beer; they do not gaze at others; and they never move. Even work like Lakoff's, which takes as central to the structure of categories the fact that humans have particular kinds of bodies, supposes that the categories that result are cognitive models, residing in "the mind," and the basis of those categories are citation forms of words and their relationships. In this sense, I think most of us, although we might reject the competence-performance distinction, nonetheless have a very "competence"-like view of syntax. Syntax for us is often the relationships between classes of citation-form words.<sup>5</sup>

Critiques of this classical view of human nature can now be found in most disciplines within the humanities and social sciences, instigated in large part by feminist scholarship, but with other sources as well (e.g. Suchman, 1987; Jaggar and Bordo, 1989; Goldenberg, 1990; Gallop, 1988; Grosz, 1994; Jacobus, Fox Keller, and Shuttleworth, 1990; Varela, Thompson, and Rosch, 1991). For example, one early critique of AI focused on the body as the locus of human being, and as at least part of the source of the failure of AI:

We have seen that, as far as we can tell from the only being that can deal with such "vagueness," a "machine" which could use a natural language and recognize complex patterns would have to have a body so that it could be at home in the world. (Dreyfus, 1972; I am using the 1993 edition, p. 304).

But as far as I know, critiques of the tradition of mind have not made a big impact on the practices of functional linguistics. I'd like to suggest here that those critiques are as important to us as to people in other disciplines. To that end, I'd like to explore another understanding of human being and human language, based on the work of Martin Heidegger. In this perspective,

---

<sup>5</sup> Ladd (1980) makes a related point with regard to the relationship in much linguistic scholarship between grammar and intonation:

It is not hard to see how the view of intonation as peripheral might arise. Most studies naturally attempt to ferret out contrasts by taking the segmental part of the sentence as fixed and varying the intonation. But this reasonable heuristic procedure, separating segmental from suprasegmental, is easily transmuted into the idea that the segmental sentence is somehow structurally independent, and that the intonation merely constitutes the way it is said. (p. 119)

the subject-object distinction is discarded in favor of a non-dualistic view of transparent everyday coping:<sup>6</sup>

Self and world belong together in the single entity, Dasein. Self and world are not two entities, like subject and object...but self and world are the basic determination of Dasein itself in the unity of the structure of being-in-the-world. (Heidegger, 1982: 297; cited in Dreyfus, 1991: 67)

[Husserl] developed an account of man as essentially a consciousness with self-contained meanings, which he called intentional content. According to Husserl, this mental content gives intelligibility to everything people encounter. Heidegger countered that there was a more basic form of intentionality than that of a self-sufficient individual subject directed at the world by means of its mental content. At the foundation of Heidegger's new approach is a phenomenology of "mindless" everyday coping skills as the basis of all intelligibility. (Dreyfus, 1991: 3)

Although Heidegger apparently does not say much about bodies, his approach is radically non-mentalistic and non-representationalist:

Heidegger shows that [the] subject/object epistemology presupposes a background of everyday practices into which we are socialized but that we do not represent in our minds. Since he calls this more fundamental way of making sense of things our understanding of being, he claims that he is doing *ontology*, that is, asking about the nature of this understanding of being that we do not *know*--that is not a representation in the mind corresponding to the world--but that we simply *are*. (Dreyfus, 1991: 3; italics in the original)

And in his descriptions of using equipment, it is clear that he means for that being to be crucially bodied (see also Jung, 1987):

---

<sup>6</sup> I am using Dreyfus' commentary on *Being and Time* as my source on Heidegger, along with Heidegger's own *On the Way to Language*.

Where something is put to use, our concern subordinates itself to the "in-order-to" which is constitutive for the equipment we are employing at the time; the less we just stare at the hammer-thing, and the more we **seize hold of it and use it**, the more primordial does our relationship to it become... (Heidegger, 1962: 98; cited in Dreyfus, 1991: 64; emphasis mine)

Heidegger gives us a non-mentalist, non-representationalist, embodied approach to human being, grounded crucially in the union of self and world, that is, in thoroughly "socially inscribed spaces", both of location and time. It is important to note here that by embodied I mean more than "distributed through a body." I really mean physicalized, or en-formed more generally, and in a social world. For example, embodied here refers to the sequential location of an utterance (e.g. "after a question"), the beer can in a speaker's hand, the picnic table at which s/he is seated, the other people present with her/him, the ground which holds the picnic table up, the air s/he is breathing in order to speak, the recipient's lateral head shake, and so on. It is not just the body that is important; and it is not just that the body is important, but that the body and all forms live in an already socially inscribed world. I am thus not proposing that we extend the limits of the self from the skull to the skin, but that we take as our starting point embodied-being-in-the-world.<sup>7</sup>

With this introduction, let us turn to some language data.

### **III. Examples of Face-to-Face Interaction**

In this section I'd like to explore some empirical evidence for the proposal that what we think of as grammatical organization is created by embodied practices in socially inscribed environments. Section IV will return to Heidegger's claims regarding transparency of action.

This section explores the embodied practices deployed in two concerns of relevance to participants in interaction: (a) turn-projection--that is, when recipients (and us as overhearing analysts) hear an utterance as possibly complete, when it would be relevant for someone else to start to speak; and (b) action interpretation—that is, how and when an utterance is heard to be

---

<sup>7</sup> Hinduism and Buddhism offer similar insights into the nature of human being. For a discussion on bringing Buddhist insights into modern cognitive science, see Varela, Thompson, and Rosch (1991).

doing a particular kind of social action (such as offering, complaining, etc.). Participants are always attuned to these two concerns, as they must know what action is being performed through a particular utterance so that they can respond with an appropriate next action; and they must know when an utterance is possibly complete so that they can respond at the appropriate moment. We can thus say that these concerns are omnirelevant.<sup>8</sup>

Let us begin with a consideration of turn projection. Within the Conversation Analysis literature, the directedness of an utterance towards completion is referred to as projection; that is, the shape of an emerging utterance projects how and when that utterance might come to possible completion. I want to focus on projection here because it is one of the major functions served by syntax in conversation and hence should be of interest to functional linguists (see Sacks, Schegloff and Jefferson, 1974; Ford and Thompson, 1996; Ford, Fox and Thompson, 1996). Turn projection is also a rich site for examining the complex relationships among what we refer to as syntax, pragmatics, prosody and gesture.

To begin, let us examine an utterance from a casual social interaction (in American English) which illustrate the embodied nature of language. This utterance is taken from a longer segment of the interaction (which occurred during a backyard picnic in central Ohio in the early 70s), which is given below. The utterance to be examined is given in bold face and is arrowed.<sup>9</sup>

The conversation involves 3 heterosexual couples. Pam and Curt are the hosts of this picnic; Mike and Phyllis are friends of theirs, and Carney and Gary are family to Curt (but not known to Mike and Phyllis). The segment starts with all 6 participants seated at a picnic table: Phyllis, Mike and Carney are seated on one side; Gary, Curt, and Pam are seated on the other. Shortly after the segment starts, Pam gets up and walks away from the table.

The segment begins with Pam suggesting that someone tell a joke:

Pam: .hh **Oh** yeah you've gotta tell Mike that. Uh-cuz they want that on fi:lm.

[

---

<sup>8</sup> They are also obviously entirely intertwined. Part of knowing when an utterance may be possibly complete is knowing when the action implemented by that utterance could be complete (see Ford and Thompson, 1996; Ford, Fox and Thompson, 1996).

<sup>9</sup> The video (with audio) of this segment can be viewed on the web at:

Car: Oh: no: here we go  
 ag(h)(h)ain o(h)o(h)o .hh=  
 [  
 Cur: Huh huh huh huh  
 Gar: =I don't think it's that funny.  
 [  
 Car: Oh:.,  
 [  
 Pam: I gotta go to the john before I hear that again.  
 [  
 Car: You'll like it, you'll really like it.  
 [  
 Cur: You do too you laugh like hell you huh!  
 [  
 Phy: ehheh huh  
 Gar: Well I.,  
 [  
 Cur: Y-  
 Gar: hadn't had a beer ye:t.=  
 [  
 Pam: **You don't like it because you didn't think of it!** <==  
 [  
 Gar: eh-heh-heh-huh-hah-huh!  
 [  
 Cur: eh!  
 Cur: =ehh-heh that's ri(h)gh(h)t nnn .hh  
 [  
 Phy: hehhhhuhh  
 Gar: I: hadn't had a beer ye:t.I: can laugh at anything get a bee:r,  
 [  
 Cur: nh huh huh. huh, huh-huh  
 [  
 Phy: ehhu::n  
 Cur: eh-heh  
 Gar: heh-heh-heh-heh-heh-ha-ha-ha-ha-ah!ah!ah!  
 [  
 ]



Cur: That's ri(h):ght. (huh!),  
 [

Mik: hah:hah:hah: hah huh huh, huh huh hah huh  
 [

Phy: hnnn n-hn-hn  
 [

Cur: (h)You want another beer, you  
 better (keep laughing)

In this interaction, Pam has suggested that someone tell a joke, perhaps for the benefit of the camera, that the others have already heard. There are four reactions to this suggestion, including Gary's *I don't think it's that funny*. Different participants then orient to Gary's reaction: (1) Carney's *You'll like it, you'll really like it*, (2) Curt's *You do too you laugh like hell you hhuh*, and (3) Pam's *You don't like it because you didn't think of it!* It is Pam's response to Gary that I want to explore further:

(1) Pam: You don't like it because you didn't think of it!

A phonetic transcription of the utterance might look like this:

\*\*

Although it was not possible to create a pitch track for this utterance (because it is produced in overlap), the general pitch movement can be described visually as follows:

You don't like it because you didn't think of it!

That is, the pitch starts high on *You*, comes down significantly on *don't* and continues to fall through *because*; there is a pitch peak on *you* and the pitch then is much lower on *didn't*. The pitch goes up again for the accent on *think*, and then is lower on *of*, with a very slight increase on *it*.

From the perspective of context-independent syntax, there are variety of places at which Pam's word string could be complete:

You.

You don't.

You don't like it.

You don't like it because you didn't.

You don't like it because you didn't think.

You don't like it because you didn't think of it.

But given the sequential location--Pam is responding to Gary's *I don't think it's that funny*, and after Carney and Curt have also responded--only a few of these are possibly relevant contributions, and each of these would have to be said with a particular intonation and rhythm to be recognizable as doing that utterance, and hence as upcomingly complete. For example, for *You* to be a hearably complete utterance in this sequential location, it would probably have to show a complete pitch contour, either terminal rising or falling. The prosody of the utterance is thus crucial, so I will now try to analyze the utterance from that perspective.

*You* is heavily accented here, through volume and pitch. The pitch comes down on *don't* and there is no accent on *like*, indicating that after the direct object may not be a place of possible completion. Furthermore, *it* is produced at the same pitch as *like*, so there is no suggestion of completion-relevant fall or rise. The pitch on *because*, at least on the syllable *-cause*, is lower than the pitch on *it*. The accent on *think* indicates a possible upcoming completion after the object.

The accent on the first *You* is interesting. While this accent could be heard as creating a contrast (*You* as opposed to others present), it is also possible that the speaker accents *You* in order to avoid accenting *don't like*; that is, *don't like* is deaccented and the accent is done on *You* as the default locus of accent (following Ladd, 1980). According to Ladd (1980), this deaccenting leads us to hear the utterance as a shifter (in the Jakobsonian sense), as interpretable only with regard to the context:

...the shifter-like quality, the interpretability only with reference to something else in the context, is clearly signalled by the lack of accent. (p. 93)

And, indeed, we are led to hear *You don't like it* as a rephrasing of Gary's earlier utterance (*I don't think it's that funny*). It is at least in part through the pattern of relative weakening and strengthening of prominence, then, that we hear Pam's utterance as directly responsive to Gary's.

The important facet of this for our discussion here is that Pam's utterance could be heard as beginning with already known, or presupposed, information, and this relationship of information is now widely considered to be a grammatical relationship. Now in this example we can see that it is largely what we think of as the prosody which carries the "presupposedness" of the first segment of Pam's utterance. To me this fact suggests not only that prosody can carry grammatical information (a proposition that is already well accepted), but that in fact prosody is grammar.

The rhythm is also interesting: ' .....!'. The whole utterances makes use of what Bolinger calls descending rhythm, which means that the accent is on the first syllable of each foot. Almost all of the words are single-syllables (with the exception of *because* and *didn't*, the last of which is marginally two syllables). The last two feet are dactyls.

It should be clear from these few observations that the prosody is a crucial thread in creating what Pam is up to with this utterance. It is not that the citation forms of words, which we think of as constituting syntax, are said and then prosody is added on top of that; prosody, interpenetrated by and with segmental articulation, is how we hear these particular words and the possibilities they project. And let us be reminded here of what prosody is. As I understand it, prosody is the deployment of vocal vibrations and their absence in a particular temporal manner, which involves most of the torso and head: the diaphragm, abdominal muscles, lungs, throat muscles, vocal folds, back (as resonating cavity), tongue, teeth, etc. Prosody is thus obviously bodied.

I have offered this example to illustrate some of the embodied nature of the practices involved in turn projection, especially some of the practices using the voice. Let us now turn to consider an illustration of action projection, this time focusing on the body beyond the voice.

In the utterance given as (1) above, Pam has (perhaps jokingly) insulted Gary. Gary responds to this with an account: *I hadn't had a bee ye:t.I can laugh at anything get a bee:r,. What Gary appears to mean by this is that he didn't think the joke was funny when it was told earlier because he hadn't yet had a beer; once he's had a beer, he'll be able to laugh at anything (including the same joke, if it is told again). Curt then responds as given below:*

(2)

Cur: That's ri(h):ght. (huh!),

[

Mik: hah:hah:hah: hah huh huh, huh huh hah huh

[

Phy: hnnn n-hn-hn

[

Cur: **(h)You want another beer, you <==**

**better (keep laughing)** <==

Curt begins his response with *That's ri(h):ght. (huh!)*, a joking acknowledgment of Gary's account. He then goes on to display his understanding of Gary's utterance (an understanding which he had only claimed with *That's ri(h):ght. (huh!)*): *(h)You want another beer, you better (keep laughing)*.

How should we hear Curt's utterance in real time? Is *You want another beer* heard as a possible offer of a beer at this moment? There are cans of beer on the picnic table, and as Curt is the host of the picnic, it would be relevant for him to offer the guests something to drink. It is thus relevant to ask if Curt's *(h) You want another beer* is hearable at any point in its production as an offer. I will claim that it is not hearable in this way; and I will suggest that it is the embodied production of the utterance—which would be so obviously understandable as an offer out of context—that makes it something other than an offer in this particular context (in fact it comes off as a joking threat). The argument offered for this analysis is meant as evidence in favor of the proposal that “grammar” is created by a vast array of embodied practices.

First, we can notice that *(h)You want another beer* is done as a continuation from *That's ri(h):ght. (huh!)* The first *You* is done without a pitch reset or new head position, so even from the very beginning we don't hear *(h)You want another beer* as a possible question, that is, as a starting of something new. We also hear the utterance as not possibly complete after *beer*, by the coming together of the following four facets of Curt's comportment:

- the tempo of the talk (*beer* is not noticeably lengthened);
- the pitch rise on the last few syllables;
- the continued gaze towards Gary;
- and the head movements done through the talk.

Since I have already discussed prosody, I want to focus now on Curt's head movements. Near the beginning of this utterance, Curt appears to straighten his neck, or at least to bring his chin in and down. There is then a sharp vertical head nod, which, together with the upward intonation, strongly indicates more to come, some talk through which another head nod will be done. That is, the first head nod is done in such a way that projects another head nod to come (and in fact there is a third head nod, post completion). It is thus at least in part through Curt's head movements (in addition to prosody) that we hear (*h*)*You wan' another beer* not as a question/offer, but as the first component of a complex utterance, probably hearable as a conditional/threat.

The rhythm of the utterance also may lend some clues as to why Curt uttered the particular words he did. In contrast to Pam's descending rhythm, Curt's utterance begins with a clearly ascending rhythm, which means that the accent is at the end of each foot (in this case, on *beer*). This may lead to a resonance of ascending rhythm in the second segment of the utterance (with accent on *laughing*). And it is possible that Curt utters the words he does in each segment at least in part to create this resonance, so that, for example, the last word (if not the last syllable) of each segment is accentable. Compare example (1), in which Pam uses forms that allow her to end the segments of her talk with unaccentable syllables (*it, of it*). If this is true (as Bolinger, 1986, suggests), then words are perhaps as much sound-possibilities as meaning-possibilities (and perhaps for native speakers those are not separate).

Notice that while this utterance is what we might call a conditional sentence, of the *if-then* sort, neither the word *if* nor the word *then* is actually uttered. It is the embodied doing of the utterance, and not just some isolatable words, which allow a native speaker to hear that a conditional is being produced. If head movements can carry at least part of a meaning like "conditional," which we think of as grammatical (as opposed, perhaps, to emotional), then do we not have to include head movements in meaning-representations? As well as prosody, gesture, gaze, body position, and so on? But then all of the body has become part of representation, and it is no longer "mental."

#### **IV. Discussion**

From the discussion above, we can see that language is embodied in at least one quite obvious sense. Two of the important roles played by syntax (however we understand that term) in conversation is the projection of what will constitute a place of possible completion (Ford and Thompson, 1996; Ford, Fox and Thompson, 1996), and the interpretation of what action is under way (Schegloff, 1995). But even if we want to maintain a fairly decontextualized, abstract notion of syntax, we must still accept that part of what participants orient to in these concerns comes from prosody, gesture, gaze, body movements, and other entirely physicalized, embodied practices. Although this is an obvious point, it is one I think we very often forget in how we actually do functional linguistics.

But there is a deeper sense in which language is embodied that I would like to explore here, and that is, that there may be, for native-speaker participants speaking in a transparent way in the setting of everyday talk, no decontextualized, disembodied, abstract, or purely mental form of words and syntax. Rather words, and their possibilities for collocation, can be seen as embodied practices, which native speakers, by virtue of having vast experience speaking in everyday social interactions, do.

Now one could argue that words are abstract mental objects, as is syntax, which create a proposition (which is a decontextualized mental object), and that prosody, gesture, gaze, body movements, etc. are (obviously) embodied parts of the performance of an utterance, parts which display the emotional state of the speaker, or the turn-taking possibilities of the utterance, or some other facet of speech which is icing on the propositional cake. For example, we could say that syntax sets up one set of expectations on how the utterance will proceed, and prosody, gesture, gaze, body movement, etc. each contribute their own, separate, projections; the recipient must calculate how these different projections are to be oriented to.

But if we go back to the examples as they unfold, we see that it is difficult to maintain this position. It might be possible to hold on to it if an utterance could start with a citation form, but every utterance is from its very beginning voiced, and often preceded by gesture; it occurs in an already physicalized setting (such as in chairs, or at a picnic table), and in a sequential context which is made manifest by prior vocalizations, gestures, movements, and in traces of current body postures, participant alignments, and so on. The gesture which precedes the utterance, if

there is one, may set up projections even before any "words" have been uttered, and the sequential environment of the utterance also sets up projections. So every utterance "begins" (in quotes because it is in fact impossible to say where an utterance starts) already in a bed of projections as to what might constitute a possibly complete utterance and as to what kind of action might be relevant. There is thus no "fat moment," as Garfinkel would say, no moment at which "syntax" alone might show itself.

To make this point concretely, let's go back to Pam's utterance. Gary has indicated that he might not want to hear the joke. There have been two responses to Gary's comment, one a cajoling and the other a disagreement. Pam's utterance thus starts in the context of counter-responses to Gary. Pam's utterance begins with the deictic pronoun *You*, one of a small class of items believed by even the most serious mentalists to require context for its interpretation. To find the referent of *you* in any utterance, one must know which body is speaking, whom that person's gaze finds, whether the voice quality is of quotation or not, and so on. This case is no exception. At this point in the interaction Pam has walked away from the picnic table and is now behind Gary, at what distance we don't know. Even though we cannot see her gaze or body orientation (and neither can two of the participants, at least when she starts), it is clear by the quite loud volume of *You* that she means to locate a recipient that is not physically close to her. And since the recent referent of *you* has been Gary, and since Gary is both physically distant and turned away from Pam, it would be reasonable to hear this *You*, at least tentatively, as referring to Gary. The high pitch and voice quality of *You* suggest disagreement and disapproval, so we might hear the utterance, from its very inception, as disagreement (or negative in some way) directed at Gary. And we can hear from the lack of pitch rise or fall that the utterance is not possibly complete after *You*. All of this we can hear from the embodied production of a pronoun which itself makes few projections (and could, in some contexts, be complete by itself). We next hear *don't*, which by virtue of a lack of pitch rise or fall on it, as well as by its reduced and unaccented form, indicates that the utterance is not possibly complete; we can thus hear that a verb may be coming. Here we can see that the projection based on words and their collocations originates in how those words are said: we only hear that a verb may be coming because we have heard that the utterance is not possibly complete at *don't*. The saying of a word and what it projects are absolutely intertwined.

There are several ways of arguing against this suggestion.

(1) One could suggest as a counter hypothesis that we have a connectionist-like network of disembodied words, and once one word is activated, all of the possible next words receive some

activation; one word or type of word "wins" by being further activated by prosody, gesture, body movement and so on. (This is similar to the word activation model proposed by cognitive psychologists like Swinney and Kintsch.) My response is this: First, such a model is meant to be a model of the brain and the brain is part of the body. I believe it is misleading to focus on the brain as the only part of the body that "does language," but I otherwise have no objection to modeling the brain. Second, if prosody, gesture, etc. can spread activation to a word or type of word, then somehow such obvious embodiments are manifest in the brain and are inextricably connected, as it were, to words, and to their meanings. Any separation between word and body, even in this model, thus seems entirely artificial. (And as Bolinger has pointed out "in intonation there is no distinction between the grammatical and the ideophonic except as they represent extremes of a scale" (Bolinger, 1986: 32). It is for this reason that some linguists have started to include intonation as part of syntax, but even in these cases intonation is seen as added on to the "word-things.") Third, I expect that in such a model words would have to have a phonological representation as part of their "abstract" representation, and since as least as far as I can see phonology is in essence about the body (even an abstract feature like [+nasal] would not be used of the sound made by wind in the trees), then words cannot escape being embodied.<sup>10</sup>

(2) A more traditional counter-hypothesis would say that what syntax projects in any particular utterance is irrelevant to syntax as an abstract system. In the case of Pam's utterance, for example, native speakers of English know that *do* can be both a form of verbal anaphora and an auxiliary; which one is intended on any particular occasion may be indicated by prosody, etc., but that has nothing to do with the abstract knowledge native speakers have. Prosody, gesture, and so on may act to signal to the hearer which syntactic structure is intended, but otherwise they have no role in syntax. This seems like a completely plausible view, until one tries to integrate it with the fact of speaking (which is why many scholars avoid "performance" altogether). If one attempts this integration, then several problems arise:

(a) It doesn't address the fact that for some obvious cases, prosody is part of what makes something a particular structure (declarative interrogatives are clear cases).

(b) It doesn't say how prosody, gesture, etc. do the work of disambiguation. If the syntactic structure is abstract, how is this disambiguation accomplished?

---

<sup>10</sup> For a critique of connectionism from a Heideggerian perspective, see Dreyfus (1993).



(c) It seems to me that if such an approach were used to "generate" utterances in a conversation, it would have to take as input to the syntactic structures something like prior meanings or intentions--that is, the speaker's intended meaning would "drive" what syntactic structure was selected from the abstract system. I'm not sure how else the speaker could select the right syntactic structure from all the possibilities.

I would like to explore this problem in more detail. The first difficulty with this approach is that, as Schegloff (1987) has pointed out, people do not start with an intended meaning: they start with the sequential location they find themselves in. For example, if someone has just asked you a question, what you say next will be heard as either answering or not answering that question--this has nothing to do with your personal intentions. Answering is relevant regardless of your intentions.

So let's say someone has invited you to go out for lunch. The meaning you want to convey is to accept, and you want to do so enthusiastically. Let's say you select some kind of syntactic structure, put in some citation-form words. But how do you know which structure to choose, and which words to choose? On the basis of the intended meaning, of course. But how is that meaning given to you? If it is given in words, then the whole model is empty--just speak the words which are given to you in the meaning. The model fails to answer where those words in the meaning might have come from. That is why scholars in certain generative traditions have written meanings with capital letters (to disguise their already English-word nature), or have searched for essential components of meaning, or have proposed cognitive models of the world which underlie the words. Because this is a horrific problem. We really have no idea how such prior-meanings could be represented without presupposing exactly what they are supposed to generate. Maybe this problem is solvable; I can't prove that it is not. But in 2000 years of theorizing, no one has yet solved this problem.

Moreover, it is clear that speakers use words and what might be considered syntactic structures for reasons that are clearly not abstract-semantic. For example, Sacks (1991: 735, volume 1) points out that in utterances like *Yes I did too*, the *too* is there not for abstract-semantic reasons but for prosodic reasons--*too* is needed as a place to put the accent when *I did* and *I did* are sequentially not appropriate. Or, as I suggested above in the case of example (2), it may be that certain words, and syntactic structures which might allow those words, are used to achieve rhythmic patterning. So here again we see it could be that words, and all of the ways of putting words together, open fields of embodied possibilities (to steal a phrase from Heritage, 1984).

And here is where Heidegger can help us. For Heidegger, no prior meaning or intention is necessary, and this is possible because speaking is an embodied skill which people know how to do. Our bodies know precisely how to answer a question delicately, how to decline an invitation so as to indicate we do not wish to be invited again. There is no need to locate a prior, propositional, meaning in the mind:<sup>11</sup>

One can say "Give me a lighter hammer," and receive one, without having either an intentional state (a desire for a lighter hammer) or the intention to communicate that state by using words. Indeed, this is the way language normally works. (Dreyfus, 1991:220)

The traditional view of practice, from Descartes on at least, is representational. Contemporary philosophers such as John Searle and Donald Davidson, who do not agree on much, do agree that action must be explained in terms of beliefs and desires, that is, mental states causing bodily movements. Heidegger's attempt to break out of the tradition is focused in his attempt to get beyond the subject/object distinction in all domains, including action....The focal problem is thus not which kind of intentionality--theoretical or practical--is more basic, but how to get beyond the traditional account of intentionality altogether. (Dreyfus, 1991: 49-50)

What is imposed on us to do is not determined by us as someone standing outside the situation simply looking on at it; what occurs and is imposed are rather prescribed by the situation and its own structure; and we do more and greater justice to it the more we let ourselves be guided by it, i.e., the less reserved we are in immersing ourselves in it and subordinating ourselves to it. We find ourselves in a situation and are interwoven with it, encompassed by it, indeed just "absorbed" into it. (Gurwitsch, 1979: 67; cited in Dreyfus, 1991: 67)

---

<sup>11</sup> Sacks seems to hold a similar view when he discusses anthropomorphizing humans:

I'll make a principled statement, which is quixotic enough but I believe in it. I figure it's okay to anthropomorphize humans. I don't think it's any worse for sociology to anthropomorphize than, say, for physics to do it. No better, but no worse. All research anthropomorphizes its objects. This is not to say that I believe humans are anthropomorphic. I certainly don't. (Sacks, Winter 1967, March 9, page 3; cited in Jefferson, 1989: 429)

Dasein ... is nothing but ... concerned absorption in the world. (Heidegger, 1985: 197; cited in Dreyfus, 1991: 67)

In order for language to be able to exist in this way, it must always be embedded in social contexts which are meaningful, and meaningful as resting in a vast experience of unspoken and unspeakable assumptions and activities (Garfinkel made this point too; see Garfinkel, 1967):

Language is used in a shared context that is already meaningful, and it gets its meaning by fitting into and contributing to a meaningful whole. (Dreyfus, 1991: 219)

...all interpreting takes place on a background understanding that it presupposes--a background, moreover, that conditions from the start what questions can be formulated and what counts as a satisfactory interpretation, yet that can never be made completely explicit and called into question...(Dreyfus, 1991: 200)

## **V. Implications**

What does this line of argumentation lead us to say about grammatical organization? According to Heidegger, in an everyday interaction, participants are not "using grammar," or "saying words," or "making phonemes," or any such linguistic-based notion; rather, they are telling jokes, insulting each other, making arrangements to get together again, comforting each other, and so on. They are speaking in a way that is transparent to them, and to any native-speaker analyst watching them. So it is very easy as a native-speaker analyst to hear what they heard, and what they heard was not passive constructions or the like. Heidegger refers to this transparentness of language (or any "equipment") as *availableness (Zuhandenheit)*. Availableness is experienced as entirely embodied, contextualized. For example, if you are going out a door, the doorknob is transparently available.

This is not to say that what we think of as grammatical organization never comes into existence. When linguists engage in transcription and grammatical analysis—that is, problematizing what

was said—grammatical patterns do seem to emerge. And it is very likely that for participants in an interaction there are moments where a kind of linguistic analysis may take place. Heidegger refers to this less transparent, more decontextualized mode, as *occurrentness* (this is Dreyfus' translation of *Vorhandenheit*), and he suggests that as beings-in-the-world, we engage in practices of *occurrentness* when we run into trouble, that is, when the transparency of coping is interrupted. When we engage in practices of *occurrentness*, other beings may arise as "objects" (but not in the sense of self-sufficient objects that we typically think of in the subject-object distinction). For example, if I am trying to go out a door and the door is stuck, the doorknob may become *occurrent* to me as I try to solve the problem of the door-stuckness. "Grammar," like other such "objects," can be the result of such practices of *occurrentness*. "Grammar" is occurrent, while transparent speaking is available. To the extent that our notion of "grammar" is disembodied, so we will find it through practices of *occurrentness* (like linguistic analysis, transcribing, etc).

Now it seems to me that this point may have important ramifications for how we practice functional linguistics. Transparent speaking, as available, is primary; "grammar," as *occurrent*, is derived from transparent speaking by decontextualizing manipulations of various sorts. It thus may be that "grammar" is not what participants are up to when they speak, unless the transparency of the interaction is interrupted and they engage in making language *occurrent* (for example, through repair, or comments on their "bad grammar," or the like). What, then, are we studying when we study "grammar"? If we are trying to understand how people speak, then is studying "grammar" the way to do that? How else can we understand transparent speaking? I don't have the answers to these questions--I just want to suggest that they are important questions to address.

One objection that could be raised against what I have been exploring is that I am talking about speech and not grammar. But this is just our old friend, the competence-performance distinction, again; and I don't believe that it is possible to understand language-in-use (and what other kind of language is there?) through this distinction. As I tried to show above, whenever this abstract notion of language is held responsible for the actual speakings of people, it runs into trouble. The distinction needs to be dissolved.

So how should we proceed? How should we approach transparent speaking? At this point, I'd say I'm not completely sure. So far, for myself, I find the methods of Conversation Analysis, with its insistence on displaying the relevance of the analysis for the participants, the most helpful. But beyond this, I think we can only proceed with caution and attention to the nature of

the categories we work with. And with a deep respect for the socially inscribed bodies and forms in and through which we dwell.

## REFERENCES

- Bolinger, D. 1986. *Intonation and Its Parts*. Stanford, CA: Stanford University Press.
- Chafe, W. 1987. Cognitive constraints on information flow. In R. Tomlin (Ed) *Coherence and Grounding in Discourse*, pp. 21-51. Amsterdam: John Benjamins.
- Dreyfus, H. 1991. *Being-In-The-World*. Cambridge, MA: MIT Press.
- Dreyfus, H. 1993. [1972]. *What Computers Still Can't Do*. Cambridge, MA: MIT Press.
- Dreyfus, H. and S. Dreyfus. 1986. *Mind over Machine*. New York: The Free Press.
- Du Bois, J. 1987. The discourse basis of ergativity. *Language* 63: 805-55.
- Eilberg-Schwartz, H. (Ed.). 1992. *People of the Body: Jews and Judaism from an Embodied Perspective*. Albany, NY: SUNY Press.
- Ford, C. and S. Thompson. 1996. Interactional units in conversation: syntactic, intonational, and pragmatic resources for the projection of turn completion. In E. Ochs, E. Schegloff and S. Thompson (Eds.) *Grammar and Interaction*, pp134 –184. Cambridge: Cambridge University Press.
- Ford, C., Fox, B. and S. Thompson. 1996. Practices in the construction of turns: the 'TCU' revisited. In C. Ford and J. Wagner (Eds), *Pragmatics*: \*\*
- Gallop, J. 1988. *Thinking Through the Body*. New York, NY: Columbia University Press.
- Garfinkel, H. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Givon, T. 1979. *On Understanding Grammar*. New York: Academic Press.
- Goldenberg, N. 1990. *Returning Words to Flesh: Feminism, Psychoanalysis, and the Resurrection of the Body*. Boston, MA: Beacon Press.

- Goodwin, M.H. 1995. Games of stance: conflict and footing in hopscotch. Paper presented at the Symposium on Conversation, Albuquerque, NM.
- Grosz, E.A. 1994. *Volatile Bodies: Toward a Corporeal Feminism*.
- Heritage, J. 1984. *Garfinkel and Ethnomethodology*. London: Polity Press.
- Hofstadter, A. 1965. *Truth and Art*. New York, NY: Columbia University Press.
- Hooper, J. and S. Thompson. 1973.
- Hopper, P. 1987. Emergent grammar. Proceedings of the thirteenth annual meeting, Berkeley Linguistics Society, pp. 139-157. Berkeley: Berkeley Linguistics Society.
- Jacobus, M., Fox Keller, E., and S. Shuttleworth. 1990. *Body Politics: Women and the Discourse of Science*. New York, NY: Routledge.
- Jaggar, A. and S. Bordo (Eds.). 1989. *Gender/Body/Knowledge: Feminist Reconstructions of Being and Knowing*. New Brunswick, NJ: Rutgers University Press.
- Jung, H. Y. 1987. Heidegger's way with Sinitic thinking. In G. Parkes (Ed.) *Heidegger and Asian Thought*. Honolulu, HI: University of Hawai'i Press.
- Ladd, D. R., Jr. 1980. *The Structure of Intonational Meaning*. Bloomington, Indiana: Indiana University Press.
- Ono, T. and S. Thompson. 1995. What can conversation tell us about syntax? In P. David (Ed) *Alternative Linguistics: Descriptive and Theoretical Modes*, pp. 213-271. Amsterdam: John Benjamins.
- Sacks, H. 1991. *Lectures on Conversation*. Cambridge, MA: Blackwell.
- Sacks, H., Schegloff, E., and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696-735.
- Schegloff, E. 1987. Talk presented at Xerox PARC.

Schegloff, E. 1995. Discourse as an interactional achievement III: the omnirelevance of action.  
*Research on Language and Social Interaction* 28: 185-211.

Scheper-Hughes, N. 1994. Embodied knowledge: thinking with the body in critical  
medical anthropology. In R. Borofsky (Ed.) *Assessing Cultural Anthropology*.  
New York: McGraw-Hill.

Suchman, L. 1987. *Plans and Situated Actions*. Cambridge: Cambridge University Press.

Varela, F., Thompson, E., and E. Rosch. 1991. *The Embodied Mind*. Cambridge, MA:  
MIT Press.