13 Contextures of Action

Charles Goodwin

This chapter will argue that human action is 1) constructed through the systematic use of public signs in diverse media with quite different properties that mutually elaborate each other (and is thus intrinsically embodied, multi modal, and tied to the world where action is occurring); 2) cooperative; and 3) distributed. Most of the work will focus on how Chil, an aphasic man with a three-word vocabulary, is able to function as a powerful speaker by building action in concert with others.

Chil's ability to do this calls into question many traditional assumptions about what constitutes a speaker, the range and types of phenomena that must be taken into account to describe how action is built within states of talk, and the analytical integrity of units such as the isolated utterance or sentence. The practices used to build his action will offer support for the argument that both the resources humans use to construct meaning – including phenomena such as language – and human agency are intrinsically cooperative, public, social, and distributed. Both action and the sign complexes used to accomplish action emerge within interactively sustained contextual configurations (C. Goodwin, 2000) that link both diverse signs and differently positioned participants into common courses of action, within a continuous process of progressive transformation.

Concrete particulars can help make clearer the nature of the arguments being made in the two preceding paragraphs. Therefore, before turning to the sequences with Chil that will be the primary focus of this chapter, it is useful to briefly illustrate the view of action being proposed here with a specific example.

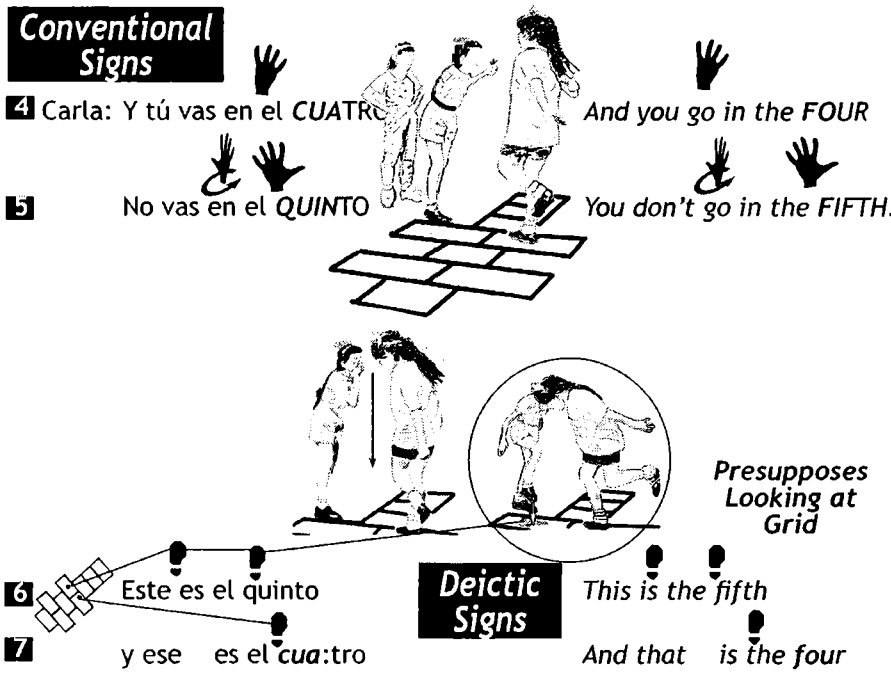
In Transcript 13.1 Carla, on the left, is accusing Diana, on the right, of making an illegal move in hopscotch, specifically, throwing a bean bag on the wrong square (the placement of the bean bag constrains what squares must be jumped on). This example is taken from Goodwin (2000), where analysis of the practices used to build action being reported here is developed in detail.

After Diana throws the bag, and while Diana is starting to hop through the squares, Carla marches onto the grid, stops Diana's progress by pushing her, and accuses her of being a cheater ("Chiriona"). Then, in the sequence in Transcript 13.1, Carla tells Diana she is supposed to go in the fourth square (line 4), not the fifth (line 5). Carla says this while leaning toward Diana and thrusting her hand toward Diana's face. The numbers in her talk are matched by hand shapes displaying the same numbers (four fingers over line 4 and five over line 5). Lines 4 and 5 use the same syntactic frame and are spoken with matching prosodic contours that highlight the difference between the relevant numbers. During most of these lines, Diana is standing on one foot (perhaps displaying that she is still treating as viable her status as a player making a legal move and entitled to jump through the grid) while gazing at Carla. She is thus positioned to not only hear what Carla is saying, and thus visibly constituting herself as a hearer to Carla's talk, but also to see her gestures. Note that neither the structure of Carla's talk with its use of numbers to describe grid squares, nor her use of her fingers to construct conventional counting gestures, presuppose that their addressee is looking toward the actual squares being talked about.

At the end of line 5, over "QUINTO," Diana's head drops away from Carla, so that her eyes are now pointing toward the grid at their feet. Diana is thus no longer visibly positioned as a hearer to Carla, or able to see her hand gestures.

As soon as this happens, Carla moves her own gaze to the grid while continuing her accusation, only now with deictic signs: The term "Este" ("this") in her talk is linked to a pointing gesture, a foot stomp on the grid square being talked about, which is placed, like Carla's earlier hand gestures, right where Diana is now gazing.

Carla's ability to adapt her utterance in mid-course to what her addressee is visibly doing demonstrates a particular kind of semiotic agency. As an actor and speaker, Carla possesses a repertoire of signs with different properties (for example, both names and conventional counting gestures that do not require actually looking at what is being talked about, versus deictic signs) that allow her to rapidly adapt to the changing environments that shape the actions being pursued through her talk



Transcript 13.1.

(the presence of a hearer as an active co-participant is a central feature of such environments).

What occurs here provides one concrete demonstration of the general arguments about the organization of action made above. First, action is being constructed through the integrated use of signs with diverse properties lodged within a variety of different media. The discrete systematic oppositional contrasts, in both sound structure and lexicon, used to build Carla's language are accompanied within the stream of speech by the continuously varying flow of her prosody. The mutual orientation of Carla and Diana, as displayed through their linked postural displays and gaze, creates a framework where other kinds of sign exchange processes, such as Carla's talk and gesture, can flourish. These postural displays ground and make intelligible the deictic signs produced through both talk and gesture by establishing public frameworks of shared orientation toward specific phenomena in the local scene, including in this sequence not only specific locations on the grid, but also each other's bodies. The hopscotch grid itself is a semiotic structure, one that makes the game itself possible. Unlike talk, it is constructed within a material medium that allows participants to build action by actually jumping through the ordered spaces it creates.

To summarize, action is being built here through the integrated use of structurally different kinds of sign phenomena, lodged within diverse media with quite different properties that mutually elaborate each other (for example, signs within the stream of speech that rapidly disappear versus the enduring presence of the painted grid that is able to support the weight of multiple bodies).

Second, the organization of action and sign use that occurs here is cooperative. This does not mean that actors are seeking solidarity and harmony. Indeed here, they are not only in an argument where each is refusing to acknowledge the position being taken by the other (note Carla's disruption of Diana's move by walking into her game space, and Diana's posture of readiness for continued play), but with the dropping of her gaze, Diana is no longer even acting as a recipient to the actions Carla is directing toward her. Instead, cooperation refers to the way in which subsequent (as well as simultaneous) action is built by performing systematic operations on the sign complexes made publicly available by others. Thus Diana's gaze drop is meaningful because it removes her face from displaying that she is acting as a hearer to the action Diana is addressing to her – a public position made possible and visible by Carla's gaze toward Diana, and Carla's deictic stomp is organized with reference to the new sign complex and public focus of attention created by Diana's postural shift.

From a slightly different perspective, the way in which Carla restructures her emerging action when Diana looks away demonstrates that neither her utterance nor her turn at talk is organized entirely within the stream of speech, or with reference to her actions alone. Instead, in face-to-face interaction, both utterances and turns are multi-party activities (see Iwasaki, this volume) in which the hearer, even when silent, is a consequential, active participant.

Because consequential displays of the hearer (as well as the speaker) are made not with talk, but with visible embodied signs, the turn – and the utterance emerging within it – have a crucial multi modal organization. They

are constructed through the interplay of structurally different kinds of signs produced by actors occupying differentiated positions within an unfolding interactive field. For example, in mundane conversation, the principal characters in stories are frequently present at their telling (a wife may tell a story about something her husband said while he is present), and this can make relevant particular forms of silent, though visible, participation in the story as crucial events within it are told (for example, the present party's stance toward what he or she is described as doing within the story). The principal character works to arrange his or her body for the forms of participation and stance displays that become relevant at specific points within the telling. As this is done, his or her body displays public analysis of the projections and possibilities for action provided by the unfolding syntactic structure of the teller's talk (C. Goodwin, 1984). The story is not lodged exclusively in the talk produced by the speaker, but instead within an interactive field being sustained through the coordinated actions of structurally different kinds of participants who constitute the telling by producing visible as well as spoken embodied signs.

Third, the construction of action is distributed in the hopscotch sequence in at least two different ways. First, as noted earlier, it is distributed across participants and encompasses, for example, both the talk of the speaker and the embodied displays of the hearer(s). As has been very well demonstrated by conversation analysis, action is distributed across turns as well: The talk of subsequent speakers is linked to the particulars of what prior speakers said (Jefferson, 1984; Sacks, 1995; Schegloff, 2007). Second, action is distributed in the sense that it is constructed from structurally different kinds of sign phenomena that mutually elaborate each other. The distributed organization of action thus encompasses both interaction between participants as separate actors and interaction between diverse sign phenomena. For analysis of the distributed organization of cognition within endogenous communities, see Hutchins (1995).

To be a speaker is thus to occupy a particular position within a dynamically unfolding interactive field structured through public sign use. As will be seen in the later discussion, this is what makes it possible for Chil to function as a consequential speaker despite a catastrophically impoverished vocabulary and syntax. However, much analysis of language and pragmatics takes as its point of departure a very different model of the speaker.

CONCEPTUALIZING THE SPEAKER

The speaker, or more broadly the speaker-hearer relationship, sits at the absolute center of human language production as the crucial site where language is constructed, organized, and understood. For example, according to Chomsky (1965: 4), "The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance."

Such a view of the speaker-hearer, which is by no means restricted to formal linguistics, contains a number of crucial assumptions about what a speaker is and what phenomena and resources are relevant to the analysis of the cognitive organization of speakers and their capacities to produce language. First, a speaker is an entity capable of producing linguistic signs. Second, rather than simply producing isolated linguistic signs, such as single morphemes or words, a speaker has the subtle, flexible ability to combine a large repertoire of linguistic signs, all with a range of structurally different properties (nouns, verbs, etc.), into larger units. A competent speaker can use the combinatorial powers of syntax to construct an infinite variety of complex sentences. The speaker is thus defined by the ability to construct sentences, that is, to perform a complex symbolic calculus with the uniquely human signs provided by language. Third, as indicated by the combination of speaker and hearer into a single hyphenated entity, the hearer is simply the mirror image of the speaker, specifically an entity capable of properly decoding the linguistic signs produced by the speaker. Fourth, this ability to produce and decode sentences exhausts the phenomena that have to be taken into account analytically. The primary task posed for the study of human language is describing how an infinite variety of grammatical sentences can be constructed with finite resources. Speakers are simply and exclusively, by definition, the entities where this process occurs.

This picture of human language is deeply consequential for all research following from it. It specifies what types of phenomena count as data for appropriate inquiry (idealized, well-formed, grammatical sentences) and what is irrelevant (the degenerate actual utterances produced by speakers in the real world, with their performance errors such as restarts and sentence fragments [Chomsky, 1965: 3–4]). A particular geography of cognition and language is created. First, the domain of analysis is the isolated sentence. Second, the processes providing for the construction and understanding of sentences are to be found within the mental life of the speaker-hearer. (Saussure's famous diagram of heads linked together by the circuit of speech provides a vivid picture of these assumptions).

There has of course been consistent and very strong opposition to this view of language with its restricted analytic geography. For example, in the first half of the twentieth century, Bakhtin and Volosinov (1973) provided powerful demonstrations of the intrinsic social and historical character of human language. In the second half of the same century, both sociolinguistics and the ethnography of speaking provided powerful analysis of how language was intricately tied to – indeed constitutive

of – a host of social phenomena. Conversation analysis took as its point of departure sequences of action by multiple speakers rather than isolated sentences. Anthropology has consistently investigated language as an indivisible component of human social, cognitive, and cultural life. Indeed, the differences in perspectives being noted here are well illustrated by the way in which in America, linguistics began the twentieth century as a subfield of anthropology and ended the century as a separate discipline with much closer ties to psychology than to any other field with a strong interest in human social life.

Note, however, that even this later work continues to take as its point of departure a view of the speaker as an entity capable of producing rich, interesting linguistic structure. Variability in phonological and lexical structures produced by speakers in different social groups or positions provides central data for the study of variation in sociolinguistics. The study of reported speech, begun with such power by Volosinov (1973), and narrative more generally, requires the existence of individual speakers capable of producing rich, linguistic structures that weave together the voices of the current speaker with the quoted voices of others.

In the present work, I intend to probe such assumptions by investigating how a man with a vocabulary restricted to three words is nonetheless able to function as a very powerful speaker in conversation. Without in any way denying the tremendous importance of individual speakers being able to produce complex linguistic structure, I want to argue that the primary locus for the analysis of human language is not the private mental life of individual speakers, but instead a public, embodied interactive field that is sustained and constituted from moment to moment by the coordinated, differentiated work of structurally different kinds of actors (C. Goodwin, 1984). Rather than simply providing through their talk the output of a complex symbolic calculus with linguistic signs, speakers become speakers by using both talk and other signs to perform relevant action in this interactive field. To build utterances, they frequently do not start sentences from scratch within the splendid isolation of a private world, but instead creatively reuse and reshape materials provided by the talk and action of others (Dubois, 2007; M. H. Goodwin & C. Goodwin, 1987). The actions of hearers are central to the dynamic processes through which utterances, and the sentences they contain, are constructed and frequently modified even as they are emerging (C. Goodwin, 1979). Because hearers are largely silent, many of the signs they use to participate in this process are made with their visible bodies. It is thus necessary to move analysis beyond exclusive focus on phenomena in the stream of speech to encompass the embodied actions of multiple participants. In brief, speakers are constituted through their ability to use language structure as a form of public practice to build relevant action within this dynamic, ever changing, interactive field.

CHIL AND HIS CAPACITIES AND RESOURCES

In 1979, when Chil was sixty-five years old, a blood vessel in the left hemisphere of his brain ruptured. He was left completely paralyzed on the right side of his body and with a vocabulary that consisted of only three words: *yes*, *no*, and *and*.[1] Despite this, he continued to function as a powerful actor in conversation, and indeed had an active social life in his community, going by himself to a coffee shop in the morning, doing some of the family shopping, and so forth.

Despite the limitations of his vocabulary, Chil retains an extensive repertoire of other semiotic resources. First, his understanding of what others are saying, the language they produce, is excellent. Second, he has very expressive prosody, which he produces over both his *yes*'s and *no*'s and over "nonsense" syllables such as "duh", which seem to be spoken precisely to carry relevant prosody (C. Goodwin, M. H. Goodwin, & Olsher, 2002). Third, though completely paralyzed on his right side, Chil uses his left hand to produce a varied and important range of gesture, including pointing (C. Goodwin, 2006; Kita, 2003) and hand shapes displaying numbers (C. Goodwin, 2003). Fourth, by living at home with his family and caretakers in the town that has been his community for almost forty years, he inhabits a world that is not only meaningful, but which can be recognized in relevant ways by those around him. He can thus use actions such as pointing to invoke meaningful phenomena in powerful ways. However, as will be seen in the following discussion, his inability to accompany this pointing with relevant language can produce puzzles to be unraveled rather than transparent reference. Finally, unlike many people suffering from aphasia, Chil's timing as a participant in interaction is rapid and fluid. This may in fact be a by-product of the severity of his impairment, because he does not spend extensive time trying to find and produce words.

A SEQUENCE OF ACTION WITH CHIL

The sequence in Transcript 13.2 will be used to investigate some of the practices through which Chil is constituted as a speaker. Chil lives in New Jersey, in a suburb of New York City. His son, Chuck, who now lives in California with his wife Candy, is visiting. Chuck and Chil are sitting in the living room of Chil's house, where Chil has just finished eating a grapefruit in the bowl on the tray on his lap. The grapefruit is part of a shipment of citrus fruit Chil has ordered from Florida, and is treated as especially tasty, a special treat.

[1] I investigate different aspects of how Chil builds meaning and action in concert with others in a number of different papers. Providing the reader with a short picture of his abilities is necessary in all of them. The description of Chil offered in this section of the current paper can therefore be found in other papers as well including Goodwin (in preparation).

A major argument in this chapter is that Chil, and speakers in general, build utterances and meaning by using a range of different kinds of sign phenomena that mutually elaborate each other. Transcript 13.2 therefore includes not only a transcription of the talk using as its point of departure the system developed by Gail Jefferson (Sacks, Schegloff, & Jefferson, 1974: 696–735), but also drawings of the participants, which show not only relevant gestures but also how they position their bodies with respect to both each other and features of their surrounding environment. Lines link pictures of gestures to the place where they occur in the talk. Chil's prosody, which is extremely important, is indicated in a very incomplete way with a few pitch tracks. How Chil uses prosody to organize action in this sequence is the focus of a separate paper (C. Goodwin, 2010a). Where relevant, a summary of that analysis will be provided. For the moment, the reader is alerted to the variety of pitch contours that occur within a very short period of time over very similar semantic units, such as versions of *No no*.

The transcript contains a lot. If I had unlimited space, I would prefer to introduce elements of it progressively, so that the reader could focus on one phenomenon at a time, and ideally to see the actual video. I hope that the use of drawings rather than verbal glosses, and the attempt to include as much white space as possible, enables the reader to clearly see what is happening. Briefly, after Chil finishes his grapefruit, he points repetitively first to his bowl, and then to the space in front of him. After each of these pointing sequences, Chuck states a guess as to what Chil might be trying to indicate, but Chil rejects each of Chuck's guesses.

Indexical incorporation

Unlike the ideal image noted previously of language, meaning, and action residing within self-contained sentences and/or utterances, none of Chil's *No No*'s (lines 8, 11, 14, 17, and 19) can be appropriately understood in isolation. Each is a response to what his interlocutor has just said. By tying to a specific bit of talk produced by someone else, Chil is able to indexically incorporate what was said there into the organization of his own utterance. Thus his "No ***No***." in line 14 is heard as indicating that he doesn't want the fruit taken away, while his "No no." in line 17 signals that he doesn't want Chuck to take some of the fruit back with him. Even though the semantic units, the double *No*'s, are the same in both cases, and each signals opposition to what was just said (a core meaning possibility for *No*), Chil tells Chuck something quite different each time he uses this expression (see also Du Bois, 2007; Ochs, Schieffelin, & Platt, 1979). What Sacks (1995) called *tying techniques* are central to the process through which human beings build meaning and action by combining different sign complexes into larger, meaningful wholes (a current utterance and the utterance[s] it is being tied to). Chil lacks the semantic and syntactic ability to construct sentences that would state either of the previously stated propositions by himself. However, he is able to vastly expand his repertoire as a speaker by sequentially tying to the particulars of the complex talk and language structure of his interlocutors. He can use their productive linguistic ability to state propositions and make moves within the complex symbolic world being constructed through the local interaction, which would be completely impossible for him as an isolated individual (C. Goodwin, 2007).

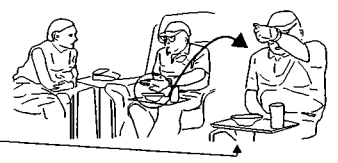
It was noted earlier that definitions of the speaker in fields such as formal linguistics take the ability to produce syntactically complex sentences as central. What we find here raises another possibility. Rather than being a hidden capacity that resides within the mental life of the individual, language is organized as a public interactive field. Actors with different abilities, or who may occupy different positions within that field (C. Goodwin, 1984), can nonetheless contribute to production and organization of action through language in different ways. By participating in this field, and using language – some of it from his interlocutor – to build relevant action, Chil acts as a speaker capable of conveying complex propositions through language use in interaction, despite his inability to produce sentences as an individual.

Combining signs

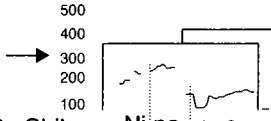
Chil's ability to produce syntactic units – that is, to combine morphemes and other linguistic units into larger wholes – is catastrophically impaired. He can link multiple *Yes*'s or *No*'s together (lines 14 and 17 in Transcript 13.1 provide examples of double *No*'s) and produce expressions such as "Yes No." However, combinations such as these mark the limits of his ability to concatenate his limited vocabulary into units that contain multiple, semantically meaningful morphemes.

However, focusing only the arrangement of linguistic signs within the stream of speech provides a very impoverished picture of Chil's combinatorial powers. Instead, it is relevant to examine how he combines structurally different kinds of signs into larger multi modal packages in order to build relevant action.

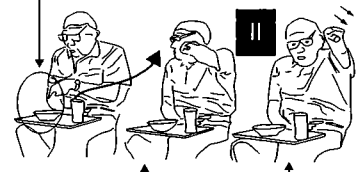
In the sequence being examined here, Chil makes use of variants of *No* extensively. Within talk-in-interaction fluent speakers who produce a *No* in response to something that someone else has just said, frequently add to their turn an account. This can provide a reason for why what the prior speaker said is being rejected and/or it can offer an alternative to a proposal made in the turn being responded to. The gestures of Chil in images I, II, and III in Transcript 13.2 are tied to his *No* in just this way. His utterance does not reside in the talk alone. Instead the complex multi modal action of *No + Gesture* not only rejects what Chuck has just proposed, but attempts to display a correct alternative to Chuck's candidate gloss of what Chil is trying to say. Chil's inability to state this



7 Chuck: Had enough

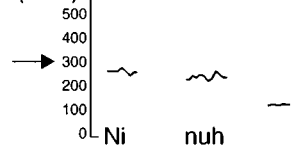


8 Chil: Ni na n o.

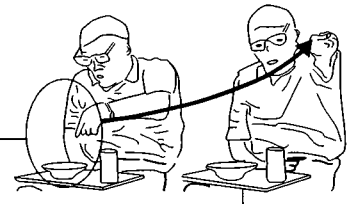


9 Chuck: Diya want-

10 Chuck: Want me ta ***get*** some.(0.24) No.

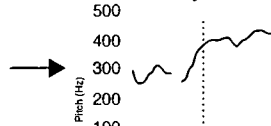


11 Chil: Ni nuh



12 (0.9)

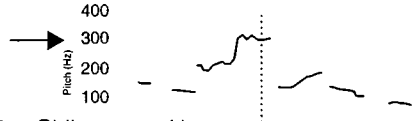
13 Chuck: Do you want me ta take that awa:y.



14 Chil: No ***No.***

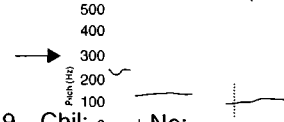
15 Chuck: Oh: you like it. Yeah yuh-

16 Oh s:- Take some back with us.=

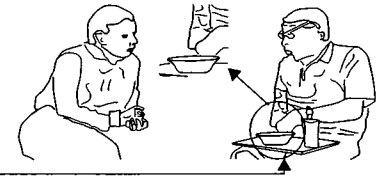


17 Chil: =No no.

18 (0.6)



19 Chil: No: No. *((deeper voice))*



20 Chuck: Oh no it's illegal.

21 Yeah.

22 Chil: Y(h)a h(h)a o(h) h n(h)o n(h)o n(h) o.

23 Chuck: Hi hi ha ha O(h)k(h)ay.

Transcript 13.2.

alternative with unambiguous conventional signs leads to failure. He is nonetheless constructing a complex action with differentiated parts, and moreover one that has the same underlying shape as those produced by fluent speakers in this sequential position.

The gestures Chil produces here have a significant combinatorial organization in their own terms. Chil is building his action to Chuck by constructing two gestures that are linked to each other as components of a single gesture package.[2] In all three cases, Chil starts by using his hand to point at, or some other way indicate (Image 1), the bowl on his lap. Then, with a second movement that uses the first one as a point of departure, he raises his arm to point in front of him.

Let me be explicit about why the pairs of gestures that occur here are being described as linked gestures that are tied to each other in such a way that the second builds from the first. First, elsewhere Chil does in fact construct action by producing sequences of gestures in which a next builds from a prior. A clear and simple example is provided by the way he uses his single functioning hand to produce numbers (see C. Goodwin, 2003). Because he cannot display a number greater than five with his fingers, larger numbers are constructed by producing a hand shape with five fingers followed by a second hand shape with more fingers, say two. His interlocutors add the two gestures together to obtain seven and state this number as part of their response to Chil's gestures. Examples such as this also demonstrate how many of Chil's gestures, including the ones being examined in this paper, are produced so that recipients can perform systematic operations on them, and how recipients are faced with the task of uncovering and explicitly stating how Chil's gestures are to be interpreted with respect to an unfolding course of action. As part of such gesture practice, Chil is able to systematically indicate that particular hand shapes are to be linked, and thus added together, and also to establish junctures between gestures to demonstrate to his addressees that a subsequent gesture is not been seen as tied to a prior one. Second, in describing Chil's two gestures as linked together into a larger action package, rather than as two separate gestures, I am relying on the way in which Chuck responded to them within the scene itself (to be discussed later) and the way I see them now when looking at the video. There are a number of features of the production of Chil's gestures in the present sequence that display their status as elements being combined into a larger whole. First is their timing. There is no break whatsoever between the production of the two gestures. Second, other co-occurring embodied activity by Chil treats the two gestures as part of a single action package. Both gestures occur within a single excursion from home position (Sacks & Schegloff, 2002). This is accompanied by gaze patterns that mark the production of a conversational move. Thus for the first pair, Chil gazes toward his addressee at the beginning of the gesture, then toward the direction indicated by the second gesture, at which point he both returns his gaze to his addressee (an action that may mark the expectation of a response to the now completed package) and starts to return his hand to home position. Third, as can be clearly seen in Transcript 13.2, when Chuck fails to understand what Chil is indicating, Chil repeats both gestures together, treating them as a single, coherent course of action.

I am using the term gesture to describe the pointing Chil does here. However, in a number of important respects, what occurs here is unlike most human gestures. First, gestures typically accompany rich, fluent talk being spoken by the same person who is producing the gesture. Chil's first pair of gestures (Image I) do not accompany any talk by Chil. Instead, they stand alone as a self-contained action, and indeed they are treated by Chuck as constituting by themselves what is functionally a complete utterance. Chil, unlike fluent speakers, is using gesture without talk (lines 5–6), or with very minimal talk (the dual No's that accompany his gestures in lines 8–12), to build action. Second, these linked gestures are ordered relative to each other to create a larger whole.

One way of seeing this sequence is to see that the first gesture is in some way locating the bowl (or what is in it) as something like a topic and the second pointing gesture as making a comment about that topic. This is in fact the way that Chuck interprets these actions in his responses to them. For example, in line 13, he inquires if the bowl [topic] is to be taken away [comment on that topic]. The sequence of actions Chuck describes, picking up the indicated bowl and carrying it into the kitchen, a place where Chil might possibly be pointing, is a very plausible – although, as it turns out, incorrect – reading of the actions being indexed by Chil's gestures.

Topic-comment constructions are hallmarks of the kinds of propositions that can be expressed linguistically by well-formed sentences. Indeed Chuck is treating Chil's package of linked gestures as in some way equivalent to a unit, such as a sentence, utterance, or turn-at-talk, that states a proposition. Thus Chil's multiple gestures are treated as a whole that is 1) referentially depicting some state of affairs (the subject of Chuck's subsequent glosses) and 2) pragmatically calling for some particular, relevant subsequent action.

Semiotic agency

Chuck is treating Chil's moving hand with its gestures as constructing signs indicating some other action that Chil wants Chuck to perform. Such a process of attempting to initiate (on most occasions successfully) a course of action through systematic sign use will be called *semiotic agency*.

There is a large, important, and relevant literature on both agency in human social life (Bourdieu, 1977; Giddens, 1984; Ortner, 2006) and the organization of agency in language (Agha, 2007; Ahearn, 2001; Duranti,

[2] See (C. Goodwin, 2003) for specific analysis of how Chil subdivides movements made by his hand and arm into discrete gestural packages.

2004; Kockelman, 2007). However, for clarity, I want to focus for the moment specifically on phenomena visible in Transcript 13.2.

First, as is clearly demonstrated by Chuck's three failed attempts to gloss what Chil wants done, Chuck doesn't understand what Chil is trying to say. He nonetheless continues to treat Chil as attempting to accomplish a course of action through sign use, even in the absence of understanding what that action might be or what the signs being gazed at might mean. Chil's semiotic agency as something that his interlocutor recognizes and uses as a point of departure for subsequent action does not require grasping what his signs mean (for example, how agency is encoded by specific signs), but instead resides simply in Chil's visible use of signs to initiate action (see also Duranti, 2004: 455).

Cooperative semiosis

Each party is producing signs so that someone else can perform systematic operations on those signs. Thus Chuck produces a candidate gloss after each of Chil's gestures, and Chil in turn operates on each of these glosses by rejecting them. Chil's signs do not stand alone as isolated self-contained entities but instead require the actions of another to lead to meaningful action. This arrangement, in which action emerges through a process in which signs are produced to be operated on by another, will be called *cooperative semiosis*.

Despite a pervasive focus in fields such as linguistics on the investigation of isolated sentences within frameworks that include only the mental life of the individual speaker, cooperative semiosis is in fact central to the ways in which participants in interaction construct utterances, sentences, and action. Consider the exchanges in Transcript 13.3. In each of the three examples (taken from audio recordings of children arguing in their peer group), subsequent speakers build next actions by performing systematic operations on the sign complexes that constitute the utterance of the prior speaker. These linked utterances, which we have examined as format tying (M. H. Goodwin & C. Goodwin, 1987), are instances of what DuBois (2001) investigates as dialogic syntax. The practice DuBois introduced of writing the transcribed talk in columns to highlight repetition in structure is used in Transcript 13.3.

A	Tony:	Why don't you		get out my yard.	
	Chopper:	Why don't you *make* me		get out the yard.	
B	Chopper:	Don't gimme that.			
		I'm not *t*alkin ta you.			
	Tony:	I'm *t*alking ta you.			
C	Malcolm:		She	a dropout.	*((talking about Ruby))*
	Ruby:	*I know*	You	a dropout.	
	Malcolm:	She says She know	She	a dropout.	

Transcript 13.3.

In **A**, Chopper incorporates all of what Tony just said into his own next utterance, while adding the words "***ma***ke me" to transform what Tony said to him into a return challenge directed at Tony. Chopper's talk is built by reusing with systematic transformation, and embedding within his own reply, the sign complex just produced by Tony. In **B**, a similar transformation and reversal of action occurs, only here it is accomplished by deleting the negation found in the first utterance. In **C**, this process of building next actions through systematic operations on the materials provided by a prior sign complex is extended over several utterances, producing by the third move a sentence of some complexity.

Such processes of cooperative semiosis constitute a central locus for the organization of cognition, action, and language practice. Crucially, for the study of language and grammar, they extend the scope of investigation simultaneously outward, beyond the individual sentence and speaker, and inward, into the range of alternative possibilities for combining and shaping into larger wholes the signs that construct utterances and sentences. Thus, among the central questions posed for both analysts and participants (such as newcomers to a language) investigating grammatical organization are: 1) what are the relevant grammatical units into which the sound structure provided by the stream of speech is to be divided; and 2) what are the possible and permissible ways of combining these units? Unlike the situation with idealized, well-formed grammatical sentences, within processes of multi-utterance cooperative semiosis, the actions of participants create an environment where 1) the stream of speech is parsed into discrete, relevant building blocks; and 2) populations of alternative possibilities for combining these signs into larger units are made publicly available (C. Goodwin, 2006). In the brief examples in Transcript 13.3, these include the contrast between a sentence with and without negation (**B** in Transcript 13.3), practices for transforming a prior structure, complete in its own terms, into a subcomponent of a new, more complex unit (**A** in Transcript 13.3), and division of the stream of speech into relevant sub-units including paradigms for alternative items that can occur in the same slot. Thus repetition with significant contrast breaks a longer strip of talk into discrete language relevant units ("talkin ta you" as something that can be preceded by either "I'm not" or just "I'm," with that difference being marked as consequential, "know" as a unit that can be both preceded and followed by alternative pronouns, etc.). This same process displays the varied pronouns – *I, you, she* – that can occur in alternation with each other in the same position within an utterance, among other things.

These practices for rapidly manipulating language structure have strong social and affective consequences. In Transcript 13.3, competing speakers contest their skill and character by elegantly reshaping the very sign complexes just produced by their opponents so that they return as attacks against their original author. Though

such argument might be considered a vivid, special case, the pervasiveness of phenomena such as tying techniques, in which subsequent utterances mark specific ties to earlier talk (Sacks, 1995) and sequential organization more generally (Sacks et al., 1974), demonstrate the central importance of cooperative semiosis in the organization of language and action. Such semiotic manipulation of sign complexes provided by others is precisely how participants build subsequent social action through interaction with each other, even at a very early age (Keenan, 1983). Engaging in these practices requires orientation to not only the process of building one's own sentences and utterances, but also fine attention to the detailed structure of signs being produced by others, and to the operations that can be performed on those signs. What is required in such language practice is the development of skill at seeing structure and grasping possibilities for action within a complex public environment being constituted through the continuous transformation of sign complexes.

Cooperative semiosis provides the construction of action through language (and with other signs) with an autopoetic organization. The very practices used to build action through language simultaneously provide a public, endogenous analysis of how the stream of speech can be decomposed into relevant sub-units, and the permissible ways in which these units can be combined to build relevant larger grammatical units. This is not only visible to participants, but is actively being used by them to construct the actions they are engaged in.

Within this process, the authorship of the individual speaker, as, for example, the party who constructs the sign complex that makes up the utterance that he or she produces, can become distributed across multiple actors and utterances. For example, most of the utterance produced by Chopper in **A** was originally assembled by Tony. Such distributed authorship does not, however, diminish the agency and power of the individual speaker. Thus, though Chopper reuses Tony's words, he constructs himself as actor who is producing a strong counter to what Tony just said with these very same words. Such reuse of language that others have sedimented with meaning in the past has a clear resonance with Bakhtin (1981). However, here the language being used is what has just been said in the local interaction by the current speaker's co-participants. This process of building utterances that incorporate language structure produced by others is central to Chil's ability to act as the speaker and author of complex, locally relevant statements (C. Goodwin, 2007).

In brief, through a process of cooperative semiosis, participants build action by secreting sign complexes into a public environment where others build subsequent action by performing systematic operations on the structures created by their co-participants.

Chil has only the most limited ability to speak linguistic signs by himself. However, cooperative semiosis encompasses the full range of signs participants use to build action in concert with each other. Embodied signs without language constitute a crucial component of Chil's semiotic repertoire.

Discovering what Chil was doing

Chuck's efforts to discover what Chil is trying to tell him with his linked gestures in Transcript 13.2 are ultimately unsuccessful. However, approximately a minute and a half later, Chuck goes to look for his wife, Candy, and is told by someone else in the house that she is out walking the household dogs on streets that wind around the back of the house, an area that falls within the trajectory of Chil's second point toward some place in front of him. On hearing this, Chil redoes the earlier point to the bowl and gets Chuck to recognize that with his earlier gestures he was trying to indicate that he wanted Chuck to offer some of the delicious grapefruit they had just eaten to Candy.

Once what Chil was trying to say is known, we can see that Chil's linked gestures provide a clear physical enactment of this request. Chil first uses the bowl to index the grapefruit he wants to give Candy, and then, with his moving point, enacts picking up the target of the first pointing gesture and carrying it to Candy by pointing toward the space where she currently is. With the benefit of hindsight, it is possible to argue that Chil's linked gestures provide an elegant solution to the problem of telling this to Chuck without language. However, when it occurs, Chuck is completely unable to figure out what this gesture complex means.

TRAPPED IN A WORLD OF ICONIC AND INDEXICAL SIGNS

It took considerable work and time for Chil to get Chuck to understand that he wants Candy to try some of the special grapefruit he has just eaten. Indeed, initially Chuck is not able to figure this out at all, despite what appear in hindsight to have been rather clear gestural images of what Chil wants done. If Chil could have produced an utterance such as "Give some of this delicious grapefruit to Candy," none of this work would have been necessary. If Chil could have spoken the single word "Candy" while producing his gestures linking the grapefruit to a depicted course of action, it seems plausible that Chuck would have grasped what Chil was trying to say quite quickly.

With the exception of his limited vocabulary, the signs Chil produces are indexical and/or iconic. Thus, with his pointing hand, he indexically locates and marks as relevant the bowl with its grapefruit. Then, with the subsequent moving point he iconically mimes picking up the bowl and taking it somewhere (or with hindsight perhaps giving it to someone in the space being pointed at). One major problem for Chuck is the intrinsic but open-ended meaningfulness of signs structured in this fashion. Chil's indexical points in Images II and III in Transcript

13.2 are clearly indicating something just under his finger. However, Chuck is unable to determine whether it is the bowl itself or the fruit that the bowl contained and, if the latter, whether it is that particular bit of fruit or some class that it instantiates, for example by functioning as a sample of the fruit from Florida that was just received. Similar but more complex issues arise when Chuck tries to figure out what Chil is trying to indicate in the space(s) in front of him with his second point. Moving from indexical to iconic signs, Chil's arm movement from the bowl can iconically depict the act of offering something to Candy, but it can just as well depict the act of taking the empty bowl back to the kitchen (line 13 in Transcript 13.2), or getting more fruit (line 10). The problem, at least in human communication,[3] for signs based entirely on either contiguity in an environment that is necessarily ambiguous and complex (note the difficulties that arise with what would seem to be an exemplary simple object: Chil's bowl), or signs constructed in terms of resemblance that is necessarily partial, is that rather than moving immediately forward and acting on what the sign indicates, an actor trying to use that sign as the point of departure for subsequent action is faced with the task of resolving its ambiguities, as Chuck is here. This is a process that can seriously disrupt the onward course of action that use of the sign was designed to initiate in the first place. Chil is trapped in a world of intrinsically meaningful, but inherently ambiguous, signs. This has consequences for not only himself, but also his interlocutor, as it structures how their interactions with each other are organized.

Conventionalized signs and the organization of action

The flow of action in interaction with Chil is disrupted by the work required to determine exactly what he is attempting to indicate through his iconic and indexical signs. However, if the signs being used were not based on either contiguity or partial resemblance to what they signify, but instead completely arbitrary, to be understood by conventions, norms or rules that are recognized by members of the community using these signs, none of the issues that so plague Chil and his interlocutor here would arise. There would be no search for what a sign might resemble or indexically indicate. Instead a conventional sign is immediately grasped by someone who has mastered the rules of its use. Whatever actions the sign is being used to accomplish can move forward without the necessity of clarification.

Signs with such structure are, of course, what are called symbols within the traditions that follow from Peirce (Deacon, 1997; Peirce, 1998) and constitute a central hallmark, indeed a defining characteristic of human language.

Much analysis of the changes made possible in human cognitive and social life by language has focused on its extrordinary representational power. Those who have mastered language are no longer tied to the here and now but can represent not only past, spatially distant and future events, but also, through symbols and their combination, construct whole new cognitive worlds. Collective knowledge can be synthesized and passed on through generations. Socially, through language, human beings were able to construct more complex and extended social relations than other primates could by simple grooming (Dunbar, 1996), and were able to conceptualize and position actors within more complex kinship and social networks.

This is certainly true. However, the phenomena being investigated here suggest that in addition to their representational power, arbitrary symbols make possible radically new forms of social action that are intricately tied to distinctively human forms of cognition. With symbols, other actors can immediately recognize and use the signs being produced by their interlocutors. Such rapid, frequently transparent recognition makes possible complex sequences of actions that can move forward fluidly.

This process links the ongoing organization of human action to cognition and culture. As is well demonstrated by the diversity of human languages and cultures, the specific phenomena that are the focus of the participants' attention as they build action together can be extraordinarily varied and complex, and may not be recognized by members of other groups who have not mastered the current symbols. However, insofar as rapid recognition is useful, indeed required for fluid action, intersubjectivity emerges within this process as a form of public practice. If participants are to build rapid and fluent action together, they must in some relevant sense understand what each other is doing and recognize as quickly as possible the specific signs being produced by others. Such frameworks of shared understanding can be built from within the process itself. The sequential structure of action in interaction provides endogenous resources for repairing misunderstandings (Schegloff, Jefferson, & Sacks, 1977), and for calibrating the forms of perception and embodied action required (C. Goodwin, 2010b) for the collaborative construction of action.

Through this process, a primary environment for human cognition linked to action not only emerges, but is continuously sustained and reconstituted through the unfolding organization of mundane human interaction. To engage in this process, participants must be able to build action by producing signs that will organize the actions of others while attending to the details of the cognitive and pragmatic environments that have been constructed through the actions of others. This suggests that factors promoting the development and evolution of human language were not simply its new and powerful representational capacities, but rather the process

[3] For other animals, processes of evolution might strongly constrain how members of a particular species interpret iconic and indexical signs.

through which the cognitive lives and abilities of different actors were linked together in ways that enabled the fluent accomplishment of radically new forms of action that changed the world while simultaneously construing it in intricate, relevant ways.

THE EMERGENCE OF THE INDIVIDUAL WITHIN THE MATRIX OF COOPERATIVE SEMIOSIS

The individual, especially the mental life of the individual, and the social world are sometimes treated as phenomena that are analytically distinct (for example, on a gross level, with notable exceptions, the organization of psychology and sociology as separate fields of study). Chil makes meaning in concert with others through processes of cooperative semiosis, a process that is intrinsically multi-party and thoroughly social. Rather than being lost in this process, Chil's cognitive life as an individual is constituted as something that is public, consequential, vivid, and flexible.

These same issues are relevant to the organization of his agency. One issue that has emerged in the study of agency is the difference between individual and "social" agency, the latter typically being investigated as agency lodged within a social or corporate group, rather than an individual. Insofar as Chil's agency is organized within ongoing processes of cooperative semiosis, it is thoroughly social (see also Al Zidjaly, 2009). Getting Chuck to offer some fruit to Candy requires that Chuck perform appropriate operations on the signs Chil is producing. His signs, like all signs, require interpretants. However, what emerges from this process is not the amorphous diffusion of his agency into a social group, but instead a very vivid recognition by his interlocutors of his agency as an individual, for example as someone who has something unique to say, even when they can't grasp what that might be.

Though being produced for others, all of Chil's signs are indexical expressions of his feelings, thoughts, and actions at the moment. His prosody provides a particularly clear example. The way that he changes prosody over the same two word expression ("No No") in Transcript 13.2 from line 14 (strong opposition to Chuck's proposal to remove his fruit) to line 17 (Chuck has proposed doing something illegal) constructs a vivid display of a creative actor with a rich mental and cognitive life that is capable of rapid change to adapt to changing contingencies as interaction unfolds (C. Goodwin, 2010a). Rather than being hidden and inaccessible, Chil's mental and emotional life is revealed and made consequential to others through the way in which it is organized within processes of cooperative semiosis.

More generally, the present paper has attempted to describe a range of interrelated phenomena that are central to the organization of human action. These include, first, the way in which action is built by combining semiotic phenomena with quite different properties (linguistic structure, prosody, gesture, embodied participation frameworks that indexically ground the signs that occur within them, etc.) into local constellations where different kinds of signs mutually elaborate each other to create a whole that is both greater than, and different from, the individual elements from which it is built (see the discussion of contextual configurations in C. Goodwin, 2000). Second, action has an intrinsically distributed organization in that it is constructed through the cooperative semiosis of multiple actors. Even when performed by a single individual, both action and meaning emerge within frameworks provided by a history constituted through the public use of signs being used to build the actions that sustain endogenous communities. From such a perspective, what is of interest is not multimodality as a phenomenon in its own terms, but rather the intrinsic organization of human action, which has distinctive forms of organization that are not encompassed within any semiotic modality as an isolated whole, or within the individual actor.

ACKNOWLEDGMENTS

I am deeply indebted to Candy Goodwin and Jürgen Streeck for insightful comments on an earlier version of this analysis.

REFERENCES

Agha, A. (2007). *Language and Social Relations*. Cambridge: Cambridge University Press.

Ahearn, L. (2001). Language and agency. *Annual Review of Anthropology, 30*, 109–137.

Al Zidjaly, N. (2009). Agency as an interactive achievement. *Language in Society, 38*, 177–200.

Bakhtin, M. M. (1981). *The Dialogic Imagination*. Translated by Caryl Emerson and Michael Holquist. Austin: University of Texas Press.

Bourdieu, P. (1977). *Outline of a Theory of Practice*. Translated by Richard Nice. Cambridge: Cambridge University Press.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Deacon, T. W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*. New York: W. W. Norton.

Du Bois, J. W. (2001). *Towards a Dialogic Syntax*.Unpublished manuscript, University of California at Santa Barbara.

Du Bois, J. W. (2007). The stance triangle. In R. Engebretson (Ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction* (pp. 139–182). Amsterdam/Philadelphia: John Benjamins.

Dunbar, R. (1996). *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Duranti, A. (2004). Agency in language. In A. Duranti (Ed.), *A Companion to Linguitic Anthropology* (pp. 451–473). New York: Blackwell.

Giddens, A. (1984). *The Constitution of Society: Outline of the Theory of Structuration*. Berkeley: University of California Press.

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday Language: Studies in Ethnomethodology* (pp. 97–121). New York: Irvington Publishers.

Goodwin, C. (1984). Notes on story structure and the organization of participation. In M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 225–246). Cambridge: Cambridge University Press.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*, 1489–1522.

Goodwin, C. (2003). Conversational frameworks for the accomplishment of meaning in aphasia. In C. Goodwin (Ed.), *Conversation and Brain Damage* (pp. 90–116). Oxford, New York: Oxford University Press.

Goodwin, C. (2006). Human sociality as mutual orientation in a rich interactive environment: multimodal utterances and pointing in aphasia. In N. Enfield & S. C. Levinson (Eds.), *Roots of Human Sociality* (pp. 96–125). London: Berg Press.

Goodwin, C. (2007). Interactive footing. In E. Holt & R. Clift (Eds.), *Reporting Talk: Reported Speech in Interaction* (pp. 16–46). Cambridge: Cambridge University Press.

Goodwin, C. (2010a). Constructing meaning through prosody in aphasia. In D. Barth-Weingarten, E. Reber & M. Selting (Eds.), *Prosody in Interaction* (pp. 373–394). Amsterdam: John Benjamins.

Goodwin, C. (2010b). Things and their embodied environments. In L. Malafouris & C. Renfrew (Eds.), *The Cognitive Life of Things* (pp. 103–120). Cambridge: McDonald Institute Monographs.

Goodwin, C., Goodwin, M. H., & Olsher, D. (2002). Producing sense with nonsense syllables: Turn and sequence in the conversations of a man with severe aphasia. In B. Fox, C. Ford, & S. Thompson (Eds.), *The Language of Turn and Sequence* (pp. 56–80). Oxford, New York: Oxford University Press.

Goodwin, M. H., & Goodwin, C. (1987). Children's arguing. In S. Philips, S. Steele, & C. Tanz (Eds.), *Language, Gender, and Sex in Comparative Perspective* (pp. 200–248). Cambridge: Cambridge University Press.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Jefferson, G. (1984). On stepwise transition from talk about a trouble to inappropriately next-positioned matters. In J. M. Atkinson & J. Heritage (Eds.), *Structures of Social Action* (pp. 191–221). Cambridge: Cambridge University Press.

Keenan, E. O. (1983). Making it last: repetition in children's discourse. In E. Ochs & B. B. Schieffelin (Eds.), *Acquiring Conversational Competence* (pp. 26–39). Boston: Routledge & Kegan Paul.

Kita, S. (Ed.). (2003). *Pointing: Where Language, Culture and Cognition Meet*. Hillsdale, NJ: Lawrence Erlbaum.

Kockelman, P. (2007). Agency. *Current Anthropology, 48*(3), 375–387.

Ochs, E., Schieffelin, B. B., & Platt, M. L. (1979). Propositions across utterances and speakers. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental Pragmatics* (pp. 251–268). New York: Academic Press.

Ortner, S. (2006). *Anthropology and Social Theory: Rethinking Culture and the Acting Subject*. Durham, NC: Duke University Press.

Peirce, C. S. (1998). *The Essential Peirce: Selected Philosophical Writings, 1893–1913*. Bloomington: Indiana University Press.

Sacks, H. (1995). *Lectures on Conversation Vols. I and II*. Edited by Gail Jefferson, with an Introduction by Emanuel A. Schegloff. Oxford: Basil Blackwell.

Sacks, H., & Schegloff, E. (2002). Home position. *Gesture, 2*(2), 133–147.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696–735.

Schegloff, E. A. (2007). *Sequence Organization in Interaction: Volume 1: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language, 53*, 361–382.

Volosinov, V. N. (1973). *Marxism and the Philosophy of Language*. Cambridge, MA; London: Harvard University Press.