2

Interactive Footing

Charles Goodwin

2.1 Introduction

In 1929, V. N. Volosinov (1973) argued that the linguistics of his time was seriously flawed because it took as its primary object of study language structures that were isolated from both context and the social life of their speakers. He proposed that this situation could be remedied by focusing on reported speech, utterances in which the current speaker in some way quotes or reports the talk of another. Noting that 'what is expressed in the forms employed for reporting speech is an *active relation* of one message to another', Volosinov (1973: 116, italics in original) proposed that reported speech constituted a crucial site for recovering the intrinsic dialogic organisation of language. The cogency and power of this argument is well demonstrated by the large body of significant research on both reported speech and the dialogic organisation of language and culture that has flourished since the 1970s in a number of different fields.

In reported speech the voices of separate actors are found in a particular place, a complex strip of talk produced by a single speaker, albeit one quoting the talk of another. While recognising both the originality and the importance of Volosinov's insights, I will argue here that the precise way in which he conceptualised reported speech actually served to hide, and render invisible to analysis, crucial aspects of the very dialogic organisation of language that he sought to probe. These include: 1) dialogue as multiparty sequences of talk in which the voices of different participants are not only heard but actually shape each other; 2) the visible

actions of hearers and thus the multi-party interactive organisation of utterances (which would seem central to Volosinov's (1973) interest in a word as shared territory); and 3) utterances which lack the syntactic and other complexity required to incorporate reported speech.

To investigate such issues I will look first at one of the most powerful and influential models for analysis of the different kinds of 'speakers' that can co-exist within a strip of reported speech: Goffman's deconstruction of the speaker in Footing (1981) (see also Goffman, 1974/1986). In presenting this model Goffman also offered an important framework for the study of participation, and indeed participation seems absolutely central to the dialogic organisation of human language (C. Goodwin, 1981, 1986a; M. H. Goodwin, 1990, 1997, 2000; Goodwin and Goodwin, in press, 1987; Heath, 1986; Rae, 2001). There are, however, serious problems with Goffman's approach to participation. What he provides is a typology of participants rather than analysis of how utterances are built through the participation of structurally different kinds of actors within ongoing courses of action. To probe how such issues are consequential for the investigation of actual talk I will first use Goffman's model of the speaker to describe the different entities visible within a strip of reported speech in a story. This model provides important analytic tools. However, its limitations become visible when analysis is expanded to include the actions of silent (though consequential) participants, such as the party whose talk is being quoted. To further examine the dialogic organisation of both utterances and the speaker I will then look at the impoverished talk of a man with aphasia so severe that he lacks the syntax to construct the rich, laminated utterances required by the frameworks of both Goffman and Volosinov. It will be argued that a quite different notion of both participation and the dialogic organisation of language is necessary to explicate the way in which this man functions as a powerful speaker by incorporating the complex talk of others within his own limited utterances. Such phenomena shed light on the constitution of the speaker and the hearer – the two participant categories that are most central to human language – and to the dialogic processes that provide organisation for the construction of talk through their interaction with each other within this framework.

```
                              (4.0)
   1   Ann:    Well- ((throat clear))          (0.4)
   2           We coulda used a liddle, marijuna.
   3           tih get through the weekend.
   4   Beth:   What h [appened.
   5   Ann:           [Karen has this new hou:se.
   6           En it's got all this like (0.2) ssilvery::g-go:ld
   7           wwa:llpaper,
   8           *hh (h) en D(h)o(h)n sa(h)ys,
   9           y'know this's th'firs'time we've seen this house.
  10           =Fifty five thousn dollars in Cherry Hill.=Right?
  11                   (0.4)
  12   Beth:   Uh hu:h?
  ┌─────────────────────────────────────────────────────────────┐
  │ 13   Ann:   Do(h)n said (0.3)                                 │
  │ 14          dih-did they ma:ke you take this [wa(h)llpa(h)p(h)er? │
  │ 15   Beth:                                   [hh!              │
  │ 16   Ann:   =er(h)di [dju    pi(h)ck  [i(h)t ou(h)t.          │
  └─────────────────────────────────────────────────────────────┘
  17   Beth:            [Ahh huh  huh [huh huh=
```

Figure 2.1. Extract (1)

2.2 Complex speakers

The deconstruction of the speaker offered by Goffman in Footing demonstrates the genuine power of an analytic framework that focuses on the dialogic interplay of separate voices within reported speech. Figure 2.1 is a story in which a teller quotes something that her husband said. The story is about one of the prototypical scenes of middle-class society. Friends have got a new house. As guests visiting the house for the first time, the speaker and her husband, Don, were in the position of admiring and appreciating their hosts' new possessions. However, while looking at the wallpaper in the house Don asked the hosts if they were able to pick it out, or were forced to accept it (lines 13–16).[1]

   Who is speaking in lines 14 and 16? The voice that is heard is Ann's, the current story-teller. However, she is reporting something that her husband, Don, said, and moreover presenting what he did

[1] This same story was analysed from a different perspective, without reference to Footing, in Goodwin (1984). I am indebted to Gail Jefferson for transcribing this talk.
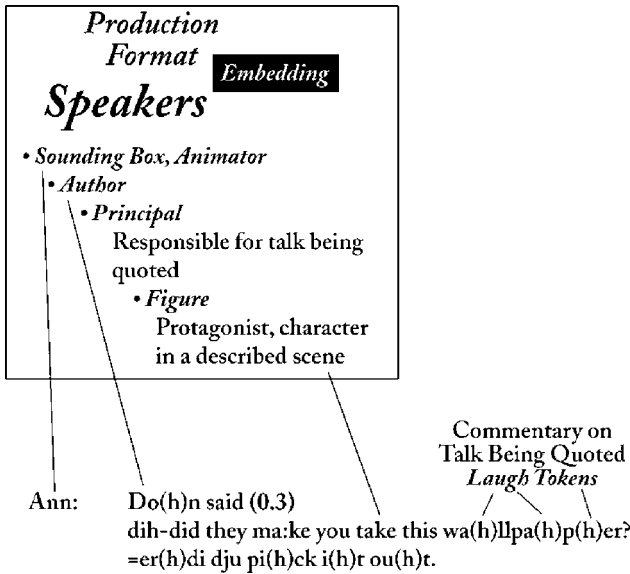
Figure 2.2. Production Format

as a terrible faux pas, an insult to their hosts in the narrated scene. She is both reporting the talk of another and also taking up a particular stance toward what was done through that talk. In a very real sense Ann (the current story-teller) and Don (the principal character in her story) are both 'speakers' of what is said in lines 14 and 16, though in quite different ways. The analytic framework offered by Goffman in Footing for what he called the Production Format of an utterance provides powerful tools for deconstructing the 'speaker' into a complex lamination of structurally different kinds of entities (see Figure 2.2).

In terms of the categories offered by Goffman, Ann is the *Animator*, the party whose voice is actually being used to produce this strip of speech. However, the *Author* of this talk, the party who constructed the phrase said, is someone else, the speaker's husband, Don. In a very real sense he is being held accountable as not only the *Author* of that talk, but also its *Principal,* a party who is socially responsible for having performed the action done by the original utterance of that talk. Goffman frequently noted that the talk of speakers in everyday conversation could encompass an entire

theatre. And indeed here Ann is putting Don on stage as a character in the story she is telling, or in Goffman's terms *animating* him as a *Figure*.

Moreover there is a complex laminated and temporal interdigitation among these different kinds of entities within the space of Ann's utterance. Thus it would be impossible to mark this as a quotation by putting quotation marks before and after what Don said. In addition to the report of this talk, the utterance also contains a series of laugh tokens, which are not to be heard as part of what Don said, but instead as the current speaker's, Ann's, commentaries on what Don did through that talk. Through her laugh tokens Ann both displays her own stance towards Don's utterance, formulating his talk as something to be laughed at, and, through the power of laugh tokens to act as invitations for others to join in the laughter (Jefferson, 1979), invites others to join in such treatment. Ann thus animates Don as a figure in her talk while simultaneously providing her own commentary on what he said by placing her own laugh tokens throughout the strip of speech being quoted.

In brief, in Footing Goffman provides a powerful model for systematically analysing the complex theatre of different kinds of entities that can co-exist within a single strip of reported speech. The analytic framework he develops sheds important light on the cognitive complexity of speakers in conversation, who are creating a richly inhabited and textured world through their talk. In addition to producing a meaningful linguistic sentence, Ann, within the scope of a single utterance, creates a socially consequential image of another speaker. His talk is thoroughly interpenetrated with another kind of talk that displays her stance toward, and formulation of both what he said (e.g. as a laughable of some type), and the kind of person that would say such a thing. Goffman's deconstruction of the speaker provides us with genuine analytic insights, and tools for applying those insights to an important range of talk.

2.3 Recovering the social and cognitive life of hearers

Goffman's speaker, a laminated structure encompassing quite different kinds of entities who co-exist within the scope of a single utterance, is endowed with considerable cognitive complexity. However,

no comparable semiotic life animates Goffman's hearers. In a separate section of the article they are described as cognitively simple points on an analytic grid listing possible types of participation in the speech situation (e.g. Addressee vs Overhearer, etc.).

However, Ann's talk is actually lodged within a participation framework that has a range of structural features that carry it well beyond either a typology of participants, or dialogic text instantiated within the talk of a single speaker. Don, the principal character in Ann's story, the party whose faux pas is being reported, is not just a figure animated through the talk of the story, but an actual person who is present at the telling. Indeed he is seated right next to the story-teller. Elsewhere Goffman defined a social situation, such as the gathering where this story was told, as 'an environment of mutual monitoring possibilities' (Goffman, 1972: 63). Central to the organisation of the participants' monitoring of each other is the way in which those present 'jointly ratify one another as authorized co-sustainers of a single, albeit moving focus of visual and cognitive attention' (Goffman, 1972: 64). Within the field created by Ann's story it is appropriate and relevant for the others present to look at Don, the author of the terrible faux pas, when it is at last revealed. That place for scrutiny of the co-present offender being animated within the talk is defined by the sequential organization of the story, that is at its climax. As principal character in the story Don is faced with the task of arranging his body for the scrutiny it will receive when that moment arrives. When a videotape of the telling is examined it can be seen that, as Ann quotes what he said during lines 14 and 16, Don's face and upper body perform visual versions of her laughter. Indeed, on looking at the video, it appears that two separate bodies are performing the same laugh. For example, there is quite precise synchrony between escalation in Ann's vocal laughter and Don's visual displays. Thus, just as laugh tokens first appear in 'wa(h)llpa(h)p(h)er' in line 14, Don's face starts to form a smile/visual laugh. As Ann's laughter becomes more intense in line 16 Don's face matches her escalation with more elaborate head movements, wider opening of his mouth, etc.

The participation framework relevant to the organisation of Ann's story, and most crucially the quoted speech within it, thus extends far beyond structure in her talk to encompass the embodied actions of others who are present. Don is faced with the task of

systematically organising the displays being made by his body with reference to the phenomenal field being constituted through the unfolding structure of Ann's story. Moreover, the precision with which he coordinates his actions with structure in Ann's talk suggests that he is not waiting to hear what she has actually said before he begins to act, but instead projecting what she is about to say. Structure in her talk provides resources for such projection. The phrase 'Do(h)n said' marks that a quotation will be produced next (the story clearly concerns what Don said to their hosts). Over 'Do(h)n said' in line 13, Don, who had been looking to his side attending to something else, moves his head back to the focus of Ann's story, and sits next to her in a posture that places his head in a position where it is available for story-relevant scrutiny by others (see 3 in Figure 2.3). However, he places his hand over his mouth, the region of his face that will break into a smile a moment later when Ann actually reports what he said. Thus, over line 13, he positions himself like an actor moving to the wings just before his projected entrance on stage.
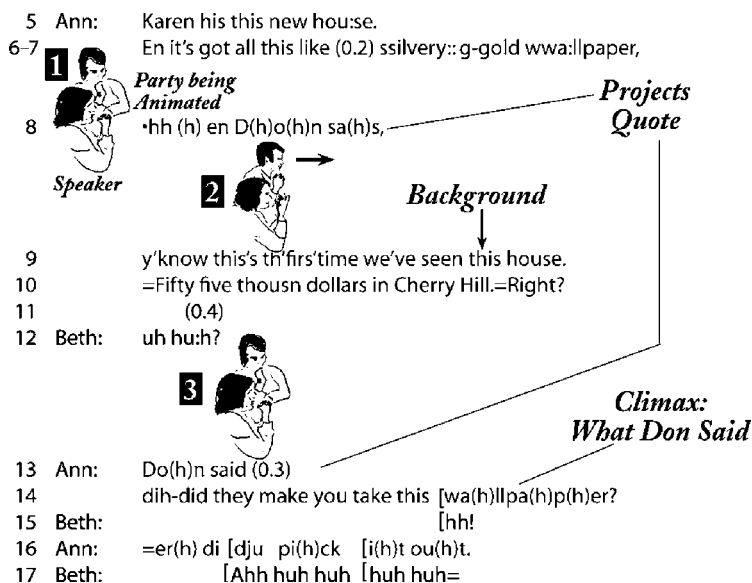

Figure 2.3. Extract (1), Detail

The way in which Don moves to a preparatory position just before he becomes the story-relevant focus of attention in line 14 strongly suggests that he is using the emerging structure of Ann's talk to make projections about what he should do next within the multi-party interactive field invoked by the story.[2] Ann's story provides materials that permit more systematic investigation of this possibility. In line 8, with 'en D(h)o(h)n sa(h)s', Ann also projects that she is just about to quote the terrible thing that Don said to their hosts. Over this talk Don moves into a preparatory position (see Figure 2.4 on page 29) that is almost identical to the one he assumes several moments later over line 13. However, right after this, in lines 9–11, Ann abandons movement into the projected quotation and provides her hearers with additional background information about the house being assessed (its price and location). In essence, lines 9–11 constitute a parenthesis as additional background information is embedded within the climax segment begun in line 8.

The effect of this is that in line 9 Don has the rug pulled out from under his feet; the event that he had moved into position to be ready for – the quotation of his faux pas – is suddenly withdrawn. If Don is in fact organising his body with reference to projections about upcoming events in the story, then he is now positioned inappropriately (i.e. he has arranged his body for the scrutiny of others as his talk is quoted, but the speaker has suddenly shifted to further description of the house itself). As soon as the parenthesis is entered, Don abandons his preparatory position, and indeed visible orientation to the telling itself, by turning his gaze to his left and looking at how the person seated next to him is ladling soup from the pot in the centre of the table to his bowl (2 in Figure 2.3).

Don thus immediately adapts to the changes in Ann's talk so as to maintain the appropriateness of his participation for the structure of the talk currently in progress. Through such changes in his visible participation Don demonstrates his understanding of how the talk in progress is consequential for his own actions. As part of this process he visibly differentiates alternative kinds of units within Ann's story in terms of the participation frameworks each invokes.

---

[2] For more detailed analysis of how Don's embodied actions are organised with reference to interactive field invoked by Ann's telling, see Goodwin (1984).

The sequence as a whole provides strong evidence for the possibility that hearers are: first, visibly co-participating in the organisation of the talk in progress; second, engaged in detailed analysis of the unfolding structure of that talk; and third, using that analysis to make projections relevant to their own participation in it.

2.4 An alternative view of participation

Goffman's decomposition of the speaker provided us a model of an entity with complex internal structure, a multi-faceted player using language to perform interesting, laminated actions within a rich cognitive environment. However, no other participant category is decomposed in an equivalent fashion. Instead the complexity of participation status emerges through the accumulation of categories for types of recipients. Each of these categories, in marked contrast to the decomposition of the speaker, is treated as structurally simple and undifferentiated. Moreover, the complex ties between linguistic structure and forms of participation so prominent in Goffman's analysis of the speaker are entirely absent from his typology of possible recipients. Indeed, if one conceptualises the production of language as essentially a speaker's activity, this might seem entirely natural.

The model offered in Footing constitutes the point of departure for one very important approach to the study of participation. The categories for types of participants offered by Goffman were considerably expanded by Levinson (1988). Hanks (1990) then opposed open-ended category proliferation by noting how a range of different types of speakers and hearers could be logically accounted for as the outcome of more simple and general underlying practices, such as systematic embedding of one participation framework within another (as happens, for example, in quotation and other forms of reported speech).

Don's actions suggest an alternative framework for the study of participation, one that does not follow Footing by focusing on the construction of typologies categorising in a static fashion structurally different types of participants. Instead participation can be analysed as a temporally unfolding process through which separate parties demonstrate to each other their ongoing understanding of the events they are engaged in by building actions that contribute

to the further progression of these very same events. Thus Don participates in the interactively sustained, multi-party interactive field that constitutes Ann's telling by organising his body with reference to how he is positioned within that field, while modifying his embodied displays as the emerging structure of Ann's talk makes relevant different kinds of participation alignments.

Shifting analysis – from the elaboration of typologies for participants, to study of the activities that parties must perform in order to participate appropriately in the events of the moment by building relevant action – has a number of important consequences. First, such a framework recovers the cognitive life of the hearer by focusing investigation on the analysis he or she must perform in order to co-construct action through differentiated participation in the talk of the moment. Second, a crucial component of the analysis hearers are engaged in focuses on distinguishing alternative units with the stream of speech in terms of the different possibilities for participation that each makes relevant. The actions of hearers thus shed important light on a key theoretical issue in the analysis of language structure; specifically the question of how participants parse the stream of speech into relevant units. Indeed, that task becomes visible here as a practical problem for participants, a constitutive feature of the forms of social organisation they build through talk, rather than simply a theoretical issue for analysts or transcribers. Third, important properties of this unit structure are provided by the sequential organisation of talk, including what different kinds of units count as forms of multi-party, multi-modal interaction. Fourth, this process also has a temporal dimension as, first, different kinds of units (with different participation possibilities) unfold through time and, second, hearers make projections about upcoming units in order to accomplish relevant simultaneous action.

2.5 Logocentrism

This view of participation has a number of consequences for how stories, reported speech and talk in interaction more generally are conceptualised. Most analysis of both reported speech and stories focuses exclusively on talk. However, the data examined here reveal that a story in face-to-face interaction is a multi-modal, multi-party field of activity. In addition to phenomena in the stream of speech,

other kinds of signs displayed through, for example, the visible organisation of the body are also relevant.

Differences between kinds of sign systems, and their potential for being captured in writing, privilege one participant, the speaker, while obscuring all others. Because of the division of labour between speaker and hearer(s), speaker(s) produce most talk. Hearers' concurrent talk, though frequently informative about a recipient's analysis of what is being said and his or her stance toward it (Jefferson, 1973, 1983, 1984a), typically lacks the semantic and syntactic complexity of the speaker's talk. Characteristically, recipient actions, such as continuers (Schegloff, 1982) and assessments (Goodwin, 1986b), take the form of brief one- or two-syllable phrases. Indeed, if hearers were to provide substantive talk within another speaker's turn, extended overlap would occur. It might be argued that the actions of hearers can be recovered by focusing on later turns where former hearers who have now become speakers can display analysis of the talk they heard earlier. However, there is no reason whatsoever to treat such subsequent action as equivalent to their concurrent analysis and co-participation in the utterance while it was in progress (Goodwin and Goodwin, 1987). There are crucial differences between a hearer and a subsequent speaker. For example, though Don participated in the laughter while Ann reported what he said, in subsequent talk he tried to counter and minimise her interpretation of his talk, e.g. 'But I said it so innocuously y'know'. He responds to what Ann quotes him as saying in entirely different ways during her talk than he does later as a subsequent speaker.

The upshot of all of this is that focusing analysis exclusively on talk treats the speaker as the primary – indeed, on occasion, the sole – actor relevant to the construction of an utterance such as a story, while obscuring, or rendering completely invisible, the simultaneous actions of the hearer. Thus Don appears only as a cited figure in the transcript of Ann's story in Figure 2.1, not as a present actor. If data for study of the story consisted only of the talk transcribed there, none of his actions that were investigated above would be accessible to investigation. These include not only phenomena centred on the visible body, such as participation displays, but also his moment-by-moment analysis of the unfolding structure of the talk in progress.

Some argue that the genuine analytic problems raised by this situation can be avoided by using data such as phone calls where the participants' only access to each other is through talk. While this may be true for such limited cases, it seems clear that the primordial site for talk-in-interaction, and human interaction in general, is not one that is restricted to sound, but instead a situation in which participants are building relevant action together through talk while attending to each other as fully embodied actors, and frequently to relevant structure in their environment, the larger activities they are engaged in, etc. (Goodwin, 2000a, 2003a).

There are powerful reasons for such logocentricism. For thousands of years human beings have been grappling with the issues raised by the task of capturing significant structure in the stream of speech in writing. Writing systems, and the insights and methodological tools they have provided for the analysis of linguistic and phonetic structure, the creation of precise records that can endure in time and be transported from place to place, etc. are major accomplishments that provide a crucial infrastructure for much of research into language structure, verbal genres and more recently talk-in-interaction. However, despite efforts in kinesics, gesture studies, Labanotation, etc., there are no systems comparable to writing for the rapid and precise annotation of the other embodied modalities that contribute the organisation of face-to-face interaction. The problems posed are not simply methodological but, more importantly, require discovery of the crucial distinctions that participants attend to for the organisation of action through embodied interaction. Moreover, such perceptual bias toward what is being said, with other modalities receding into a more amorphous ground, seems to reflect in part the way participants themselves focus their explicit attention on the talk in progress. Thus, if asked what happened in an encounter, participants typically report what was said, not the work of constructing the embodied frameworks of mutual orientation that made interactive talk possible in the first place (Goodwin, 1981; Kendon, 1990).

Contemporary video and computer technology makes it possible to repeatedly examine the bodies as well as the talk of participants in interaction, and thus to move analytically beyond logocentricism. And indeed some evidence suggests that neither talk, nor

language itself, are self-contained systems, but instead function within a larger ecology of sign systems (Goodwin, 2000a).

2.6 Mutual reflexivity

Within interaction participants treat their co-participants as reflexive actors. They expect each other to take into account for the organisation of subsequent action the projective frameworks provided by both talk and visible embodied displays. Don's actions provide one example of how hearers demonstrate ongoing analysis of emerging talk by building actions that make visible appropriate participation in it. Data beyond that included in this chapter demonstrate that speakers take the actions of hearers into account in ways that have strong consequences for the future trajectory of stories and other units of talk (C. Goodwin, 1981, 2002; M. H. Goodwin, 1980). A story is constructed, not by the speaker alone, but instead through the coordinated actions of different kinds of participants. Moment by moment each party must take into account: 1) the emerging structure of the activities in progress; 2) what precisely other parties are doing; and 3) the implications that this has for the trajectory of future action. Major resources for this process include the signs present in the structure of the talk and the displays being made by the bodies of the participants.

It is precisely this organisation of mutual reflexivity that is missing from Goffman's models of Footing and participation.

In Footing, instead of collaborating together to build talk, speakers and hearers inhabit separate worlds, with quite different frameworks being used for the analysis of each. One reason for this would seem to lie in the way in which speakers and hearers are described in quite separate sections of Footing. Such a rhetorical arrangement makes it difficult (perhaps impossible) to build a model in which utterances are constructed through processes of interaction in which different kinds of participants are building action in concert with each other. In Footing, building utterances is exclusively the work of speakers, who are thus endowed with all relevant cognitive structure.

It is most ironic that one of Goffman's most influential legacies is a powerful analytic framework that focuses on the talk of the speaker in isolation from the simultaneous actions of the hearer.
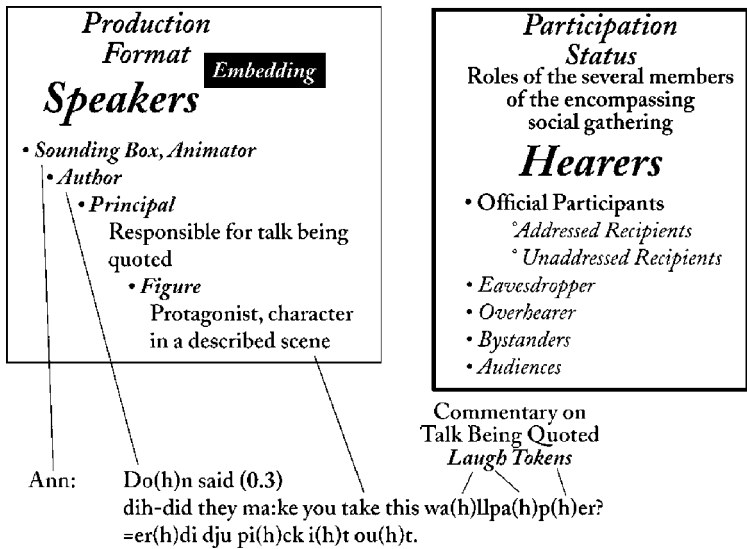
Figure 2.4. Participation Status

Goffman introduced into the study of human interaction terms such as 'mutual monitoring'. His whole career was devoted to intensive studies of processes of human interaction, and in work such as that on strategic interaction and mutual monitoring he showed deep concern for the ways in which participants were analysing each other to manipulate meaning and action. Why then in Footing did he develop an analytic framework that treats the talk of the speaker as an isolated, self-contained system?

   Candy Goodwin and I were studying with Goffman while he was working on Footing. One day he urged us to read Volosinov (1973), which had only been recently published in English. Though I have no other evidence[3] whatsoever, I suspect that Goffman's thinking in Footing was influenced in part by his reading of Volosinov. His decomposition of the speaker, noted above, provides powerful tools for working with some of Volosinov's insights about reported speech. Most crucially, both Footing and Volosinov share a vision of how dialogic interaction can be embedded within the

---

[3] Volosinov is not listed in the references to Footing and is cited only once in Goffman's *Forms of Talk* (1981).

talk of a single speaker that has deeply shaped how subsequent scholars conceptualised both reported speech and participation status. Sapir (1968) notes that a word with its associated concept 'is not only a key; it may also be a fetter.' The genuine insights of both Volosinov and Goffman created a key through which important phenomena could be investigated in novel and important ways, while simultaneously constraining such study in hidden but powerful ways.

2.7 Volosinov's conception of dialogue

Marxism and the Philosophy of Language was published in the Soviet Union in 1929. Its author, V. N. Volosinov, was a member of the circle of scholars who clustered around Mikhail Bakhtin. Indeed it has been strongly argued that Bakhtin, not Volosinov, was the work's author. I take no position on that issue. Regardless of who might be named as author, the volume and its arguments are deeply tied to the analytic frameworks associated with Bakhtin, and indeed might be considered one of the canonical Bakhtinian texts. Here I wish to raise questions about the precise way in which Volosinov formulated the dialogic nature of language, essentially arguing that both the participation of the hearer and multi-party talk are rendered invisible in his conception of reported speech. However, in doing this I want to emphasise my deep appreciation for, and agreement with, the insights and perspectives so brilliantly enunciated in this book. My disagreement is not an attack on Volosinov, but a dialogue with him, and an attempt to expand the perspectives he offers by looking seriously at how he formulated crucial analytic phenomena. Moreover, I recognise only too well how subsequent scholars have used the work of Bakhtin and Volosinov to probe the interrelationship between language, ideology, stance, social positioning, voice and culture to develop powerful, original and important new ways of thinking about, and acting within, the lived social world.

  When, at Goffman's urging, I first read Volosinov (1973) in the early 1970s, I was amazed that the book could have been written in the 1920s. It made arguments about the social and interactive organisation of language that prefigured my own interest in the

participation of the hearer in the construction of utterances, but which I had not seen anywhere else. For example:

word is a two-sided act. It is determined equally by *whose* word it is and for *whom* it is meant. As word, it is precisely *the product of the reciprocal relationship between speaker and listener, addresser and addressee*. Each and every word expresses the 'one' in relation to the 'other'. I give myself verbal shape from another's point of view ... A word is territory shared by both addresser and addressee, by the speaker and his interlocutor ... constituting, as it were, the border zone between [the speaker] and his addressee.   (Volosinov, 1973: italics in original)

All of this – for example, proposing that the speaker shaped himself or herself from the point of view of the addressee – resonated directly with what I was finding in my own video analysis of utterance construction at that time, e.g. speaker's reconstructing their emerging utterances, and displays of knowledge and certainty, as gaze was moved from one type of addressee to another (C. Goodwin, 1979, 1981, 1987).

Moreover, in a critique that remained relevant to the formal linguistics of the late twentieth century (though originally directed against what Volosinov termed the Abstract Objectivism of Saussure), and which seemed to anticipate fields such as conversation analysis, Volosinov (1973: 117) argued that the primary locus for language was not the isolated monologic utterance or sentence, but instead 'the interaction of at least two utterances – in a word dialogue.'

However, in a series of subtle but most crucial moves, Volosinov lodged the problems posed by the study of dialogue, not in sequences of multi-party talk, i.e. what one would consider to be the most natural, straightforward notion of dialogue, but instead within the consciousness of the individual speaker: 'How…is another speaker's speech received? What is the mode of existence of another's utterance in the actual, inner-speech consciousness of the recipient? How is it manipulated there?' (Volosinov, 1973: 117).

How can the interplay between the utterances of separate individuals be analysed within the mind, and talk, of a single speaker? Volosinov's answer to this problem was original and important. He argued that reported speech, in which the current speaker incorporates the talk of another into the current utterance, provides a place where the reception and transformation of another's talk can be

systematically investigated. Ann's report of what Don said in lines 13–16 of Figure 2.1 provides a typical example, and indeed, as is demonstrated by Goffman's decomposition of the speaker in Footing, and much work by a host of other scholars, what Volosinov draws our attention to here is a rich and important phenomenon. Note, however, that dialogue is now transformed into something that can best be investigated within the talk and consciousness of a single speaker, albeit one quoting the talk of another.

Indeed Volosinov argues explicitly that what he is proposing as the key to the dialogical organisation of language, the process of 'words reacting on words' found in reported speech,

is distinctly and fundamentally different from dialogue. In dialogue, the lines of the individual participants are grammatically disconnected; they are not integrated into one unified context. Indeed, how could they be? *There are no syntactic forms with which to build a unity of dialogue.*   (Volosinov, 1973: 116, italics in original)

Volosinov thus comes to the rather paradoxical conclusion that dialogue, multi-party sequences of talk, does not provide appropriate data for study of the dialogic organisation of language, 'the *active relation* of one message to another' (1973: 116, italics in original). The evidence he offers to support this, lack of syntactic ties linking the contributions of separate speakers, is simply not true as work such as that of Sacks (1992a: 144–147), on collaborative utterances and tying techniques (1992a: 716–721), clearly demonstrates.[4]

However, by making this argument Volosinov was able to define dialogue so that syntactically complex texts, rather than talk-in-interaction, constitute the primary site for the dialogic organisation of language and culture. And, indeed, within the larger politics of research fields competing to occupy the most relevant site for the analysis of human language, cognition and social action (and by offering his work as a critique of, and alternative to, Saussure Volosinov was explicitly engaging in such politics), defining *the*

---

[4] Note in this connection the significant body on research on reported speech, and the diverse grammatical practices of different languages, which have crucial differences in their sequential placement, which takes talk-in-interaction and research in conversation analysis as its point of departure (e.g. Hayashi, 1997; Holt, 1996; the introduction to this volume).

perspicuous site for research in this fashion, i.e. as reported speech, could be most useful to a literary critic such as Bakhtin. This position has strong implications for methodology and data as well. Since all relevant phenomena are located in the kinds of language structure that writing captures, it is not necessary for the analyst to look beyond the printed text, for example to investigate the bodies of either the speaker or active (though silent) participants such as hearers, or multi-party action. A radical argument about the importance of both the hearer and dialogue is thus subtly domesticated so that it fits comfortably within the boundaries of the traditional textual artefacts that thus continue to define where language will be analysed.

Volosinov and Footing thus share a number of key assumptions. In both, syntactically complex language in which the current speaker reports in some fashion the talk of another is used to build very interesting analysis of how single utterances constitute a site where the voices of multiple speakers dynamically interact with each other. However, by virtue of the way in which the multiple voices that constitute the dialogue being analysed are embedded in a single utterance, there is no need to investigate actual multi-party sequences of talk or phenomena outside the stream of speech. Parties other than the speaker are thus excluded from analysis. The crucial mutual reflexivity of speakers and hearers is lost. It becomes impossible to investigate how utterances are built through processes of interaction that include the participants' ongoing analysis of each other. In essence the world being analysed is lodged within a single speaker's speech.

2.8 Multi-party speakers and participation

While the insights of both Volosinov and Footing lead to very interesting analysis of a rich and important class of utterances, serious problems arise if models such as this are used for the general analysis of the dialogic organisation of language and culture, of stance and footing, and of the practices used by speakers to incorporate another's talk into a current utterance. For example, both Volosinov and Footing require, as a point of departure for the kinds of analysis they propose, utterances that have rich syntax, e.g. clauses in which the talk of another that is being reported is

embedded within a larger utterance by the current speaker. The necessity of rich syntax not only excludes important activities – such as many greetings which, at least in English, are frequently done with one to two word utterances (e.g. 'Hi') – but also certain kinds of speakers. Because of a severe stroke Chil, whose actions we will now investigate, was able to say only three words: 'Yes', 'No' and 'And'. It is impossible for him to produce the syntax that Goffman's Production Format and Volosinov's reported speech seem to require (i.e. he can't produce a sentence such as 'John said X'). Someone such as Chil appears to fall beyond the pale of what counts as the competent speaker required for their analysis.

Chil in fact acts as a powerful speaker in interaction, and moreover one who is able to include the talk of others in his utterances. Describing how he does this requires a model of the speaker that moves beyond the individual. The sequence in Figure 2.5 provides an example. Chil's son Chuck and daughter-in-law Candy are talking with him about the amount of snow the winter has brought to the New York area where Chil lives. After Candy notes that not much has fallen 'this year' (which Chil strongly agrees with in talk omitted from the transcript), in line 11 she proposes that such a situation contrasts markedly[5] with the amount that fell 'last year'. Initially, with his '°yeah-' Chil seems to agree (in the interaction during the omitted talk Chil was strongly agreeing with what Candy was saying, and thus might have grounds to expect and act as though that process would continue here). However, he ends his agreement with a cut-off (thus visibly interrupting and correcting his initial agreement) and moves to strong, vivid disagreement in line 13. Candy immediately turns to him and changes her 'last year' to 'the year before last'. Before she finishes, Chil (line 15) affirms the correctness of her revised version.

Despite his severely impoverished language Chil is able to make a move in the conversation that is both intricate and precise: unlike what Candy initially proposed in line 10, it was not 'last year' but 'the year before last' when there was a lot of snow. Chil says this by getting someone else to produce just the words that he needs. The

---

[5] This contrast is signalled strongly by both the contrastive 'But', which begins line 11, and the displays of heightened affect and stance that follow.

The participants have been discussing snow in the area where Chil lives

1   Candy: You haven't had **that** much this year have you.

… ((8 lines omitted))

10   Candy: But **la** [**st** year. Whoo!
11   Chuck: [°mm
12   [In the l [ast year-
13   Chil: [°yeah- [**No Na. No:.**
14   Candy: er the year before la [st. ─
15   Chil: [Yes.

Figure 2.5. Extract (2)

talk in line 14 is semantically and syntactically far beyond anything that Chil could say on his own.

Though not only spoken, but constructed by Candy, it would be clearly wrong to treat line 14 as a statement by her. First, just a moment earlier, in lines 10 and 12 she voiced the position that is being contradicted here. Second, as indicated by Chil's agreement in line 15, Candy is offering her revision as something to be accepted or rejected by Chil, not as a statement that is epistemically her own. Line 14 thus seems to require a deconstruction of the speaker of the type called for by Goffman in Footing, with Candy in some sense being an animator, or 'sounding box', for a position being voiced by Chil. However, the analytic framework offered in Footing does not accurately capture what is occurring here. Though Candy is in some important sense acting as an Animator for Chil, he is not a cited figure in her talk, and no quotation is occurring. Intuitively the notion that Chil is in some sense the Author of line 14, and its Principal, seems plausible (what is said here would not have been spoken without his intervention, and he is treated as the ultimate judge of its correctness). However, how could someone completely unable to produce either the semantics or the syntax of line 14 be identified as its Author?

Clarifying such issues requires a closer look at the interactive practices used to construct the talk that is occurring here. Chil's intervention in line 13 is an instance of what Schegloff et al. (1977)

describe as Other Initiated Repair. With his *'No No. No:.'* Chil forcefully indicates that there is something wrong with what Candy, the prior and still current speaker, has just said. She can re-examine her talk to try and locate what needs repair, and indeed here that process seems straightforward. In response to Chil's move, Candy changes '*last* year', the crucial formulation in the talk Chil is objecting to, to an alternative 'the year before last'.

   Such practices for the organisation of repair, which are pervasive not only in Chil's interaction, but in the talk of fully fluent speakers as well (Schegloff et al., 1977), have crucial consequences for both Chil's ability to function as a speaker in interaction, and for probing the analytic models offered by Goffman and Volosinov. First, through the way in which Chil's instances of 'Yes' and 'No' are tied to specific bits of talk produced by others (e.g. what Candy has just said), they have a strong indexical component which allows him to use as a resource detailed structure in the talk of others, and in some sense incorporate that talk into his own, linguistically impoverished utterances. Thus in line 13 he is heard to be objecting not to life in general, but to precisely what Candy said in line 12, and to be agreeing with what she said in line 14. Second, such expansion of the linguistic resources available to Chil is built upon the way in which his individual utterances are embedded within sequences of dialogue with others, or more generally the sequential organisation of interaction. However, this notion of dialogue, as multi-party sequences of talk, was precisely what Volosinov (1973: 116) worked to exclude from his formulation of the dialogic organisation of language. Nonetheless, Chil's actions here provide a clear demonstration of the larger Bakhtinian argument that speakers talk by 'renting' and reusing the words of others.

   Third, what happens here requires a deconstruction of the speaker that is relevant to, but different from, that offered by Goffman in Footing. What Chil says with his 'No' in line 13 indexically incorporates what Candy said in line 11, though Chil does not, and cannot, quote what she said there. Instead of the structurally rich single utterance offered in Goffman's model of multiple voices laminated within the complex talk of a single speaker, here we find a single lexical item, a simple 'No', that encompasses talk produced in multiple turns (e.g. both lines 11

and 13) by separate actors (Candy and Chil). Unlike Ann's story in Figure 2.1, Chil's talk cannot be understood or analysed in isolation. Its comprehension requires inclusion of the utterances of others that Chil is visibly tying to.

Rather than being located within a single individual, the speaker here is distributed across multiple bodies and is lodged within a sequence of utterances. Chil's competence to manipulate in detail the structure of emerging talk by objecting to what has just been said – that is, to act in interaction – constitutes him as a crucial Author of Candy's revision in line 14, despite his inability to produce the language that occurs there. Though not reporting the speech of another Candy speaks for Chil in line 14, and locates him as the Principal for what is being said there. All of this requires a model of the speaker that takes as its central point of departure not the competence to quote the talk of another (though being able to incorporate, tie to and reuse another's talk is absolutely central), but instead the ability to produce consequential action within sequences of interaction.

Fourth, the action occurring here, and the differentiated roles parties are occupying within it, are constituted not only through talk, but also through participation as a dynamically unfolding process. As line 13 begins Candy has turned away from Chil to gaze at Chuck. Chil's talk in line 13 pulls Candy's gaze back to him (her eyes move from Chuck to Chil over the last of his three uses of 'No'). Such securing the gaze of an addressee is similar to the way in which fluent speakers use phenomena such as restarts to obtain the gaze of a hearer before proceeding with a substantive utterance (C. Goodwin, 1980, 1981).

In this case, however, it is the addressee, Candy, rather than Chil, the party who solicited gaze, who produces the talk that follows. Nonetheless, through the way in which he organises his body Chil displays that he is acting as something more than a recipient of Candy's talk, and instead sharing the role of its speaker. Typically gestures are produced by speakers. Indeed the work of McNeill (1992) argues strongly that an utterance and the gesture accompanying it are integrated components of a single underlying process. Line 14 is accompanied by gesture. However, it is performed not by the person speaking, Candy, but instead by Chil (see Figure 2.6).
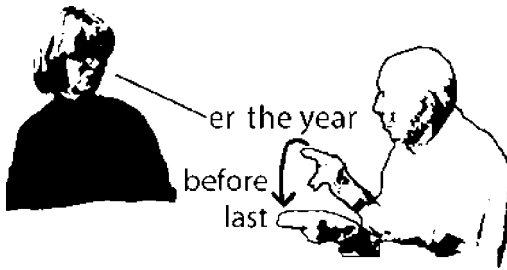
Figure 2.6. Extract (2), Detail

Chil thus participates in Candy's utterance by performing an action usually reserved for speakers, and in so doing visibly displays that he is in some way acting as something more than a hearer. The gesture seems to provide a visual version of what Candy was saying, and specifically to illustrate the notion that one unit (which can be understood as a 'year' through the way in which the gesture is temporally bound to Candy's talk) has another that precedes it. As Candy says 'the year', Chil raises his hand toward her with two fingers extended. Then, as she says 'before last', he moves his gesturing hand down and to the left (see Figure 2.6). Even if this interpretation of the gesture must remain speculative (for participants as well as the analyst) because of Chil's inability to fully explicate it with talk of his own, the gesture is precisely coordinated with the emerging structure of Candy's talk, and vividly demonstrates Chil's participation in the field of action being organised through that talk. Note once again that Participation is being investigated here, not as static categories constructed by the analyst (addressee, speaker, hearer, etc.), but instead as forms of temporally unfolding, interactively organised action through which participants demonstrate with precision (as Chil does here word by word as Candy's talk emerges) their understanding of the events in progress by building action that helps to produce these very same events.

The following provides another example of how the position of speaker is distributed across multiple bodies, and lodged within the sequential organisation of dialogue. Here Chil's daughter Pat and son Chuck are planning a shopping expedition. Once again Chil intercepts a speaker's talk with a strong 'No' (lines 6–7 in Figure 2.7). Pat is talking about the problem of finding socks that

```
 1   Pat:        So  you always need the black socks:
 2   Chuck:      °umm.
 3                  (1.7)
 4   Pat:        It's too bad we found those there too
 5               at the store that went outa busi[ness.
 6   Chil:  →                                    [Na-dee No:=
 7               =dih dih dih-
 8                  (0.5)
 9   Pat:   →    Yeah=you went to Bergenfield.
10   Chil:       Ye:s.
```

Figure 2.7. Extract (3)

fit over Chil's leg brace, since the store where she bought them last went out of business.

What occurs here has is structurally similar to the 'year before last' sequence examined in Figures 2.5 and 2.6. After Chil uses a 'No' to challenge something in the current talk, that speaker produces a revision, which Chil affirms. Once again Chil is operating on the emerging sequential structure of the local dialogue to lead another speaker to produce the words he needs. However, while Candy in Figure 2.5 could locate the revision needed through a rather direct transformation of the talk then in progress (changing 'last year' to 'the year before last'), the resources that Pat uses to construct her revision are not visible in the transcript. How is she able to find a completely different store and, moreover, locate it geographically? When a visual record of the exchange is examined, we find that in addition to talk Chil produces a vivid pointing gesture as he objects to what Pat is saying. Pat treats this as indicating a particular place in their local neighbourhood, a store in an adjacent town in the direction Chil is pointing (see Figure 2.8).

Chil constructs his action in lines 6–7 by using simultaneously a number of quite different meaning-making practices that mutually elaborate each other. First, as was seen in the 'last year' example in Figures 2.6 and 2.7, by precisely placing his 'No' (again overlapping the statement being challenged), Chil is able to use what is being said by another speaker as the indexical point of departure for his own action. His hearers can use that talk to locate something quite specific about what Chil is trying to indicate (e.g. that his action concerns something about the place where the socks were

Multiple Contextual Fields

| LocalTalk | No |
| Pointing Gesture | Structured Space of the Local Lifeworld |

```
1   Pat:     So you always need the black socks:
2   Chuck:   °umm.
3            (1.7)
4   Pat:     It's too bad we found those there too
5            at the store that went outa busi  [ness.
6   Chil:                                      [ Na-°dee No:=
7            =dih dih dih-
8            (0.5)
9   Pat:     Yeah.=You went to Bergenfield.
10  Chil:    Ye:s
```

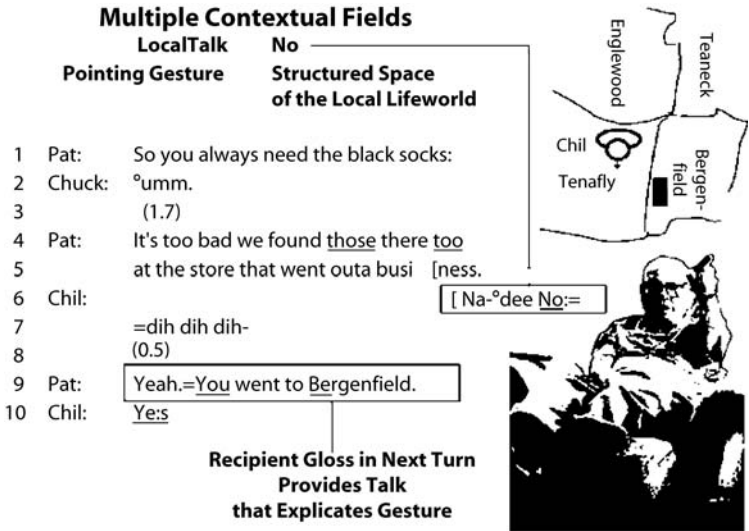**Recipient Gloss in Next Turn Provides Talk that Explicates Gesture**

Figure 2.8. Multiple Complex Fields

bought). Nonetheless, as this example amply demonstrates, such indexical framing is not in any way adequate to specify precisely what Chil is attempting to say (e.g. in lines 4–5 there is no indication of a store in Bergenfield). However, Chil complements his 'No' with a second action, his pointing gesture. In isolation such a point could be quite difficult for an addressee to interpret. Even if one were to assume that something in the environment was being indicated, the line created by Chil's finger extends indefinitely. Is he pointing toward something in the room in front of them or, as in this case, a place that is actually miles away?[6] However, by using the co-occurring talk a hearer can gain crucial information about what the point might be doing (e.g. indicating where the socks being discussed were bought). Simultaneously, the point constrains the rather open-ended indexical field provided by the prior talk by indicating an alternative to what was just said. By themselves both

[6] The task of locating what is being pointed at is not simply a theoretical problem, but a genuine practical one for those who interact with Chil. On occasion they searched for something in the room when in fact he was pointing well beyond the wall, and vice versa (Goodwin, 2000b, 2003b).

the talk and the pointing gesture are partial and incomplete. However, when each is used to elaborate and make sense out of the other, a whole that is greater than the sum of its parts is created (see also Wilkinson et al., 2003).

The ability to properly see and use Chil's pointing gesture requires knowledge of the structure of the environment being invoked through the gesture. As someone who regularly acts and moves within Chil's local neighbourhood, Pat can be expected to recognise such a structure. A stranger would not. Chil's action thus encompasses a number of quite different semiotic fields (C. Goodwin, 2000a) including his own talk, the talk of another speaker that Chil's 'No' is tied to, his gesturing arm, and the spatial organisation of his surroundings. Though built through general practices (negation, pointing, etc.), Chil's action is situated in, and reflexively invokes, a local environment that is shaped by both the emerging sequential structure of the talk in progress, and the detailed organisation of the lifeworld that he and his interlocutors inhabit together.

One pervasive model of how human beings communicate conceptualises the addressee/hearer as an entity that simply decodes the linguistic and other signs that make up an utterance, and through this process recovers what the speaker is saying. Such a model is clearly inadequate for what occurs here. To figure out what Chil is trying to say or indicate, Pat must go well beyond what can actually be found in either her talk or Chil's pointing gesture. Rather than in and of themselves encoding a proposition, the signs Chil produces presuppose a hearer who will use them as a point of departure for complex, contingent inferential work. Chil requires a cognitively complex hearer who collaborates with him in establishing public meaning through participation in ongoing courses of action.

The participation structures through which Chil is constituted as a speaker are not lodged within his utterance alone, but instead distributed across multiple utterances and actors. In line 9 Pat responds to Chil's intervention by providing a gloss of what she takes him to be saying: 'You went to Bergenfield'. Chil affirms the correctness of this with his 'Yes' in line 10. If this action is analysed using only the printed transcript as a guide it might seem to constitute a simple agreement with what Pat said in line 9. However, when a visual record of the interaction is examined Chil can be seen to move his gaze from Pat to Chuck as he speaks this word.

```
 4   Pat:     it's too bad we found those there too
 5            at the store that went outa busi [ness.
 6   Chil:                                     [Na-°dee No:=
 7            =dih dih dih-
 8                 (0.5)
 9   Pat:     Yeah.=You went to Bergenfield.
10   Chil     Ye:s.
11   Chuck:   [°uh huh. uh huh.((Nods to Chil))
12   Pat:     [To the men's shop.
```
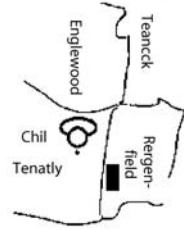
Figure 2.9. Extract (3), Detail

Chuck, who is visiting, lives across the continent. He is thus not aware of many recent events in Chil's life, including the store in Bergenfield that Pat has just recognised (though Chuck, who grew up in this town, is familiar with its local geography). With his gaze shift (and the precise way in which Chil speaks 'Yes', which is beyond my abilities to indicate appropriately on the printed page), Chil visibly assumes the position of someone who is telling Chuck about this store. He thus acts as not only the author, but also the speaker and teller, of this news. He has of course excellent grounds for claiming this position. A moment earlier, in lines 4–5, Pat said something quite different, and it was only Chil's intervention that led her to produce the talk he is now affirming. Within the single syllable of line 10 Chil builds different kinds of action for structurally different kinds of recipients: first, a confirmation of what Pat (someone who knows about the event at issue and now recognises it) see Figure 2.9, has just said and, second, a report about that event to Chuck, an unknowing recipient.[7]

[7] For other analysis of how utterances can be built to include both knowing and unknowing recipient, see Goodwin (1979, 1981, 1987).

Both Volosinov's analysis of reported speech and Goffman's deconstruction of the speaker focused on the isolated utterance of a single individual who was able to constitute a laminated set of structurally different kinds of participants by using complex syntax to quote the talk of another. By way of contrast the analytic frameworks necessary to describe Chil's speakership in line 18 must move beyond him as an isolated actor to encompass the talk and actions of others, which he indexically incorporates into his single-syllable utterance in line 10. Moreover, grasping his action requires attending not only to structure in the stream of speech but also to his visible body, and relevant structure in the surround. Chil's speakership is distributed across multiple utterances produced by different actors (e.g. Pat's talk in both lines 13 and 17 is a central part of what is being reported through his 'Yes'), and encompasses non-linguistic structure provided by both his visible body and the semiotic organisation of the environment around him. His talk is thoroughly dialogic. However, analysis of how it incorporates the talk of others in its structure requires moving beyond the models for reported speech and the speaker provided by Volosinov and Goffman.

2.9 Conclusion

In *Marxism and the Philosophy of Language*, Volosinov offered, in opposition to what he termed the 'abstract objectivism' of Saussure, a powerful vision of language as something thoroughly lodged within human dialogue, e.g. 'A word is territory shared by both addresser and addressee, by the speaker and his interlocutor' (Volosinov, 1973). However, at the same time Volosinov was careful to situate the 'dialogic' play of multiple voices, not in sequences of talk by different participants (e.g. the most straightforward notion of dialogue), but in reported speech, talk in which a speaker incorporates in some fashion the talk of another. Volosinov's insights into reported speech constituted the point of departure for a large, diverse and important body of research by subsequent scholars that shed important new light on a host of issues implicated in the organisation of language structure, stance, metalanguage, story organisation and the sedimentation of historically shaped social phenomena in the details of language structure.

However, by restricting investigation to the interplay of voices within the talk and consciousness of a single speaker, the phenomenon of reported speech renders inaccessible to analysis a host of social practices that are crucial to the dialogic organisation of language and action that Volosinov discusses with such insight in the first half of his book. This chapter has attempted to demonstrate the importance of the task posed by Volosinov of working to disentangle the different voices within a single strip of talk, while at the same time expanding such analysis to incorporate genuine multi-party interaction within the organisation of the utterance, and to explore, as an alternative to quotation, sequential practices for assimilating another's talk into a current utterance.

Goffman's deconstruction of the speaker in Footing provides powerful analytic tools for specifying the different kinds of speakers who can co-exist with a strip of reported speech. Indeed, Footing was the most influential article of Goffman's later career. It constituted one point of departure (with Volosinov and Bakhtin) for important lines of research on phenomena such as reported speech, participation, and the dialogic organisation of language and culture. The frameworks developed by Goffman in Footing continue to shape how crucial phenomena that are central to the organisation of language as a social process are analysed. However, as analysis of the story in Figures 2.1–2.4 demonstrated, parties whose talk is being quoted can exist not only as cited figures within the talk of the current speaker, but also as actual present participants. Through the way in which he visibly organized his body the principal character whose talk was being quoted helped to co-construct the multi-party, dialogic field constituted through the reporting of his speech.

Coming to terms with this analytically requires expanding the framework provided by Volosinov in at least two ways. First, rather than being constituted entirely within the stream of speech, the actions constructed within an utterance can incorporate other semiotic modalities as well, such as visible, embodied displays. One crucial consequence of focusing on this larger ecology of sign systems (C. Goodwin, 2000a) which can encompass speech without being restricted to talk, is that the relevant actions of parties who don't speak, such as hearers of various types, can be taken into account analytically. Second, participation is central to the organisation of

this multi-party interactive field. Footing actually contains a framework for the analysis of participation that has had enormous influence. However, with respect to the phenomena being investigated here, the framework Goffman offered has major problems. The model of participation in Footing consists of a typology of participant categories that are not linked to the model of the speaker presented in a different section of the article. It thus cannot provide the analytic resources necessary to describe how participants build utterances and action by taking each other into account within an unfolding process of interaction as talk unfolds, i.e. the essential mutual reflexivity of speaker and hearer(s). In this chapter an alternative view of participation is offered. Participants demonstrate their understanding of what each other is doing and the events they are engaged in together by building both vocal and non-vocal actions that help to further constitute those very same events. One consequence of this is a multi-party, interactively sustained, embodied field within which utterances are collaboratively shaped as meaningful, locally relevant action. Within such a framework the speaker is no longer positioned as the locus of all semiotic activity, and the cognitive life of the hearer – including his or her analysis of the details of emerging language structure – is recovered.

Both Volosinov's view of reported speech, and Goffman's speaker, require an actor capable of producing rich, complex language structure. To probe such assumptions, the utterances of a man able to say only three words ('Yes', 'No' and 'And') because of a stroke that left him with severe aphasia were examined. Despite his restricted vocabulary, and almost complete lack of syntax, this man was not only able to act as a powerful speaker in conversation, but also to incorporate the talk of others in his own catastrophically impoverished utterances. Clearly he could not do this with grammatical structure signalling that another party's talk was being quoted. However, he could indexically tie his own limited talk to rich structure in the talk of others and, moreover, position himself, and not that other speaker, as the ultimate author and principal of what was being said. Though not reported speech, this process would seem to constitute a clear example of action built through 'an *active relation* of one message to another' (Volosinov, 1973: 116), and to have strong relevance to the dialogic organisation that Volosinov was trying investigate through reported speech.

However, analysing this process requires moving beyond the domain of scrutiny of both Volosinov and Goffman, the talk of a single speaker, to focus instead on multi-party sequences of talk. Moreover, the interactive construction of meaning and action that occurs here requires a reflexive, cognitively complex hearer and frequently orientation to semiotic structure that extends beyond the stream of speech. Goffman deconstructed the single speaker, and the isolated utterance of that speaker, into multiple entities. Here, by way of contrast, we find a speaker who is distributed across different participants and turns (e.g. the aphasic man prompts, incorporates and claims authorship for things said by others).

Both Volosinov and Goffman offer us powerful insights into the intrinsically dialogic organisation of human language. However, both restrict their analysis to the talk and consciousness of a single speaker. In this chapter I have attempted to demonstrate that the phenomena they draw our attention to can be investigated more richly by focusing instead on how separate parties build meaning and action in concert with each other through their mutual participation in embodied sequences of talk- and action-in-interaction.