1

Introduction

Rebecca Clift and Elizabeth Holt

(I)n real life people talk most of all about what others talk about – they transmit, recall, weigh and pass judgement on other people's words, opinions, assertions, information; people are upset by others' words, or agree with them, contest them, refer to them and so forth.

(Bakhtin, 1981: 338)

1.1 Introduction

This volume is an investigation of reported speech in naturally occurring spoken interaction. We recurrently use talk to report talk, whether we are reporting the compliment someone gave us or conveying how we made a complaint or told a joke. In the following extract, for example, the speaker uses reported speech as part of a story relating how she was the victim of a nasty put-down (arrowed):[1]

(1) [Holt: C85: 4: 2–3] (Lesley has been looking around the stalls at a church fair)

```
1    Lesley:   AND uh  ↑we were looking rou-nd the
2              ↓sta:lls 'n poking about 'n he came
3              up t'me 'n he said Oh: hhello Lesley, (.)
4       →      ↑still trying to buy something f'nothing,
5       →      .tch! .hh [hahhhhhhh!
6    Joyce:            [.hhoohhhh!
```

Lesley's animation of the man's words is the culmination of her reporting of a series of actions. It is this phenomenon – the reproduction of prior talk in a current interaction – that the studies

---

[1] For a key to transcription symbols, see pages xi–xvii.

in this volume are concerned with. Together they bear witness to the use of reported speech and its variant forms across the range of interactional contexts from ordinary conversation to so-called institutional talk such as political interviews and debates. While engaging with material as diverse as story-telling, witness testimony in court, interaction between spiritual mediums and their sitters and video data of an aphasic man, the chapters have a central focus: the design and placement of reported speech – and thought – in sequences of interaction. Aspects of design include its lexical and prosodic construction; issues of placement relate to how turns in reported speech are built to follow particular others, and the responses that they in turn generate. In the extract above, for example, Lesley introduces reported speech as the climax of the story she has been telling; story climaxes, as we shall see, are one of the recurrent interactional sites for reported speech. The design and sequential placement of reported speech thus display systematicities which are only available by close analytic attention to several instances of the same phenomenon; the chapters in this volume are characterised by a commitment to such analytic attention.

A more detailed survey of the contents follows in due course, but first we sketch the background to existing work on reported speech and the main theoretical issues to have emerged from it. As we shall see, the relatively recent advent of interactionally grounded studies of reported speech has promised to illuminate many of the theoretical issues formerly regarded as intractable. The rationale for adopting the rigorously empirical approach of conversation analysis is duly set out here, followed by some of the earlier findings from conversation analytic work on reported speech; it is in this work that the current contributions have their origins.

1.2 Background and main themes

Work on reported speech in recent years has emerged from a variety of disciplines, most prominently literary theory, philosophy, linguistics and sociology.[2] The proposal of the Bakhtin/Volosinov circle

---

[2] For a comprehensive bibliography of work on reported speech, see Güldemann et al. (2002).

that much of what we say is permeated with the voices of others has proven highly influential beyond the domain of literary theory; much subsequent empirical work has pursued Bakhtin's notion of 'polyphony' and his claim that any utterance contains 'the half-concealed or completely concealed words of others' (1981: 92). Within philosophy, reported speech has been of interest in its reflexive capacity (D. Davidson, 1968–9, 1984; Quine, 1960) and in this respect converges with work on metapragmatics within linguistics (see, for example, the collection in Lucy, 1993), which has its origins in Jakobson's concern with reported speech as 'a speech within speech, a message within a message...' (1971: 130). It is the work in linguistics that has produced the most diverse range of perspectives. Across this diversity it is nonetheless possible to identify three central concerns in the literature: that with **forms** of reported speech; with its **authenticity**, and with **what it does**. While all three, as we shall see, continue to be the focus of ongoing research, it is evident that the concern with forms of reported speech generally predated work on its authenticity, and it is only in relatively recent years that research has focused on what reported speech does in interaction. This latter focus marks the increasing influence on linguistic research of work in sociology, and it is at the intersection of these two domains that much conversation analytic work on reported speech has emerged and where the current study has its starting point. To chart the route to this point, we now briefly sketch the three main preoccupations of previous work in reported speech.

1.2.1 Forms of reported speech

Of structural linguistic studies, a major focus has been the distinction between so-called direct reported speech (DRS) and so-called indirect reported speech (IRS). Jespersen proposed that:

When one wishes to report what someone else says or has said (thinks or has thought) – or what one has said or thought oneself on some previous occasion – two ways are open to one. Either one gives, or purports to give, the exact words of the speaker (or writer): *direct speech*. Or else one adapts the words according to the circumstances in which they are now quoted: *indirect speech* (oratio obliqua).   (1924: 290)

On Jespersen's account, extract (1) – cited above – shows an instantiation of the former; the extract below, in which a speaker is summarising or conveying the gist of a previous thought or locution, is an example of the latter:

```
(2)  [Rahman: II:4]   (Simplified)
1       Jenny:    An' I:van had said to me in the mo:rn ing
2          →      would I run 'im through to Saltburn .hh
```

Jenny here conveys what Ivan purportedly said without claiming fidelity to his original utterance, the presence of the pronoun 'I' clearly indicating that Jenny is speaking from her perspective. Besides this proposed distinction in the linguistics literature between DRS and IRS, more recent work has focused on what has come to be known as either 'free indirect' or 'quasi-direct' speech (Coulmas, 1986; Banfield, 1973, 1982; for a survey, see McHale, 1978), an amalgam of direct and indirect reported speech:

```
(3) [NB: II: 2: 10]
1     Nancy:    ... I only had o:ne (0.3) .hhhhhh (0.4)
2               dero:gatory rema:rk? if: you c'd call it
3               tha:t a:nd ah,h (0.6) u-it ca:me from a
4          →    gi:rl (0.2) and she said she fe:lt thet
5          →    I: would of gott'n more out'v the cla:ss
6          →    if I hed not been en eVOIder, h w'tever
7          →    sh'meant by tha:t, .hhhhh u-but that
8          →    ah:::, (0.5) I will c'ntinue t'remember
9          →    th'class en gro:w from it. Er sump'n (.)
```

The majority of Nancy's report here is indirect: the pronouns are from the point of view of the current speaker, not the original speaker. However, 'en eVOIder' (line 6) appears to be directly reported. Elements of the last part of the reported speech – 'will c'ntinue t'remember th'class en gro: w from it' (lines 8–9) – appear also to be directly reported.

Much linguistic research has been grounded in this proposed three-way distinction between direct, indirect and quasi-direct speech. Thus Li (1986) provides a detailed characterisation of the differences between direct and indirect reported speech in lexico-syntactic and prosodic terms; Banfield (1973), Partee (1973),

Mayes (1990) and Longacre (1985) have also compared direct and indirect reported speech with respect to their distinctive structural characteristics. Of more functionally oriented research, Coulmas claims that, while IRS is related from the current speaker's point of view (see also Leech and Short, 1981), DRS:

> is not the reporter's speech, but remains the reported speaker's speech whose role is played by the reporter.   (1986: 2)

And according to Li (1986), DRS is used to convey both the form and content of the reported utterance, including gestures and facial expressions. In IRS, however, the speaker has the option of communicating a comment on the utterance as it is uttered. Thus, if the utterance is reported in an angry voice, in direct form the anger will be heard as the reported speaker's, and in indirect form it will be heard as the current speaker's comment on the utterance.

The concern with different forms of reported speech has led to lively interest in its introductory components, sometimes called 'quotatives' (Mathis and Yule, 1994), most commonly in English – as in extract (1) – a pronoun and a *verbum dicendi* such as 'say'. Such quotatives may be present in what is identifiably both DRS and IRS, although in English one common characteristic of indirect reports is that the quotative is followed by the complementiser 'that' (Li, 1986).[3] However, while variants of *pronoun + say* may be considered the paradigmatic introductory component of reported speech, research has identified a number of alternatives. So Tannen's (1989) survey of quotatives includes 'tell', 'go' and 'like'. The apparently increasing use of *be + like* as an introductory component has been the focus of recent attention by Blyth et al. (1990), Romaine and Lange (1991), Ferrara and Bell (1995), Tagliamonte and Hudson (1999), Macaulay (2001) and Cukor-Avila (2002). The claim by Romaine and Lange that 'like' blurs the boundary between DRS, IRS and reported thought, claiming less commitment to the original than 'say' does, touches on the second of the three main concerns in the linguistics literature in this domain: the authenticity of reported speech.

---

[3] See Haakana (this volume) for Finnish as a contrast case.

1.2.2 The authenticity of reported speech

Research into reported speech began with the assumption (derived from the lay assumption (see Mayes, 1990: 330–31)) that direct speech is more accurate than indirect speech. Thus, Bally (1914) viewed DRS as 'a phonographic reproduction of the thoughts and words' of the original speaker (quoted in Clark and Gerrig, 1990: 795). But more recent work has shown how DRS is, in fact, rarely an accurate rendition of a former locution. Volosinov (1971) was the first to criticise the assumption that reported speech is an authentic rendition of the original, proposing that the meaning of the original utterance is inevitably altered in the reporting context (see Dubois (1989) on what she calls 'pseudoquotation', and Sternberg (1982) on claims regarding the reframing of reported speech). This claim has been supported by psycholinguistic research. Thus Lehrer (1989) shows that, in experiments to test the memory of prose, subjects tend to remember the meaning of utterances rather than the form, and that verbatim recall is unusual. Mayes (1990: 331) investigated the authenticity of the reported speech in her corpus and claimed that at least 50 per cent were inventions by the current speaker. Included in her collection, along with 'plausible quotes' and 'improbable quotes' (for example, a speaker reporting an utterance made twenty years earlier), were 'highly improbable quotes' (such as a 'Greek chorus' where a quote is attributed to more than one person) and 'impossible quotes' (including hypothetical quotations). Thus it would seem that the term 'reported speech' is somewhat of a misnomer;[4] as we shall see, one of the concerns of this volume will be to engage with the reasons for this.

1.2.3 What does reported speech do?

While early linguistic studies of reported speech were overwhelmingly concerned with structural questions for which the use of constructed exemplars or literary texts was perceived to be adequate, the past twenty years have seen an increasing number of empirical studies of reported speech. In part this is due to a convergence of

[4] Tannen (1989) goes so far as to adopt the term 'constructed dialogue' for these reasons.

structural and comparative linguistic concerns: many languages grammaticalise quotative constructions (see, for example, Cohen et al. (2002) on a range of East African languages, and the collection in Aikhenvald and Dixon (2003), and there has been keen interest amongst typologists in this grammatical encoding of reported speech (see the collections in Lucy (1993) and Güldemann and von Roncador (2002)). This move away from literary and textual materials towards naturalistic speech data in a variety of languages has also engendered an increasing interest in functional and pragmatic aspects of reported speech. So comparative linguistic studies, grounded in the ethnographic tradition, have investigated aspects of reported speech in the languages of North America (see, e.g. Collins, 1987; Moore, 1993; Urban, 1993), Austronesia (see, e.g. Besnier, 1993; Parmentier, 1993; McGregor, 1994), South and Central America (see, e.g. Adelaar, 1990; Basso, 1986; Shoaps, 2004) and Africa (see, e.g. Aaron, 1992; Clements, 1975).

Of linguistic studies concerned with the generic properties of reported speech, many have remarked on the dramaturgical quality of DRS in particular (see Li, 1986; Tannen, 1989; Wierzbicka, 1974). It has been proposed that reported speech is used in stories not only to replay an interaction but also to enable the speaker to simultaneously convey his or her attitude towards the reported utterance. Labov (1972) distinguishes between 'external evaluation', where the point of a story is explicitly explained, and 'internal evaluation' where it is conveyed through the story itself. DRS is, he argues, a means of internally evaluating the story and is therefore more effective because it allows the recipient to draw his or her own conclusions about the characters and events recounted. Mayes (1990) notes how reported speech is often used at the climax of stories and proposes this as an effective way of conveying the point of a narrative.[5]

Much research in recent years, aiming to pursue the interactional motivations for the use of reported speech, has shown the influence of the sociologist Erving Goffman's observations on social interaction. Goffman noted that:

[5] The association between reported speech and the climax or punchline of stories is not restricted to English (see, for example, Polanyi, 1982; Li, 1986; Larson, 1987).

> In daily life the individual ordinarily speaks for himself, speaks, as it were, in his 'own' character. However, when one examines speech, especially the informal variety, this traditional view proves inadequate...When a speaker employs conventional brackets to warn us that what he is saying is meant to be taken in jest, or as mere repeating of words by someone else, then it is clear that he means to stand in a relation of reduced personal responsibility for what he is saying. He splits himself off from the content of the words by expressing that their speaker is not he himself or not he himself in a serious way.   (1974/1986: 512)

In observing that reported speech is an intrinsic feature of the way we interact, Goffman echoes Bakhtin; but Goffman subsequently proposed that reported speech is a natural upshot of a more general phenomenon in interaction: shifts of 'footing', defined as 'the alignment of an individual to a particular utterance...' (1981: 227). Goffman is concerned to break down the roles of speaker and hearer into their constituent parts. The speaker subsumes the roles of 'animator' – 'the sounding box', the 'author' – 'the agent who scripts the lines' and the 'principal' – 'the party to whose position the words attest' (1981: 226). All three roles may be played by a speaker at the same time, but often they are not. For instance, the vice-president reading out the speech on behalf of the president is only the animator. The author may be the president in conjunction with a scriptwriter. The principal is the president, as well as the represented political party she represents. In reporting the speech of another person the speaker is the animator but not the author or principal. Thus, our ability to use reported speech stems from the fact that we can adopt different roles within the 'production format', and it is one of the many ways in which we constantly change footing as we interact (see Levinson (1988) for an elaboration of Goffman's proposal).

The 'reduced personal responsibility' that Goffman claims for reported speech therefore appears to account for much of the licence that speakers seem to take in using it; thus, Goffman (1981) notes how curses and taboo utterances may be used with greater freedom than if speakers are speaking 'in their own voice'.

Goffman's work has proven foundational in the investigation of reported speech in interaction because it recognises that as much is to be learned from examining the context of reported speech – and the switch from non-reported to reported speech – as examining (as

many structural studies had) just the reported speech itself. While Goffman is not in his own work concerned with the analysis of actual instances of interaction (for a critique, see Schegloff, 1988), it provides a framework for researchers concerned with investigating reported speech in its most basic environment of occurrence: ordinary conversation. Before examining some of the products of this research, we provide a brief sketch of some of the basic tenets of conversation analysis.

1.3 Conversation Analysis: a brief sketch

Conversation analysis (CA) – the adopted name for what is perhaps more accurately termed the study of talk-in-interaction – takes as a basic tenet the fact that social interaction is not haphazard but orderly, and that the methodical, organised nature of our social life can be studied by close attention to naturally occurring materials (for more detailed explication of the methods of CA, see Atkinson and Heritage, 1984; Heritage, 1984a, Chapter 8; and Psathas, 1995). The transcription of these audio- or video-recorded materials according to the system devised by Gail Jefferson (see 'Transcription conventions' on pages xi–xvii) involves registering features of the production and articulation of talk – and its absence – which capture the temporal unfurling of turns-at-talk. So features such as overlapping talk, in-breaths, the infiltration of laughter into talk, aspects of pace and prosody – all elusive to memory or intuition – are captured in the transcript and so accessible for their possible interactional import. These transcriptions then make the data available for repeated inspection and analysis. This has two important consequences: it allows for methodological transparency, such that the presence of the data makes any analysis accountable to it, and disputable because of it; and it also enables the collection of multiple examples of the same phenomenon, which reveals the systematicities underlying the apparent disorder and fragmentation of interaction. It is in establishing these systematicities that interpretation becomes analysis. And, because the analysis focuses on patterns observable in the data, analysts are able to avoid speculating about participants' intentions and understandings, or external constraints and influences that might impact on their conduct. Schegloff and Sacks note of their pioneering work in this field:

We have proceeded under the assumption (an assumption borne out by our research) that in so far as the materials we worked with exhibited orderliness, they did so not only to us, indeed not in the first place for us, but for the co-participants who had produced them. If the materials (records of natural conversation) were orderly, they were so because they had been methodically produced by members of the society for one another, and it was a feature of the conversations we treated as data that they were produced so as to allow the display by the co-participants to each other of their orderliness, and to allow the participants to display to each other their analysis, appreciation and use of that orderliness. Accordingly, our analysis has sought to explicate the ways in which the materials are produced by members in orderly ways that exhibit their orderliness and have their orderliness appreciated and used, and have that appreciation displayed and treated as the basis for subsequent action.   (1973: 290)

Turns are, in the first instance, built to contribute to the sequence of actions in which they occur; thus to analyse them in isolation is to ignore the way they are built to display analysis of, and participation in, the actions embodied by prior turns. Every turn-at-talk therefore displays the participant's definition of the situation; it displays an understanding of the activity sequence to which it contributes, and of what is an appropriate contribution to that sequence. This has an important methodological upshot: the analyst can use the sequential nature of turns at talk as a resource for accessing the participants' analysis of the nature of the actions engaged in.[6] From this perspective we can see how Goffman's observations on footing and the relationship between reported and non-reported speech have been an important influence on conversation analytic research into reported speech. It is to this work – the foundation for the current volume – that we now turn.

1.4 CA studies of reported speech

In some respects, detailed analysis of reported speech in context has highlighted differences between claims by linguists and sociolinguists and conversation analytic ones, while in others CA research has supported and extended previous findings. We begin by considering some of the discrepancies illuminated by existing CA work.

[6] For a more detailed consideration of CA method and its contribution to linguistics, see Clift (2005).