Managing problems in speaking

Herbert H. Clark

Department of Psychology, Building 420, Stanford University, Stanford, CA 94305-2130, USA

Abstract

The problems that participants in conversation have, it is argued, are really joint problems and have to be managed jointly. The participants have three types of strategies for managing them. (1) They try to prevent foreseeable but avoidable problems. (2) They warn partners of foreseeable but *un*avoidable problems. And (3) they repair problems that have already arisen. Speakers and addressees coordinate actions at three levels of talk: (1) the speaker's articulation and the addressees' attention to that articulation; (2) the speaker's presentation of an utterance and the addressees' identification of that utterance; and (3) the speaker's meaning and the addressees' understanding of that meaning. There is evidence that the participants have joint strategies for preventing, warning about and repairing problems at each of these levels. There is also evidence that they prefer preventatives to warnings, and warnings to repairs, all other things being equal.

Zusammenfassung

Die kommunikativen Probleme, die Gesprächspartner in Unterhaltungen haben, sind gemeinsame Probleme, so die vorliegenden These, und müßen gemeinsam bewältigt werden. Gesprächspartner besitzen drei Strategien, diese Probleme zu bewältigen. (1) Verhindern von Problemen, die voraussehbar und abwendbar sind. (2) Warnen des Gesprächspartners vor Problemen, die voraussehbar doch unabwendbar sind. (3) Korrigieren von Problemen, die bereits aufgetreten sind. Sprecher und Adressat koordinieren ihre Sprechhandlungen auf drei Ebenen. Diese Ebenen betreffen: (1) die Artikulation von Äußerungen durch den Sprecher und die Aufmerksamkeit des Addressaten; (2) die Präsentation der Äußerung durch den Sprecher und die Identifizierung der Äußerung durch den Adressaten; (3) was der Sprecher mit der Äusserung meint und wie der Adressat sie versteht. Anhand empirischer Evidenz wird gezeigt, daß Gesprächspartner Probleme auf allen drei Ebenen bewältigen, indem sie gemeinsame Strategien anwenden, nämlich Verhindern, Warnen und Korrigieren. Verhinderungen werden (ceteris paribus) Warnungen vorgezogen und Warnungen werden Korrekturen vorgezogen.

Résumé

Lorsqu'on trouve des problèmes dans les conversations, ce sont tout d'abord des problèmes collectifs et les participants doivent les co-gérer. Les participants à une conversation disposent de trois sortes de stratégies pour les gérer. (1) Ils essayent de prévenir les problèmes qui sont prévisibles mais évitables. (2) Ils avertissent leurs associés contre les problèmes prévisibles mais inévitables. (3) Ils réparent les problèmes qui se sont déjà révélés. Alors, les

interlocuteurs et leurs destinaires agissent en commun à trois niveaux de la parole. Ils coordonnent (1) l'articulation de l'interlocuteur à l'attention des destinaires. Ils coordonnent (2) la présentation d'une expression par l'interlocuteur à l'identification de cette expression par les destinaires. Et ils coordonnent (3) la signification signalée par l'interlocuteur à la compréhension de cette signification par les destinaires. En somme, les participants à une conversation travaillent ensemble afin de prévenir, d'avertir et de réparer les problèmes à chacun de ces niveaux. En plus, toutes choses égales par ailleurs, ils préfèrent les préventions aux avertissements, et les avertissements aux réparations.

1. Introduction

When the participants of a conversation have problems, they manage most of them quickly, skillfully and without apparent effort. These problems arise in everything they do, from maintaining attention to maintaining face. Some result in disfluencies – pauses, repairs, fillers (like "uh" and "um"), repeats, word fragments, fresh starts – but others result in a variety of other phenomena. How are these problems managed? A common view is that speakers monitor for them and repair them when they discover them. In this paper I suggest that this view is too narrow. Managing problems is really part of a larger system in which repairs are only one strategy.

Language use is fundamentally a joint activity, and that is reflected in the way problems are managed (Clark and Schaefer, 1989; Clark and Wilkes-Gibbs, 1986; Schegloff et al., 1977). When Ann and Bob talk to each other, they each perform individual actions such as uttering words, identifying sounds and forming interpretations, but many of these actions are really parts of actions performed by the *pair* of them Ann-and-Bob. I will call actions by the pair Ann-and-Bob *joint actions*, and I will call Ann's and Bob's individual actions within them *participatory actions* (Clark and Carlson, 1982; Clark and Schaefer, 1989; Cohen et al., 1990). In conversation – the fundamental site of language use – speaking and listening are participatory actions.

Ann's actions in talk are not independent of Bob's, or vice versa, and that goes for their problems as well. When Ann needs extra time to plan an utterance, that is not her problem alone. The time she needs belongs to Ann-and-Bob, so she has to coordinate with Bob on her use of that time. Likewise, when Bob does not understand Ann, the problem is not his alone, or hers alone. It is Ann-and-Bob's, and it takes the two of them working together to fix it. There are two principles here: (1) the problems that arise in language use are *joint* problems; and (2) dealing with these problems requires *joint* management.

To complicate things, problems arise at several levels of conversation. Suppose Ann is saying something to Bob. Here are four levels of action, starting at the bottom:

Level 1. Vocalization and attention. At the lowest level, Ann vocalizes sounds, getting Bob to attend to those vocalizations. She cannot vocalize those sounds unless she has Bob's attention, and Bob cannot register her vocalizations without attending to them. That takes Ann's and Bob's coordination.

Level 2. Presentation and identification. One level up, Ann presents an utterance for Bob to identify. She must be sure Bob has identified the utterance she has presented, and he must be sure of it too, and that also takes coordination.

Level 3. Meaning and understanding. One more level up, Ann gets Bob to understand what she means by her utterance. The two of them must reach the mutual belief, called the *grounding criterion*, that Bob has understood what Ann meant well enough for current purposes (Clark and Schaefer, 1989; Clark and Wilkes-Gibbs, 1986).

Level 4. Proposal and uptake. At the highest level, Ann gets Bob to commit to some *joint project* (Clark, 1994). To accomplish this, she must propose a project Bob is willing and able to take up, and he must take it up. That requires still further coordination.

All four levels consist of joint actions. They each

Repairs

problem detect **repair**

time

Warnings

anticipate **warn** problem

time

Preventatives

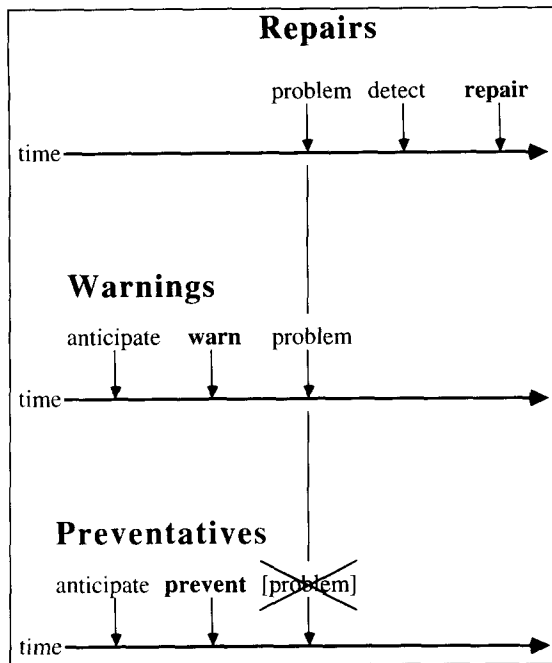anticipate **prevent** [problem]

time

Fig. 1. Temporal patterns of problem, detection and management for repairs, warnings and preventatives.

require Ann and Bob to coordinate on their individual actions. At each level, the problems Ann and Bob have as individuals are also problems for their joint action.

So how *do* Ann and Bob manage problems in conversation? Ordinarily, we think of them repairing problems that have already occurred, but that conception is too narrow. Problems in conversation are like infections: people prefer to deal with them before they grow into something worse. People's strategies for managing problems in conversation are much like physicians' strategies for managing infections:

Preventatives: these are like inoculations in averting anticipated but avoidable problems.

Warnings: these are like palliatives in helping participants prepare for anticipated but *un*avoidable problems.

Repairs: these are like antibiotics in remedying problems that have already appeared.

So problem, detection and remedy can emerge in three patterns, as shown in Fig. 1. With repairs,

the order is problem, detection and repair, but with warnings and preventatives, the order is anticipation, management and problem. With preventatives, the problem is never in fact realized.

In conversation as in medicine, people prefer preventatives to warnings, and warnings to repairs, all other things being equal. Let us consider levels 1, 2 and 3 of language use – leaving level 4 to another occasion – and see how they provide evidence for preventatives, warnings, repairs and their ordering.

2. Vocalization and attention

At level 1 (vocalization and attention), Bob must attend to Ann while she vocalizes her utterance or they will fail. Joint actions like this depend on the participants doing their parts, so for Ann and Bob to be sure of success, they need *evidence* that they are each doing their parts. Ann should look for evidence that Bob is attending to her, and he should try to provide that evidence. Consider this invented example:

Ann: Bob
Bob: [3 sec of no response]
Ann: Bob [louder]
Bob: What?

Ann tries to summons Bob with her first utterance, but gets no response. She takes that as evidence that Bob *was not* attending to her vocalization, a problem she has to repair. She does that by repeating the summons – only louder to capture his attention. This time he responds, giving her evidence that she has succeeded.

Whose problem is this – Ann's or Bob's? Neither of them can be held solely responsible. The problem arose from the mis-coordination of Ann's vocalization and Bob's attention. Perhaps Ann should have been more certain of Bob's attention before vocalizing, or he should have been paying closer attention, or both. In any event, Ann and Bob's joint action led to a joint problem, which required a joint remedy.

If Ann and Bob had worked together, they might have avoided the problem in the first place.

There are effective preventatives for just this purpose. Several such strategies have been described by Goodwin (1981), one of which is illustrated here with Lee talking to Ray:

> Lee: Can you bring – (0.2) Can you bring me here that nylon?

A videotape of this utterance shows Lee watching Ray's eye gaze. Just as Lee wants to start speaking, he sees that Ray is looking away, so if he were to start then, Ray would not be attending, and that would create a problem they would have to repair later. Instead, Lee tries to prevent the problem by producing "can you bring" not to say something, but to get Ray's attention. And he starts again only once he has Ray's attention – precisely as Ray begins to turn his head toward Lee. Lee's strategy, which itself requires a joint action, was designed not merely to remedy an existing problem – Ray was attending to something else – but to prevent a future problem – Ray would fail to understand. Ray's response – turning his eyes toward Lee – was part of that strategy. Ray provided it to show that he was ready to attend to Lee's vocalization.

Most models of language use take the coordination of vocalization and attention for granted. In fact, trying to coordinate on these processes often causes problems that the participants are forced to deal with. The strategies I have illustrated are only two ways they do that. There are many more.

3. Presentation and identification

At level 2 (presentation and identification), Ann must present an utterance and get Bob to identify its words, constituents and other relevant structure, and that too requires coordination. For them to succeed, Ann needs evidence that Bob is identifying her utterance, and he must provide that evidence.

The evidence addressees provide shows them to be in one of several states (Clark and Schaefer, 1989). (1) They may not have identified any utterance at all. When speakers detect such a problem, they usually repair it by repeating the utter-

ance, as in this spontaneous example (from (Svartvik and Quirk, 1980): [1]

> A: ((where are you))
> B: m?
> A: where *are* you
> B: well I'm still at college
> A: [continues]

Or (2) addressees may have identified only part of an utterance. To remedy this, they may ask for a confirmation or a repeat of that part, which speakers then provide, as here (from (Svartvik and Quirk, 1980)):

> Roger: now, – um do you and your husband have a j– car?
> Nina: – have a car?
> Roger: yeah
> Nina: no –

Or (3) they may have misidentified all or part of a presentation. When speakers detect such a problem, they usually remedy it by repeating the misidentified part, as in this exchange of a street address (from (Svartvik and Quirk, 1980)):

> A: yes forty-nine Skipton Place
> B: forty-one
> A: nine . nine
> B: forty-nine, Skipton Place,

The point is that all three problems are really joint problems, and they are managed with joint remedies – remedies that require the coordination of speakers and addressees.

Speakers also discover problems as they monitor their own presentation, and they generally repair the problems immediately – before they mislead addressees too far. Here is one example (from (Svartvik and Quirk, 1980)):

> Ann: they still talk about rubbish tins, which is the American the Australian
> Beth: yeah
> Ann: expression, . for that thing you put all the . stuff in at the back gate, you know

---

[1] In this and later examples from Svartvik and Quirk (1980), I will use the following symbols: "." for a brief pause (of one light syllable); "–" for a unit pause (of one stress unit or foot); "," for the end of a tone unit, marked only if it comes mid-turn; "((words))" for incomprehensible words; and ":" for lengthened vowels.

Ann catches the error in "American" on her own and instantly repairs it to "Australian" in a *self-correction*. She does not wait for Beth to catch the error and correct it for her – an *other-correction* (see (Schegloff et al., 1977)). Immediate self-corrections are preferred to other-corrections for at least two reasons. First, they are not as costly – they require only an extra word or phrase instead of two extra turns. And second, although they repair one problem, they prevent deeper and more costly misunderstandings down the line. They are not only repairs but preventatives.

There are many problems that speakers anticipate even before they become evident. Speakers recognize, for example, that most utterances have an *ideal delivery* – a pronunciation that is fluent, correct and optimal for identification (Clark and Clark, 1977). They also recognize that any deviation from the ideal might cause their addressees problems, so they should try to achieve the ideal delivery. The trouble is that they usually cannot formulate an entire presentation before they begin speaking. They are forced to formulate one phrase at a time, and they often have to interrupt their utterances to do that. Since they recognize

that interruptions and pauses pose problems for their addressees, how should they proceed?

Whenever speakers foresee a delay or interruption they cannot prevent, they can help their addressees prepare for it by warning them about it. One way is with the filler "uh" or "um", as here (from (Svartvik and Quirk, 1980)):

Reynard:	i is . is it this year, that $u:h$ Nightingale goes
Sam:	– – $u:h$ no next year,
Reynard:	– – $u:m$. – sixty – f–
Sam:	sixty–five
Reynard:	– four sixty – five
Sam:	yeah

Evidence shows that speakers use "uh" to signal short interruptions, and "um" to signal more serious ones. Twenty-five university students were each asked 40 factual questions in conversational settings, and there was often a delay in their answers, like this (with pause length indicated in seconds):

| Experimenter: | In which sport is the Stanley cup awarded? |
| Respondent: | (1.4) um (1.0) hockey |

When the respondents began without a filler, the delay averaged 2.23 seconds; when they began with "uh", it averaged 2.65 seconds; but when they began with "um", it averaged 8.83 seconds (Smith and Clark, 1993). The delays of answers are shown in Fig. 2. In a study of the London–Lund corpus of English conversation in (Svartvik and Quirk, 1980), Fox Tree and I computed the percentage of times that "uh" and "um" were preceded and followed by perceptible pauses (Clark and Fox Tree, In preparation). The percentages are summarized in Fig. 3. Speakers produced "uh" and "um" quite often after pauses. But they were even more likely to use "um" than "uh" when they anticipated further pauses. So in both studies, speakers used "uh" and "um" to warn addressees about the size of interruption they were anticipating.

Speakers also warn addressees about problems in formulating noun phrases. Although "the" is ordinarily pronounced "thuh", it is sometimes pronounced "thee" when speakers foresee a problem in formulating the current definite noun phrase. In the London–Lund corpus, there were
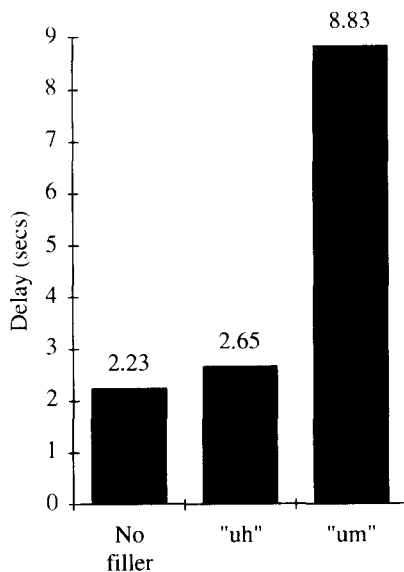

Fig. 2. Delay in answer preceded by "uh", "um" or no filler (from (Smith and Clark, 1993)).
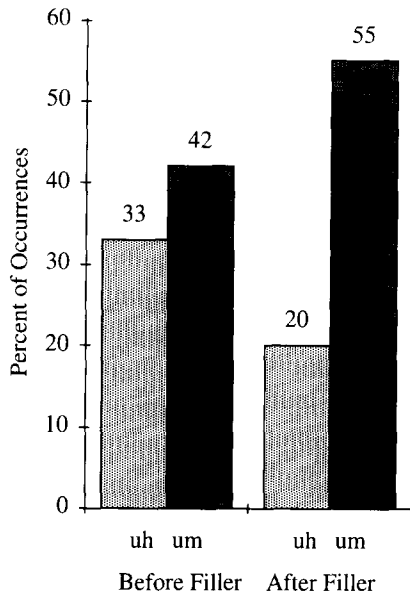
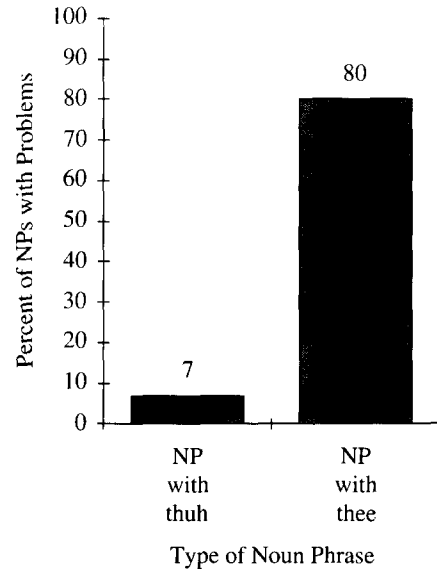Fig. 3. Percentage of pauses before and after "uh" and "um" (from (Clark and Fox Tree, In preparation)).

Fig. 4. Percentage of disfluencies in noun phrases immediately after "thuh" or "thee" (from (Fox Tree and Clark, 1994)).

disruptions in 7% of the NPs introduced by "thuh", but in 80% of those introduced by "thee" (Fox Tree and Clark, 1994). The percentages are shown in Fig. 4. Apparently, speakers choose "thee" to warn of an approaching disruption, and that should help addressees prepare for it. Speakers appear to have other warning devices as well.

At the level of presentation and identification, then, the participants Ann and Bob not only repair existing problems, but try to prevent future problems and warn of approaching but unavoidable problems.

4. Meaning and understanding

At level 3 (meaning and understanding), Ann must get Bob to understand what she means with her utterance. To succeed, they must reach the mutual belief that he has understood her well enough for current purposes, and for that, he must provide her with evidence of his understanding (Clark and Schaefer, 1989; Clark and Wilkes-Gibbs, 1986). Ordinarily, addressees pro-

vide evidence that confirms their understanding, as here (from (Svartvik and Quirk, 1980)):

Alan: how far is it from Huddersfield to Coventry .
Barbara: um . about um a hundred miles –
Alan: so, in fact, if you were . living in London during that period, . you would be closer – .

Barbara displays her interpretation of Alan's question by answering it appropriately. If her answer had revealed a misunderstanding, Alan would have corrected her, and because he did not, he displays that he believes she understood him well enough for current purposes. Not so in the next example (from (Svartvik and Quirk, 1980)):

Ken: k who evaluates the property – – –
Ned: uh whoever you asked, . the surveyor for the building society
Ken: no, I meant who decides what price it'll go on the market –
Ned: (– snorts) . whatever people will pay – –

Ned displays his interpretation of Ken's question in his answer, but that answer reveals that he

misunderstood Ken, so Ken corrects him, beginning "no, I meant". More often, it is the addressee who notices the problem and then asks for a repair, as Dar does here (from (Svartvik and Quirk, 1980)):

Sam: well wo uh what shall we do about uh *this* boy then
Dar: Duveen
Sam: m
Dar: well I propose to write, uh saying . I'm very sorry [continues]

When Dar is not certain which boy Sam is referring to, he gets Sam to confirm that it is Duveen.

It is even more common for speakers to find meaning problems and repair them before they cause further misunderstanding, as in this example (from (Svartvik and Quirk, 1980)):

Jane: this is the funny thing about academics, – . that if you're no– u:h *you know*, I I've . come to it, so late, . *I mean* I've had a lifetime of experience, rolling around,

At one point, Jane says "that if you're no–" then cuts herself off and offers the repair "I I've . come to it, so late". She then gives a second repair, "I've had a lifetime of experience, rolling around". But she does more than make the repairs. She signals the type of repairs she is making by means of *editing terms* – "u:h you know" for the first and "I mean" for the second (see (Levelt, 1983)). She helps her partner prepare for the repairs by warning him both about when they are coming and about why.

Speakers have other less obvious strategies for preventing misunderstandings. One is the use of hedges such as "kind of", "sort of" and "like". Consider "sort of" in this example (from (Svartvik and Quirk, 1980)):

Duncan: the funny thing wa– about it was that, he apparently . played . cricket, this always seemed to be rather odd, – that . at . I don't know how old he was, but – probably getting on, . you know I mean, . sort of fortyish, – . he should go on playing cricket

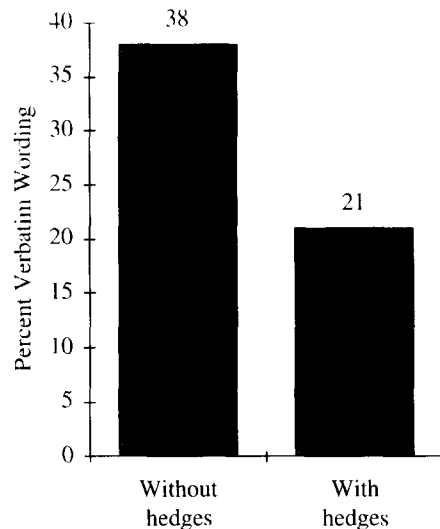Duncan would have implied that the adjective "fortyish" captured precisely what he meant if he

Fig. 5. Percentage of original wording reproduced verbatim with and without hedges (from (Wade and Clark, 1993)).

had not used "sort of". But by adding "sort of", he implies that "fortyish" is only approximate. Note all the other hedges he used in coming to that word – "I don't know how old he was, but" "probably getting on", "you know", "I mean" and the "-ish" on "forty". Experimental evidence shows that speakers use hedges such as "kind of", "sort of" and "like" to indicate that they are being less accurate (Wade and Clark, 1993). University students were asked to retell stories they had just heard, and in their retellings, they produced direct quotations both with and without hedges. When they did not use hedges, they correctly reproduced 38% of the wording from the original stories. When they did use hedges, the percentage was only 21%, as shown in Fig. 5. These speakers were right to warn their addressees about their imprecision (see also (Wade, 1993)).

At the level of meaning and understanding, then, people exploit a variety of strategies in managing problems. The first point is that the participants in a conversation methodically establish the mutual belief that the addressees have understood what the speaker meant well enough for current purposes. To do this, speakers look

for positive evidence of understanding, and their addressees try to provide that evidence. The two of them have a toolbox full of joint strategies for dealing with the misunderstandings and lack of understanding that inevitably occur. They also have strategies for warning about and preventing anticipated problems of understanding. Speakers use "I mean" and "you know" to warn about the type of self-repair they are making. And they use hedges like "kind of" and "sort of" to prevent interpreting certain words or phrases too precisely, too literally.

5. Conclusion

In a common view of conversation, the participants manage the problems they encounter by monitoring for them and by repairing them when they arise. I have argued that this view is too narrow. For one thing, managing problems is something the participants do together. All problems are ultimately joint problems and have to be managed with joint strategies. For another thing, speakers do more than make repairs. They have strategies for *preventing* certain problems from arising at all. For problems that are unavoidable, they have strategies for *warning* their partners to help them prepare for the problems. And for problems that arise anyway, they work with their partners in repairing them. In the management of problems, preventatives are preferred to warnings. Repairs are the last resort.

References

H.H. Clark (1994), "Discourse in production", in *Handbook of Psycholinguistics*, ed. by M.A. Gernsbacher (Academic Press, San Diego).

H.H. Clark and T.B. Carlson (1982), "Hearers and speech acts", *Language*, Vol. 58, pp. 332–373.

H.H. Clark and E.V. Clark (1977), *Psychology and Language* (Harcourt Brace Jovanovich, New York), pp. 261–262.

H.H. Clark and J.E. Fox Tree (In preparation).

H.H. Clark and E.F. Schaefer (1989), "Contributing to discourse", *Cognitive Science*, Vol. 13, pp. 259–294.

H.H. Clark and D. Wilkes-Gibbs (1986), "Referring as a collaborative process", *Cognition*, Vol. 22, pp. 1–39.

P.R. Cohen, J.L. Morgan and M.E. Pollack, eds. (1990), *Intentions in communication* (MIT Press, Cambridge, MA).

J.E. Fox Tree and H.H. Clark (1994), "Pronouncing 'the' as /thiy/ to signal trouble in spontaneous conversation", Presented at Psychonomics Society, St. Louis, MO, 11 November 1994.

C. Goodwin (1981), *Conversational Organization: Interaction between Speakers and Hearers* (Academic Press, New York).

W.J.M. Levelt (1983), "Monitoring and self-repair in speech", *Cognition*, Vol. 14, pp. 41–104.

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

V.L. Smith and H.H. Clark (1993), "On the course of answering questions", *J. Memory and Language*, Vol. 32, pp. 25–38.

J. Svartvik and R. Quirk (1980), *A Corpus of English Conversation* (Gleerup, Lund, Sweden).

E. Wade (1993), Communicating about narrative (in)accuracy, Unpublished Ph.D. dissertation, Stanford University, Stanford, CA.

E. Wade and H.H. Clark (1993), "Reproduction and demonstration in quotations", *J. Memory and Language*, Vol. 32, pp. 805–819.