

SHOWING STRUCTURE: USING *UM* IN THE ACADEMIC SEMINAR

Johanna Rendle-Short

Abstract

Um and *uh* are generally considered to be indicative of dysfluency and uncertainty in speech production. However, analysis of the academic seminar indicates that the distribution of *um* and *uh* is not random. In specific well-defined environments *um* is used to indicate the underlying structure of the talk. Although Swerts (1998) has already suggested that fillers such as *um* and *uh* could be treated as discourse markers in Dutch, the notion that such tokens are functioning as discourse markers has not been developed in detail. This paper analyses the role played by *um* in a series of computer science seminars. Using traditional conversation analysis techniques, the paper focuses on the way in which *um* indicates structure in the academic seminar by maintaining coherence across bits of talk. It thus argues that in specific well-defined environments *um* functions as a discourse marker. This paper therefore addresses such issues as the role and function of *um* in seminar talk, the environments in which it occurs, and its use in indicating the structure of the talk to the listening audience.

Keywords: Academic monologue, Discourse markers, Um, Uh, Repair, Institutional talk.

1. Introduction

Although *ums* and *uhs* are frequently used in spoken discourse, they are generally considered to be indicative of dysfluency or problems of speech production (e.g. Chomsky 1965: 3). As such, they are discouraged and when presenting spoken discourse in a written form, *ums* and *uhs* are generally omitted. They are considered to be representative of performance difficulties, rather than being an integral part of the utterance itself. It is usually assumed that *ums* and *uhs* are disruptive for listeners (cf. Clark 1994; Fox Tree 2001), an issue that is particularly relevant in, for example, monologic talk and broadcasting (Goffman 1981: 198), where speakers are generally encouraged to remove such words from their speech.

Spontaneous speakers across many languages regularly produce filled pauses - pauses accompanied by tokens such as *um* or *uh* - although the actual sounds or tokens themselves may differ between dialects and languages (Cruttenden 1997: 30). Of particular concern to researchers interested in language use is, what is the role and function of such fillers in talk. One proposal is that when speakers are unsure of their response, they use *ums* and *uhs* to indicate production trouble. Thus, commencing with

an *um* or an *uh* is a useful mechanism for holding the floor (Maclay and Osgood 1959; Schegloff 1982; Schegloff, Jefferson and Sacks 1977). It indicates that there is more talk to come, even if momentarily there is a problem in production. This may be why such fillers are generally accompanied by a pause or a delay, as speakers try to find the correct or appropriate word(s) (Jefferson 1974; Schachter et al. 1991).

Fox-Tree's (1995, 2001) analysis of both English and Dutch listeners' responses to *ums* and *uhs* in on-line speech comprehension tasks suggests that listeners do not experience dysfluencies as disruptive. In fact, Brennan and Williams (1995) argue that in spontaneous speech, the filler *um* seems to help listeners recognise an upcoming target word faster; and Fox-Tree (2001) shows that hearing an *uh* has a beneficial effect on listeners' ability to recognise words in upcoming speech. This contrasts with listener's responses to *um* for which there is neither a beneficial nor a detrimental effect on listeners' ability to recognise words (Fox-Tree 2001). Thus fillers such as *um* and *uh* can function as a way of cueing the listener to be more attentive to the upcoming talk because it may be difficult to process. Therefore, although it is often assumed that such dysfluencies present obstacles to comprehension, it has been shown that dysfluencies can actually help listeners compensate for disruption and delays in spontaneous speech (Brennan and Schober 2001).

Understanding the way in which listeners respond to *ums* and *uhs* has also been the focus of Fox-Tree's more recent work on interpreting pauses and *ums* at turn exchanges. She has shown that overhearers respond differently depending upon whether a second speaker starts to talk immediately or whether there are filled or unfilled pauses prior to the talk (Fox-Tree 2002). She reports that when speakers said *um*, or paused, or both said *um* and paused, overhearers thought speakers had more production difficulty, were less honest, and were less comfortable with topics under discussion. Brennan and Williams (1995) have also demonstrated that listeners are sensitive to the presence or absence of fillers. Thus answers preceded by a filled pause, are less likely to be rated as being correct than those preceded by unfilled pauses (Brennan and Williams 1995: 395). Christenfield's research (1995) indicates that *ums* or *uhs* can actually make speakers sound more relaxed, although lay people might think that such people are inarticulate, ill-prepared and nervous.

One question that has arisen in the literature is whether *um* and *uh* function in the same way or whether they function differently. Although they are often discussed together, prior research suggests that *um* and *uh* are separate elements. Fox-Tree (2001), for example, demonstrates that *uhs* are used to indicate short delays as opposed to *ums* that are used to indicate longer delays. In addition, analysis of Dutch data shows that *ums* are more likely to occur at the beginning of intonation units rather than within intonation units (Swerts 1998). Brennan and Williams (1995) however found no support for what they call the *filler-as-morpheme* hypothesis. They concluded that *uh* and *um* do not contrast in meaning.

An additional proposal in the literature is that such tokens are used to mark discourse structure (Maclay and Osgood 1959; Swerts 1998). Maclay and Osgood (1959) note that the use of such fillers is not randomly distributed. They show that filled pauses occur before content words (for example, nouns, adjectives and verbs), at phrase boundaries, and at junctures of larger syntactic units. In contrast, filled pauses tend not

to occur before function words (for example, pronouns or particles). Thus, they argue, the speaker is operating on two levels, a lexical level and a grammatical level. Swerts (1998) in his work on elicited monologues in Dutch, further shows that filled pauses are more typical in the vicinity of major discourse boundaries. He also reports that the *ums* and *uhs* that occur in these positions are prosodically different. He thus concludes that discourse structure can be predicted from the characteristics of filled pauses, and that at least in Dutch, they could be treated as discourse markers, because their function is similar to expressions that are typically called discourse markers.

Other studies have focussed on the discourse structure of monologic talk. Chafe (1979), for example, notes that in monologues, a change in a major idea unit is often accompanied by dysfluencies. Schachter et al. (1991) argue that such dysfluencies at the beginning of topic units are related to the complexity of the subject matter. They show that lecturers in humanities use more filled pauses than do science lecturers due to the nature of the subject matter. They argue that in more formal structured and factual disciplines, because lecturers have fewer options available to them, they are less likely to use *ums* and *uhs*. Swerts and Ostendorf (1997) note that unit initial utterances are often more hesitant than others. Swerts and Geluykens (1994) demonstrate the role of prosody in the structuring of information in monologic discourse, as well as the role played by discourse markers at the beginning of a new topic. They also note that one of the speakers in their study uses *um* before each new topical unit (Swerts and Geluykens 1994: 26).

Clark and Fox-Tree (2002) have also raised the issue of whether *um* and *uh* should be referred to as fillers or as words. They argue that *um* and *uh* should be considered words because, as for any other word, they must be “planned for, formulated, and produced” (p. 75). However, they argue further that although they should be considered words, unlike most of the words produced by a speaker, *ums* and *uhs* are exclusively reserved for comments about the on-going performance of the utterance. In this paper, I hope to demonstrate that tokens such as *um* and *uh* are doing more than simply indicating performance difficulties. Instead the analysis will examine their position within the discourse and argue that they are playing a structuring role within the academic seminar.

The following example illustrates the frequency of *um* (bolded) in seminar talk.

(For transcription conventions, see Appendix.)

Eg 1 [Ar:1-2]

1. ...**um** ((moves slide up)) possible applications;
2. we've got basically anything that involves signals;
3. so; (.) °**um**° i've just written those down;

4. (2.0) ((moves slide up))
5. (1) → **U::M::** (1.0) SO (1.0) well firstly i suppose i should define wavelets themselves;
6. (1.0) **UM** (1.0) THEY form into two classes or two two classes of function;
7. one is the scaling function; (.) **um** which is used to represent the low frequency
8. information; of the signal; **UM** and the other is the wavelet. which is used to represent
9. the high frequency (.) **um** component. (.) **UM** what happens; these are translated into
10. scales; **um** then they're summed together; with uh with various coefficients,

11. °multiplied by them, to uh reproduce a function.°
- 12.
13. (1.5)
14. (2) → **U:M::z** (1.0) ((moves slide up)) NOW. i've just got an example here; of a particularly
15. useful application of wavelets; °**um**° which is the saw-tooth wave; (.) **um** the wave looks
16. like this; (9.0) ((draws on board)) **um**; and if you sample that at a rate of °**um**° (.) 256
17. uh samples per second; (.) **um**; then you will in fact need **um** 256 coefficients in your
18. frame expand to represent it. **um**; the wavelet that you get is in fact only sixteen
19. () which is the right wavelength; ((takes slide off)) °so it's a lot more
20. efficient.° ((puts slide on pile))
21. (2.0) ((picks up new slide))
22. (3) → **UM**; (1.5) ((puts new slide on OHP)) THE OTHER ISSUE is speed. wavelet transforms
23. **um** can be computed in order n cycles; **UM**; this- most transformers are computed in n
24. squared; or n squared cycles; and even the n squared transform **um** is order n log n. so
25. in fact this is **um** quite good; uh what multiple you've got before that n; °**um**° >depends
26. on the wavelet; < °so a complex **um** wavelet; will in fact require an order of something
27. big times n;°
28. (1.5)
29. (4) → **U::M**; ((moves slide up)) (1.0) t! OKAY. THE OTHER ONE IS uh sparcity; (.) **UM::z**
30. that issue is about basically (.) **um** how few co-efficients **um** you can represent in simple;
31. uh in signal; uh data, and uh °sparcity is good because it >means that you have
32. represented your entire signal, in very few **um** co-efficients; <° ...

Example 1 shows the seminar presenter using *um* in a variety of environments, with arrows 1-4 indicating the presence of *um* in one particular environment. The aim of this study is to analyse *um* in this and other environments. In so doing, it will argue that although traditionally *um* tends to be associated with repair or dysfluency of speech production, in certain environments within the academic seminar, *um* is functioning as a discourse marker. In other words, it is not simply functioning as a filler; in specific well-defined environments it is playing a signposting role of indicating to the audience the structural organisation of the talk as a whole. In order to demonstrate this structuring role, the paper will examine three environments in which *um* overwhelmingly occurs. The presence of *um* in these environments will also throw light on the issues of whether within the academic seminar, *um* and *uh* function similarly. The analysis will show that in contrast to *um*, *uh* is not found in these environments. Therefore it is suggested, as argued by Fox Tree (2001) and Swerts (1998), that the functions of *um* and *uh* are different.

2. Method

The data for this research originated from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in Canberra, Australia. A series of weekly seminar presentations organised by CSIRO's Division of Information Technology (DIT) were videoed and the first six native speaker presentations were transcribed using conversation analysis methods. The seminars were part of an in-house series of

seminars, where either DIT staff or visitors from other CSIRO Divisions or university departments gave seminars on research-in-progress. All seminars were given in the conversational style, with presenters ‘talking to the overhead’. The seminars ranged from 40 to 80 minutes in length. All presenters were male. They tended to use a large number of visual images, such as overhead slides, images on the computer, videos, models, and the whiteboard. Permission for the seminars to be videoed and analysed was obtained. As DIT seminars are often videoed, it is assumed that presenters were not adversely affected by the presence of the camera.

Conversation analysis (CA) is a useful tool for examining spoken discourse, with its emphasis on taping, transcribing and analysing naturally occurring data. The importance of analysing naturally occurring data cannot be overemphasised, for it is only by examining talk within the context of the activity that surrounds the talk that it is possible to yield the “technology of the conversation” (Sacks 1984: 413). There are two basic assumptions of CA (Heritage 1989). Firstly, CA argues that all interaction is structurally organised. This means that talk is ordered, rather than a series of random utterances. Secondly, CA emphasises the importance of the context in which a particular interaction takes place. As a result, *all* aspects of linguistic behaviour are important, and nothing can be dismissed as *a priori* disorderly, accidental or interactionally irrelevant. Such an approach is clearly useful in examining *um* in interaction, as it ensures that the analysis is not clouded by pre-existing ideas as to the role and function of such a token.

3. Discourse markers

The term ‘discourse marker’ typically refers to “a more or less open class of syntactically optional, non-truth-conditional connective expressions” (Schourup 1999: 242). However, classifying the class of discourse markers remains elusive, with lack of agreement as to what counts as a discourse marker, the nature of the connection discourse markers express, and the grammatical status of the discourse markers themselves (e.g. Fraser 1999; Schiffrin 1987; Schourup 1999). Thus although words such as *okay* and *so* would clearly be called discourse markers, tokens such as *um* and *uh* are seldom included in lists of what is considered a discourse marker. They have been specifically excluded by Fraser (1999: 942) as being simply fillers or pause markers. This contrasts with Swerts’ (1998) comment that at least in Dutch, fillers such as *um* and *uh* could be treated as discourse markers.

To date, discussions of discourse markers have mainly concentrated on how such tokens function in conversation (Schiffrin 1987; Jucker 1993), in closings (Schegloff and Sacks 1973; Button 1987; 1990), in classroom discourse (Sinclair and Coulthard 1975), in service encounters (Merritt 1984), in family decision-makings (Condon 1986), in meetings (Beach 1990) and in interpreting children’s stories (Segel, Duchan and Scott 1991). Emphasis has also been placed on analysing the role and function of individual discourse markers, such as *well* (e.g. Jucker 1993), *oh* (e.g. Fox Tree and Schrock 1999; Heritage 1984), *you know* (e.g. Macaulay 2002; Östman 1981; Stubbe and Holmes 1995), and *like* (e.g. Romaine and Lange 1991; Underhill 1998).

However, the role and function of discourse markers in academic monologues, such as seminars, conference papers or lectures has not been looked at in detail. Goffman (1981) mentions how footing is communicated through cues and markers in speech, and Chaudron and Richards (1986) and Flowerdew and Tauroza (1995) have carried out research on the effect of discourse markers on second language lecture comprehension. However, no detailed analysis of discourse markers in the academic monologue has been carried out (cf. Rendle-Short 1999; 2002; 2003).

Discourse markers such as *okay*, *so*, *now* and *well* are important structuring devices in the academic seminar (Rendle-Short 1999; 2003). They occur in specific, well-defined environments to inform the audience as to the unfolding structure of the talk. Analysis of seminar talk indicates that the talk is divided into smaller, more manageable sections. Presenters talk for a bit, they pause and then they talk for a bit more. Brown and Yule (1983) called these bits of talk paratones, in that they resemble the paragraph indentation of written text. However, the more neutral term of ‘sections’ will be used in this paper, in order to minimize associations with written discourse. Previous analysis shows that sections of talk display a number of characteristics (Rendle-Short 1999, 2002). They are surrounded by pauses. Talk at the beginning of a section tends to be more prominent than preceding talk. Such prominence is achieved by increased volume, raised pitch, marked inhalation, and dental clicks. This contrasts with talk at the end of a section which tends to be quieter and faster than surrounding talk. The discourse markers *okay* and *so* frequently occur at the beginning of a section of talk. In this position, they play an important signposting role of indicating to the listening audience how the subsequent discourse should be interpreted.

4. *Um* and *uh* in repair routines

Although to date *um* and *uh* have not been the subject of detailed research, they have been referred to in other contexts, such as their use in repair routines (Schegloff, Jefferson and Sacks 1977; Schegloff 1979), as hesitation fillers (Brown and Yule 1983: 15, 106), and as interjections (Fraser 1990: 391). Discussions of self-correction or repair indicate that *um* or *uh* can be used to initiate a repair following dysfluencies or hitches, such as cut-offs, pauses, sound stretches, or repetition (Schegloff, Jefferson and Sacks 1977: 367; Schegloff 1979: 272). Numerous examples demonstrating the way in which *um* and *uh* can indicate hesitation fillers or repair are to be found in the seminar talk data.

Repair can be evidenced by a cut-off.

Eg 2 [Ph:5]

1. ...uh in sydney there's uh a room_i↓ where uh a (1.0) um about a dozen operators sit and
2. look at the traffic, and when seats isn't doing uh quite as well as it might be doing, they
3. → jump in and manually uh change things, .h um_i **but (0.6) wor- uh traffic is** an incredibly
4. complicated thing, and it's just as easy to fowl something up, by ↓ interfering with it
5. manually...

The example shows how the presenter changes his mind about a particular word (line 3). He cuts the word short and then initiates a repair sequence with *uh* followed straight away by the corrected version, *traffic is an incredibly complicated thing*.

Repair can also be evidenced by pauses.

Eg 3 [Ma:8]

1. ...and you can't generate a new ((points to OHP)) scene in response to new input. >°data
2. → that you haven't seen before.°< and so **(0.4) um (1.0) .h y-you get (0.6) uh (0.8) you can't**
3. **generate scenes**, new scenes. *um* and you can sort of bend this a little to start »widening the
4. frames of each of the scenes, but >°then you end up sort of back here.°< ((points to top of OHP))

The presenter initially indicates uncertainty (line 2) by the presence of a pause, followed by *um* and another pause. He then breathes in, but is still unable to formulate a repaired version. It is only following a further pause that the repair, initiated by *uh*, is successful.

Repair can also be evidenced by sound stretches or elongation.

Eg 4 [Mi:9]

1. ...but, as the water is pushed out the end the mass of the rocket is getting less and less,
2. → so::, **um the uh the acceleration, sorry**. >the weight of the rocket is falling, and the gravity
3. force drops.< ↑>so it's quite an interesting little system of things going up and down.< ↓...

In this example, the presenter indicates uncertainty by saying an elongated *so::*, (line 2). An unsuccessful repair is initiated by *um*, then a further unsuccessful repair is initiated by *uh*. The repair sequence is closed when the presenter says *sorry* with falling intonation.

Repetition of words also feature in repair routines.

Eg 5 [Ph:2]

1. ...there was a car on the detector. ↓ *um*; the the little *um* (0.6) ↓ >rectangles
2. → that sit in front of *uh* the stop lines, **in um uh on the on the street**. *um* have a little metal
3. detector underneath it. and that's that's the only information that *uh* scats uses...

In this example, the presenter says the wrong preposition (line 2). A repair sequence, initiated by *um*, and then *uh*, follows.

The above examples (Egs 2 - 5) clearly demonstrate how both *um* and *uh* are involved in repair sequences. Repair sequences are quite distinctive in that they are associated with uncertainties and dysfluencies of speech production and often occur in conjunction with pauses. A break in the normal flow of talk occurs because the presenter is unable to continue the talk for some reason, and such discontinuity has to be dealt with before the presenter can proceed. The way in which *um* and *uh* are used in the repair sequence is quite characteristic. *Um* and *uh* appear to be interchangeable. They can co-occur within

the same repair sequence. They can also be repeated within the same repair sequence. They are generally said with no intonation contour; in particular, they are not said with rising intonation. They also tend to be said at the same volume as the surrounding talk.

However, such repair *ums* and *uhs* are in the minority in seminar talk (see Table 1). Analysis of the computer science seminars indicates that *um* and *uh* do not only function in repair sequences, they also play a more structural role within the talks. Not all instances of *um* and *uh* occur randomly throughout the talks and not all instances of *um* and *uh* are associated with dysfluencies and uncertainty of speech. The majority of occurrences (87%) of *um* and *uh* in seminar talk are found in quite specific, well-defined environments, and they display distinctive characteristics. Therefore to simply view *um* or *uh* as only being associated with repair sequences or as hesitation fillers, is to overlook an important aspect of their function in seminar talk. The following analysis will demonstrate in more detail how *um* functions as a discourse marker in seminar talk.

Presenters	<i>Um and uh in Seminar Talk</i>		
	Total	Repair device	Discourse marker
1		43	391 (90.1 %)
2		11	282 (96.2 %)
3		20	392 (95.1 %)
4		162	491 (75.2 %)
5		54	218 (80.1 %)
6		10	203 (95.3 %)
TOTAL		300	1977 (86.7 %)

Table 1: Number of times *um* and *uh* are used in seminars by the 6 presenters.

5. *Um* as discourse marker

Presenters overwhelmingly use discourse markers to structure their talk. For example, the discourse marker *okay* frequently occurs at the beginning of a section of talk. In this position it plays the double role of indicating that the previous bit of talk is complete and that the presenter is ready to move onto the next topic (Rendle-Short 1999; Beach 1993).

- (a) (5.5) kh:: ↑**OKAY**. but I'D LIKE TO TALK MORE ABOUT THE UHM (2.5) THE CONCEPTUAL level...
- (b) (2.0) ↑**oka:y**. what i'm going to be talking about today_i is the uh the tritram project.
- (c) (1.5) **OKAY**. HERE WE'VE USED soft thresholding. (.) uhm_i but the main difference between the two methods_i is the fact that uhm (.) i've decreased the thresholds.

- (d) (1.0) ↑**OKAY. THE::** (2.0) sort of uh (0.5) the sort of opposite viewpoint_i uh is characterised by (phil's) top down paradigm.

Discourse markers in this position display prosodic characteristics typical of talk at the beginning of a section; they often occur following a lengthy pause; they are often said more loudly than preceding talk; they may be said with raised pitch; they often have a distinctive intonation pattern; and they are often associated with non-verbal activities, such as putting slides on or off the overhead projector. The discourse marker *so* also occurs in this position, at the beginning of a section of talk (Rendle-Short 2003).

- (a) (1.0) **SO THAT'S THAT'S THE FIRST BIT**_i UHM_i THE SECOND THING YOU NEED_i
- (b) (2.0) ↑**so** that's one way of doing it_i (2.0) uhm_i that takes (0.3) well not very long,
- (c) (2.0) ↑**SO HOW DOES THE UH**↓ (2.5) how does the (0.5) model and the algebra help you do that_i
- (d) (1.5) ↑**so** how do we apply this to the rocket? ↑well in the rocket_i we start off up the top here_i

The two discourse markers *okay* and *so* can also work together to indicate to the audience that a new section is about to begin.

- (a) (4.5) khh:: ↑**OKAY. SO** (1.5) HE:::RE IS AN EXAMPLE_i (3.0)
- (b) (6.5) khh:: ↑**OKAY. SO** HERE'S ANOTHER EXAMPLE_i and this is a sort of constructive theory_i↓
- (c) (2.5) ↑**OKAY. SO**↓ (2.0) without any mo::re (1.5) .h introduction, i'm going to (.) put up the super model_i
- (d) (2.5) t! **okay. so** what is uh titram? ↑uhm_i it's a traffic simulation system.↓ so: what that means is that uhm_i (2.0)

To demonstrate the way in which *um* also functions as a discourse marker in seminar talk, it is necessary firstly, to examine more closely the specific environments in which the token *um* occurs. If *um* were only functioning as part of a repair routine, its occurrence would be random because, as noted by Schegloff, Jefferson and Sacks (1977: 363), nothing is excluded from the class of the repairable. In other words, *um* should be equally distributed at the beginning of a Turn Construction Unit (TCU), within a TCU or at the end of a TCU. However the following examples demonstrate that the occurrence of *um* in seminar talk is not random. Secondly, it is necessary to examine whether within these environments *um* displays features typical of repair routines, as exemplified in examples 2-5, or whether it displays the characteristic features of discourse markers, in terms of prosody and the way in which presenters are concurrently interacting with the overhead slides.

5.1. *Um at the beginning of a section of talk*

The earlier example (Eg 1) showed *um* occurring at the beginning of four sections of talk. It showed how *um* either occurs alone (arrow 3), or in combination with another discourse marker, such as *so*, *now*, or *okay* (arrow 1, 2, 4). In this position, at the beginning of a section, *um* assumes the characteristic features of this position: it occurs at the beginning of a TCU; it occurs following a pause; it is said with increased volume; and it is said with rising intonation. It is also associated with the speaker interacting with the overhead slides, either by moving them up (arrow 1, 2, 4) or by placing a new slide on the screen (arrow 3).

The following example from a different speaker demonstrates a similar picture.

Eg 6 [Ro:2]

1. Pres: ...°so hopefully this is a bit of a rehearsal for me, to get my thoughts together.°
2. (2.5)
3. → Pres: ↑t! **um** (1.0) ((stands and looks at OHT, hand on chin)) what i'll cover
4. ((points down at OHP)) briefly↓ today_i is just look at some of the g-i-s
5. requirements that we looked at. the g-i-s on the internet. **um**_i and then i'll look at
6. some of the systems that are currently available_i (1.0) **um**_i (1.5) t! then some of
7. the
8. formats_i that you might be (.) uh using if you're a developer. >°uh looking at°°
9. the way of developing °systems to deliver g-i-s over the internet_i° or the web_i<
10. (1.0) **um**_i ((reads from OHT)) some of those data types, one in particular actually
11. or one protocol ↓in particular that °we used for our work.° there's a couple of
12. them
13. floating around.↑ none of which are comprehensive. °all that comprehensive.°
14. ↑i'll look at handling some of the g-i-s data requirements.↓ >uh at the client end.<
15. given that the client's basically a web browser. one of the popular web browsers
16. that we know and love. ↑and then (1.5) uh lastly_i↓ i'll have a look at a couple (.)
17. i'll show you hopefully a couple of our demos that we've done. °in the last few
18. months.°
19. (8.0) ((takes slide off and puts on pile. organises next slide))
20. → Pres: .hh **um** (2.0) ((stands and looks at OHT))
21. Pres: look [ing at the require[ments for g-i-s_i (1.5) uh (1.0) that we were given_i=
22. Action: [picks up new slide [puts new slide on OHP
23. Pres: =um (0.8) ↑we've done a number a number of collaborative projects.↓ with with
24. °um° commercial partners. ...

Once again, the speaker commences the sections with *um*. Line 2 shows the speaker pausing for 2.5 seconds, before commencing a new section with a raised pitch ↑t! **um**_i (line 3). The presenter then stands and looks at the slide for 1.0 second, before telling the audience what the topic of the section is about. Similarly at line 17, the presenter pauses for 8.0 seconds while he organises his slides. He then breathes in, says *um* with a rising intonation before pausing for a further 2.0 seconds while looking at the slide (line 18). Following the pause, the presenter moves onto the topic of the next section.

The question is, what is going on here. The picture resembles the way in which more commonly accepted discourse markers, for example, *okay* and *so*, occur at the beginning of a section of talk.

Eg 7 [Ma:19]

1. (1.5)
2. → **um** (1.0) ((puts new slide onto OHP)) ↑and here maybe is a ↓(.) some algebra for a
3. generalised uh (1.0) uh property, generalised version of that; uh which is really just
4. defining a section »as i've said of pairs °which is a function model.° ((takes slide off))
5. °details aren't ((puts slide onto pile)) really important.°
6. (2.5)
7. t! ↑**OKAY. SO** ((puts new slide onto OHP)) ONE THING THAT I I HAVEN'T
8. MENTIONED YET, AND WHICH I THINK ((walks over to 2nd OHP)) IS ACTUALLY
9. REALLY IMPORTANT; uh is this arrow here. (.) ((points to OHP)) which is the
10. change of space. (.) ((walks back)) um (1.0) .h now what that's about, is uh as i sort of
11. alluded to earlier; it's not always the case that you find your ((points to OHP)) information
12. space here; we find a scene that's directly isomorphic to it and that's it as the end of your
13. display. it's often the case that you need to start in one information space and transform um
14. to a different space. in order to make sense. now here are some cases where that has to
15. happen; first of all if the uh information spaces out here; ((points to OHP)) where the
16. where
17. there aren't any isomorphic scene spaces. you have to find a mapping to one here; but often
18. the just the task that you want to perform; uh »requires a different structure than the one
19. that you started with. ↓ um and °so this change of space is important for that.° .h ↑um; it
20. also happens uh by accident, if you don't really understand the structure of your scene
21. space; and you think you're mapping to one space; when in fact you're mapping to
22. something else; ↓ and in that case we could sort of say well that's a change of space that
23. happened. °even though it was unintentional°
24. (6.0) ((takes off slide, puts onto pile and puts new slide onto OHP))
25. → ↑**okay. so** here's the sort of just a quick example of what a task based changed
26. space uh
27. might be like; um; suppose we had just a whole bunch of x y pairs; that we collected; in
28. some sort of experiment; sampling; we can ((points)) sort of look at that as a function from
29. some unstructured uh enumeration space to pairs; but what we really want ↓ is to transform
30. that to a function of two variables x and y. and a °sort of scattergram.° so (.) the (.) well the
31. (.) »the way ((takes slide off)) the data is collected isn't necessarily °the way ((puts onto
32. pile)) you want to wor- use it. i suppose.°
33. (2.0)

This example shows three sections, bounded by longish pauses, ranging from 1.5 seconds (line 1) to 6.0 seconds (line 23). The first short section commences with *um* (line 2); the second long section commences with the combination, *t! OKAY. SO* (line 7); the final section again commences with *Okay. so* (line 24). At the boundary of each section, the old slide is removed and the new one put on the overhead projector.

Both *okay* and *so* occur at the beginning of a section, following a pause, and in connection with changing the overhead slide. They are generally said with increased prominence, such as raised pitch or louder voice, characteristic of talk at the beginning of a section. Examples 1, 6, 7 show *um* functioning in exactly the same way. Therefore,

it could be argued that *um* is functioning as a discourse marker, because it occurs in the environment where discourse markers occur and it displays the prosodic features of discourse markers in this position. It is always said with rising intonation in this position. It clearly belongs to the beginning of the section because it is given prominence associated with section beginnings. It is clear that in this position *um* is not functioning as a repair device. There is no evidence of production trouble, no repetition of words, false starts, or sound stretches. *Um* is never repeated at the beginning of a section. It is not being used as part of a word search, because the presenter does not indicate that he is confused or uncertain about being able to continue. If it were simply being used as a repair device or a filler when the presenter was accountable in some way for not speaking, its occurrence would be random. This is not the case. It occurs systematically at the beginning of a section.

Although *um* is not being used as a repair device, it could be argued that it is being used as a hesitation filler, either for planning what to say next (Brown and Yule 1983) or to fill in or mark a pause (James 1974; Schourup 1985). Evidence from the data indicates that this is not the case. Firstly, in this position, *um* is always said with rising intonation, with increased prominence, and late in the total pause time. In addition, *um* is often preceded by an audible intake of breath or a dental click indicating that the speaker is about to commence the next section (Eg 6, line 3). In other words, *um* belongs to the beginning of the new section. If *um* were simply filling a pause between sections, there would be no evidence of it belonging to the new section. A filler *um* would be said more quietly than surrounding talk, without rising intonation, and would occur earlier in the total pause time, as the following example illustrates.

Eg 8 [Ma:2]

1. Pres: ...maybe by the end; *um* °some of it (0.3) ((picks up sheet)) will've soaked into
2. your subconscious.° ((puts sheet on pile))

3. → Pres: (1.0) °°**um**°° (1.0) [(1.0)=
4. Action: [picks up new slide

5. → Pres: =.h ↑SO FIR [ST OF ALL, (1.5) uh (.) what's it for; (.)=
6. Action: [puts new slide on OHP. title: "what's it for"
7. Pres: =t! *um*; (.) well really what it is, is a is a (.) uh:: (.) is an attempt to provide a
8. mathematics for↓ information display. a descriptive mathematical language. (.)
9. *um*; ↑but here are some of the things that (.) that a good model of some area
10. should do;↓ ((runs hand down list)) <and i'm hoping this one does; > ...

Line 3 shows a filler *um* said more quietly than surrounding talk, without rising intonation and earlier in the total pause time. Such an *um* does clearly not 'belong' to the new section, which commences with the discourse marker *so* at line 5.

Therefore we can conclude that when the louder, raised pitch, rising intonation *um* occurs in this position in the seminar, it is playing a function similar to other discourse markers which occur at the beginning of a section of talk. In other words, presenters are using *um* as a discourse marker to indicate the macrostructure of the talk to the audience by letting them know there is more talk to come. This type of *um* contrasts with the characteristic features of the repair *um*, which is associated with

dysfluencies and uncertainties, is said with no intonation contour, and is said either at the same volume as surrounding talk or more quietly than surrounding talk. In addition, the repair *um* often co-occurs with *uh* or is interchangeable with *uh*; whereas *uh* almost never occurs at the beginning of a section of talk.¹ The data shows that beginnings of sections in computer science talk are characterised by clear discourse, uncluttered by hesitations, uncertainties, stumblings or elongation of vowels.

5.2. *Um after the orientation phrase*

A second environment in which *um* is overwhelmingly found is immediately following the orientation phrase at the beginning of a new section of talk.

Eg 9 [Ma:19]

1. (6.0) ((takes off slide, puts onto pile and puts new slide onto OHP))
2. ↑okay. so here's the sort of just a quick example of what a task based changed space *uh*
3. → might be like_i **um**_i suppose we had just a whole bunch of x y pairs_i that we collected_i in
4. some sort of experiment_i sampling_i we can ((points)) sort of look at that as a function from
5. some unstructured *uh* enumeration space to pairs_i ...

The characteristic pattern at the beginning of a section is a discourse marker followed by an orientation phrase:

discourse marker(s) + orientation + topic talk

In line 2 of Example 9, the discourse markers *↑okay. so* are followed by the orientation phrase, *here's the sort of just a quick example of what a task based changed space *uh* might be like_i*. Analysis of the data shows that the orientation phrase is overwhelmingly followed by a rising intonation *um*_i (line 3). The presenter then moves onto further topic talk. In this example, the move to further talk occurs without a pause, although pauses between the orientation phrase and the commencement of the actual topic of the section may occur (see Eg 8).

A further example illustrates a similar occurrence.

Eg 10 [Mi:5]

1. ...and we think that's >probably a problem with this ((points to one on demo table))
2. one,< where °they're not really far enough down.°
3. (1.5)
4. OKAY. AND FINALLY >YOU NEED SOME SORT OF AN AIR COMPRESSOR.<
5. → **um**_i you can use that hand pump if you want to, but it >rapidly ceases to be a joke.< an
6. electric air compressor's is one form_i, again the >people from the energy research centre
7. have excelled themselves_i ...

¹ The data shows 110 instances of *um* occurring at the beginning of a section of talk. This contrasts with one instance of *uh* in this position.

This example shows the discourse marker *okay* occurring at the beginning of the section of talk (line 4), following a 1.5 second pause. The presenter then continues, in a louder voice, to indicate the topic of the section, *AND FINALLY >YOU NEED SOME SORT OF AN AIR COMPRESSOR.<* This is immediately followed by *um*_ç, said with rising intonation (line 5).

It is clear therefore that *um* in this position is not associated with dysfluencies or uncertainties in the presenter's talk. There is no evidence of repair in the form of cut-offs, elongations, or repetition. This environment is the most frequent place where *um* occurs.²

- (a) (1.0)
SO THAT'S THAT'S THE FIRST BIT_ç UHM_ç THE SECOND THING YOU NEED_ç is ↑you need a thing to go in the end. a plug. (2.0) **uhm**_ç and_ç if you're going to withstand a 120 p-s-i, it has to be pretty tight.↓ otherwise you get leaks.
- (b) (2.5)
t! okay. so what is uh titram? ↑**uhm**_ç it's a traffic simulation system.↓ so: what that means is that uhm_ç (2.0) we've set it up so that it uhm mimics the behaviour,
- (c) (1.5) OKAY. HERE WE'VE USED soft thresholding. (0.5) **uhm**_ç but the main difference between the two methods_ç is the fact that uhm (.) i've decreased the thresholds.
- (d) (1.5)
OKAY. AND FINALLY >YOU NEED SOME SORT OF AN AIR COMPRESSOR.< **uhm**_ç you can use that hand pump if you want to, but it >rapidly ceases to be a joke..<
- (e) (2.0)
OKAY. THE ONE method that pretty much everyone starts out with_ç uhm_ç with the wavelets, is the highway version. **UHM**_ç THE BASIC IDEA is you've got the sequence_ç.....

As with *um* at the beginning of a section of talk, *um* in this position occurs at the beginning of a TCU, prior to the presenter commencing topic talk. It may or may not be accompanied by pauses. In this position *um* always occurs with rising intonation. It may or may not be given prominence, depending upon the level of prominence given to the rest of the utterance at this point in the talk.

5.3. *Um as each point on the slide is mentioned*

The final environment in which *um* is to be found is when it is used to mark points on an overhead slide.

Eg 11 [Ro:1]

1. (2.5)
2. → t! ↑**um**_ç (1.0) ((stands and looks at OHT, hand on chin))
3. what i'll cover briefly↓ today_ç is just look at some of the g-i-s requirements that we

² Out of a total number of 2277 instances of *um* in the data, *um* occurs in this position 786 times (34.5%).

4. looked at. the g-i-s on the internet.
5. → **um** and then i'll look at some of the systems that are currently available;
6. → (1.0) **um** (1.5) .t! then some of the formats that you might be (0.3) uh using if you're a
7. developer. >uh looking at the way of developing systems to deliver g-i-s over the
8. internet or the web<
9. → (1.0) **um** ((looks at OHT)) some of those data types, one in particular actually or one
10. protocol ↓in particular that we used for our work. there's a couple of them floating
11. around. ↑none of which are comprehensive. °all that comprehensive.°
12. → ↑i'll look at handling some of the g-i-s data requirements. ↓ >uh at the client end.<
13. given that the client's basically a web browser. one of the popular web browsers that we
14. know and love.
15. → ↑and then (1.5) **uh** lastly ↓ i'll have a look at a couple (0.5)
16. i'll show you hopefully a couple of our demos that we've done. °in the last few months.°
17. (8.0) ((takes slide off and puts on pile. organises next slide))

This example (a repeat of Eg 6) shows the presenter near the beginning of his presentation telling the audience what his talk will be about. An overhead slide with 6 points is projected onto the screen as he talks. As he mentions each new point on the overhead, he marks the talk in some way. The first 4 points are marked with a rising intonation *um* (lines 2, 5, 6, 9). The fifth point, however, is not marked by an *um* (line 12), although the point is still marked by the presenter raising the pitch of the talk. The final point is more overtly marked as the final point in a list, *↑and then (1.5) uh lastly ↓* (line 15).³

The following example similarly shows a different presenter marking all the points on the overhead with *um*.

Eg 12 [Ar:2]

1. (2.0)
2. → .h **UM** AND SO the better ((points at OHP)) the velocity can pressure;
3. → (2.0) **UM::** what happens when you compress the signal (1.0) *um* is that you basically
4. ((points at OHP)) take the transform be it a cosine or wavelet and then store >>those
5. coefficients or send those coefficients.<<°
6. → (1.5) **UM** and of course. the wavelet ((points at OHP)) that you press is better,
7. °cos almost all of the information is contained in very few co-efficients.
8. → (2.0) **UM** ((moves slide up)) and finally- or another application, sorry. is noise removal;
9. → (1.0) **UM** this is in fact related to the sparsity issue.
10. → (1.0) **UM** ((moves slide up)) because the essential signal in the wavelet domain, is
11. represented in so few coefficients.
12. (1.0) **UM** whereas noise is spread °um° over very many small co-efficients.
13. if you cut out all the small coefficients,
14. → (1.5) **UM** ((moves slide up)) you uh in fact get uh a very *um* (0.4) good de-noised image.
15. **UM** the advantage of the wavelets is that they don't merge sharp edges. and so if you've
16. got for instance ((goes to blackboard)) a nuclear magnetic resonance in () °sort of
- like
17. this, (2.0) ((writes on blackboard)) and you use traditional () methods to get that,

³ Jefferson (1990) discusses the way in which speakers and listeners orient the three-part nature of lists. Therefore choosing not to mark every point on the list with *um*, as in Eg 11, could be orienting to the fact that once it is clear that a list is underway, the need to overtly indicate each point becomes less necessary.

18. () whereas *um* the wavelets won't ((walking back to OHP)) in fact do that.^o
19. (2.0)

In this example, the presenter marks each point on the overhead slide with a rising intonation *um*. The *ums* are also said slightly louder than surrounding talk.

One final point to note is that each time this presenter indicates he is talking about a different point on the overhead slide, he also moves the overhead slide up. Therefore the audience is left in no doubt as to the fact that he is talking about the next point on the slide. Thus the structure of the talk is made apparent to the audience, not only through the use of tokens such as *um*, and the way in which they are given additional prominence within the talk, but also through the co-occurring nonverbal activities. By using a raised pitch, rising intonation *um* as he adjusts the slide, the audience is invited to focus on the next point. Thus the structure of the talk is made apparent.

6. Conclusion

Discourse markers are overwhelmingly used in computer scientist seminar talk as a way of signposting, of indicating the structure of the talk to the listening audience. They play an important role in indicating the beginning of each section of talk, in ensuring that the talk is presented as a coherent piece of spoken discourse. The most frequently occurring discourse markers in seminar talk are *okay* and *so*. Such discourse markers occur in specific, well-defined environments and display characteristic prosodic features.

An important premise underlying the analysis of talk, both in everyday and in institutional settings, is that talk is highly ordered and structured. As a result, nothing can be dismissed as being accidental or unworthy of attention. This is most clearly demonstrated in the above analysis of *um*, a token that would appear to lack semantic content, yet in the institutional setting of the seminar presentation can be shown to function as a discourse marker. It is only through detailed analysis of naturally occurring data that it is possible to determine how *um* actually functions in seminar talk, and to dispel some of the myths surrounding tokens like *um*.

The analysis has shown that *um* systematically occurs in predictable environments and that in these environments there is no evidence of dysfluency or uncertainty of speech production. It has been argued that in this role *um* is functioning as a discourse marker. This is firstly because *um* occurs in the environments where discourse markers occur. It occurs at the beginning of a section of talk; it overwhelmingly occurs following the orientation phrase; and it occurs as presenters discuss points on an overhead slide. In these environments, *um* indicates to the audience that there is more talk to come. Secondly, *um* resembles the way in which discourse markers function, in that it resembles the way in which discourse markers are characterised by specific prosodic features. Rising intonation, increased volume, or raised pitch ensure that at the beginning of a section *um* is made more prominent than surrounding talk. This contrasts with a general lack of specific prosodic features evident when *um* occurs as part of a repair routine. Thirdly, as for other discourse markers

(Rendle-Short 2002), the saying of *um* can be integrated with non-verbal activities, such as moving the overhead slide or putting it onto the projector.

Thus the analysis has shown that although *um* can indicate production problems within spoken discourse, this is not its only function. In this respect, *um* seems to function differently from *uh*. Whereas both *um* and *uh* can occur in instances of repair, only *um* shows the underlying structure of the academic seminar through its occurrence in specific well-defined environments.

Appendix: Transcription conventions

.	a stopping fall in tone, not necessarily the end of a sentence
,	continuing intonation, not necessarily between clauses of sentences
?	rising inflection, not necessarily a question
˘	rising intonation weaker than that indicated by a question mark
-	cut-off
t!	dental click
=	connecting talk
> <	talk is faster than surrounding talk
< >	talk is slower than surrounding talk
◦ ◦	a passage of talk that is quieter than surrounding talk
SO	a passage of talk that is louder than surrounding talk
↓↑	marked falling and rising shifts in pitch
::	an extension of a sound or syllable
()	transcription doubt
(())	analysts comments
(1.0)	time intervals
(.)	a short untimed pause
»	talk that becomes gradually softer and faster, usually at the end of a section
hh	audible aspirations
.hh	audible inhalations
<u>so</u>	emphasis
[]	overlapping utterances or actions
[]	
—>	a marker to indicate something of importance
okay	bold type to emphasise important words
OHP	overhead projector
Pres:	presenter
Action:	line in the transcript indicating what the presenter is doing while he talks

Transcription conventions are based on Gail Jefferson's notation in Sacks et al. (1974), Atkinson and Heritage (1984), Button and Lee (1987).

References

- Atkinson, J. Maxwell, and John Heritage (eds.) (1984) *Structures of Social Action: Studies in Conversational Analysis*. Cambridge: Cambridge University Press.
- Beach, Wayne (1990) Language as and in technology: Facilitating topic organisation in a videotex focus group meeting. In: M.J. Medhurst, A. Gonzalez, and T.R. Peterson (eds.), *Communication and the Culture of Technology*. Pullman, WA: Washington State University Press. pp. 197-220.
- Beach, Wayne (1993) Transitional regularities for 'casual' "Okay" usages. *Journal of Pragmatics* 19: 325-352.
- Brennan, Susan, and Michael F. Schober (2001) How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language* 44: 274-296.
- Brennan, Susan E., and Maurice Williams (1995) The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language* 34: 383-398.
- Brown, Gillian, and George Yule (1983) *Discourse Analysis*. Cambridge: Cambridge University Press.
- Button, George (1987) Moving out of closings. In G. Button, and J.R.E. Lee (eds.), *Talk and Social Organisation*. Clevedon: Multilingual Matters. pp. 101-151.
- Button, George (1990) On varieties of closing. In G. Psathas (ed.), *Interaction Competence: Studies in Ethnomethodology and Conversation Analysis*. Washington, D.C.: International Institute for Ethnomethodology and Conversation Analysis and University Press of America, pp. 93-147.
- Button, George Lee, and R.E. John (eds.) (1987) *Talk and Social Organization*. Clevedon: Multilingual Matters.
- Chafe, Wallace L. (1979) The flow of thought and the flow of language. In T. Givón (ed.), *Syntax and Semantics: Discourse and Syntax Vol 12* New York: Academic Press.
- Chaudron, Craig, and Jack C. Richards (1986) The effect of discourse markers on the comprehension of lectures. *Applied Linguistics* 7.2: 113-127.
- Chomsky, Noam (1965) *Aspects of the theory of syntax*. Cambridge MA: MIT Press.
- Christenfeld, Nicholas (1995) Does it hurt to say um? *Journal of Nonverbal Behaviour* 19.3: 171-186.
- Clark, Herbert H. (1994) Managing problems in speaking. *Speech Communication* 15: 243-250.
- Clark, Herbert H., and Jean E. Fox-Tree (2002) Using uh and um in spontaneous speaking. *Cognition* 84: 73-111.
- Condon, Sherri (1986) The discourse functions of OK. *Semiotica* 60: 73-101.
- Cruttenden, Alan (1997) *Intonation*. 2nd edn. Cambridge: Cambridge University Press.
- Flowerdew, John, and Steve Tauroza (1995) The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition* 17: 435-458.

- Fox-Tree, Jean E. (1995) Effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 34: 709-738.
- Fox-Tree, Jean E. (2001) Listeners' uses of um and uh in speech comprehension. *Memory and Cognition* 29.2: 320-326.
- Fox-Tree, Jean E. (2002) Interpreting pauses and ums at turn exchanges. *Discourse Processes* 34.1: 37-55.
- Fox Tree, Jean E., and Josef C. Schrock (1999) Discourse Markers in spontaneous speech: Oh what a difference an *oh* makes. *Journal of Memory and Language* 40: 280-295.
- Fraser, Bruce (1990) An approach to discourse markers. *Journal of Pragmatics* 14: 383-395.
- Fraser, Bruce (1999) What are discourse markers? *Journal of Pragmatics* 31: 931-952.
- Goffman, Erving (1981) *Forms of Talks*. Oxford: Basil Blackwell.
- Heritage, John (1984) A change of state token and aspects of its sequential placement. In: J.M. Atkinson, and J. Heritage (eds.), *Structures of Social Action: Studies in Conversational Analysis*. Cambridge: Cambridge University Press: pp. 299-345.
- Heritage, John (1989) Current developments in conversational analysis. In D. Roger, and P. Bull (eds.), *Conversation*. Clevedon Philadelphia: Multilingual Matters, pp. 9-47.
- James, Deborah (1974) The syntax and semantics of some English interjections. University of Michigan dissertation. University of Michigan Papers in Linguistics 1.3.
- Jefferson, Gail (1974) Error correction as an interactional resource. *Language in Society* 2: 181-199.
- Jefferson, Gail (1990) List construction as a task and resource. In G. Psathas (ed.), *Interaction Competence: Studies in Ethnomethodology and Conversation Analysis*. Washington, D.C.: International Institute for Ethnomethodology and Conversation Analysis and University Press of America, pp. 63-92.
- Jucker, Andreas H. (1993) The discourse marker well: A relevance-theoretical account. *Journal of Pragmatics* 19: 435-452.
- Macaulay, Ronald (2002) You know, it depends. *Journal of Pragmatics* 34: 749-767.
- Maclay, Howard, and Charles E. Osgood (1959) Hesitation phenomena in spontaneous English speech. *Word* 15: 19-44.
- Merritt, Marilyn (1984) On the use of O.K. in service encounters. In J. Baugh, and J. Sherzer (eds.), *Language in Use: Readings in Sociolinguistics*. New Jersey: Prentice-Hall, Inc., pp. 139-147.
- Östman, Jan-Ola (1981) *You know: A Discourse Functional Approach*. Amsterdam: John Benjamins.
- Rendle-Short, Johanna (1999) When 'okay' is okay in computer science seminar talk. *Australian Review of Applied Linguistics* 22.2: 19-33.
- Rendle-Short, Johanna (2002) Talk and action in the computer science seminar. Unpublished Ph.D. Thesis, Department of Linguistics, ANU.

- Rendle-Short, Johanna (2003) So what does this show us: Analysis of the discourse marker 'so' in seminar talk. *Australian Review of Applied Linguistics* 26: 2.
- Romaine, Suzanne, and Deborah Lange (1991) The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66: 227-279.
- Sacks, Harvey (1984) On doing "being ordinary". In J.M. Atkinson, and J. Heritage (eds.), *Structures of Social Action: Studies in Conversational Analysis*. Cambridge: Cambridge University Press, pp. 413-429.
- Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson (1974) A simplest semantics for the organisation of turn-taking in conversation. *Language* 50.4: 695-735.
- Schachter, Stanley, Nicholas Christenfeld, Bernard Ravina, and Frances Bilous (1991) Speech dysfluency and the structure of knowledge. *Journal of Personality and Social Psychology* 60.3: 362-367.
- Schegloff, Emanuel A. (1979) The relevance of repair to syntax-for-conversation. In T. Givón (ed.), *Syntax and Semantics: Discourse and Syntax* Volume 12. New York: Academic Press, pp. 261-286.
- Schegloff, Emanuel A. (1982) Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In D. Tannen (ed.), *Analysing Discourse: Text and Talk*. Georgetown Roundtable on Languages and Linguistics. Washington, DC: Georgetown University Press, pp.71-93.
- Schegloff, Emanuel A., and Harvey Sacks (1973) Opening up closings. *Semiotica* 8: 289-327.
- Schegloff, Emanuel A., Gail Jefferson, and Harvey Sacks (1977) The preference for self-correction in the organization of repair in conversation. *Language* 53.2: 361-382.
- Schiffrin, Deborah (1987) *Discourse Markers*. Cambridge: Cambridge University Press.
- Schourup, Lawrence (1985) *Common Discourse Particles in English Conversation*. New York: Garland.
- Schourup, Lawrence (1999) Discourse Markers. *Lingua* 107: 227-265.
- Segel, Erwin M., Judith F. Duchan, and Paula J. Scott (1991) The role of interclausal connectives in narrative structuring: Evidence from adults' interpretations of simple stories. *Discourse Processes* 14: 27-54.
- Sinclair, John, and Malcolm Coulthard (1975) *Towards an Analysis of Discourse: The English used by teachers and pupils*. Oxford: Oxford University Press.
- Stubbe, Maria, and Janet Holmes (1995) *You know, eh* and other 'exasperating expressions': An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language and Communication* 15: 63-88.
- Swerts, Marc (1998) Filled pauses as markers of discourse structure. *Journal of Pragmatics* 30: 485-496.
- Swerts, Marc, and Ronald Geluykens (1994) Prosody as a marker of information flow in spoken discourse. *Language and Speech* 37.1: 21-43.
- Swerts, Marc, and Mari Ostendorf (1997) Prosodic and lexical indications of discourse structure in human-machine interactions. *Speech Communication* 22: 25-41.
- Underhill, Robert (1998) *Like* is, like, focus. *American Speech* 63: 234-246.