

InLiSt No. 1

Interaction and Linguistic Structures

Coherent Voicing. On Prosody in Conversational Reported Speech

Elizabeth Couper-Kuhlen

Konstanz
January 1998

In recognition of the enthusiasm he has brought to all aspects of the study of spoken verbal interaction, we dedicate this series to Professor Dr. Aldo di Luzio, University of Konstanz.

Prof. Dr. Elizabeth Couper-Kuhlen
Dr. Susanne Günthner
Universität Konstanz
FG Sprachwissenschaft
Postfach 55 60
D-78457 Konstanz

Prof. Dr. Margret Selting
Universität Potsdam
Institut für Germanistik
Postfach 60 15 53
D-14415 Potsdam

Prof. Dr. Peter Auer
Universität Hamburg
Germanisches Seminar
Von-Melle-Park 6
D-20146 Hamburg

Additional copies may be ordered from:
Universität Konstanz, Sekretariat LS Angl. Sprachwissenschaft, InLiSt
Postfach 55 60, D 180, D-78457 Konstanz, Tel.: +49-7531-88-2552, fax: +49-7531-88-4157
<http://www.uni-konstanz.de/FuF/Philo/Sprachwiss/home/couperku/>

Coherence should be findable for everything that is a demonstrably relevant aspect of the talk for the parties, or there should be evidence of trouble or of its suppression. (Schegloff, 1990)

Goffman has pointed out that interlocutors in the course of any natural conversation are constantly changing the footing of their talk. In Goffman's usage, this term refers to the alignment which speakers take up to themselves and to others as evidenced by the way they handle the production and reception of utterances (1981:128). Changes in footing may involve different reception roles or different production roles or both (Goffman 1981:226ff; also Levinson 1988), and they are commonly understood to be signaled inter alia by prosodic cues and code-switching, which contextualize the particular footing or participant framework currently relevant (Gumperz 1982, Tannen, ed. 1993). Yet precisely how this contextualization is accomplished and what specific contribution prosody makes to the 'management' of footing has not yet been fully spelled out¹ — at least not for all types of shift.

The present paper addresses one of the most frequent shifts of footing, namely that occasioned by the use of reported speech in conversation. What happens with reported speech is that the unity within a single speaker of the three production roles which Goffman identifies — animator, author and principal — dissolves, leaving the role of animator separate from, and independent of, those of author and/or principal. The 'reporting' speaker animates or voices a 'reported' figure without necessarily composing the words which this figure is made to utter or espousing the beliefs which the figure's words will be heard as attesting to.² The question which the 'voicing' of figures raises for a prosodist is whether and to what extent the speaker's phonatory voice is instrumental in the process.³ Using a methodology developed by crossing prosodic analysis with conversation analysis (Couper-Kuhlen/Selting 1996), this paper attempts to pin down exactly which tasks the 'voicing' of reported speech confronts conversationalists with and how speakers' prosodic and paralinguistic voice resources contribute to the accomplishment of these tasks.

¹ The collection of articles in Auer&di Luzio, eds. 1992, takes one step in this direction.

² This description is intended to be general enough to apply to all forms of reported speech, including canonical direct as well as canonical indirect speech.

³ To keep the two notions apart, I shall use scare quotes around *voice* when reference is to the animation of a figure; *voice* without scare quotes refers to phonatory voice.

Coherence as a conversationalist's practice and an analyst's object

Schegloff has suggested that in conversation the issue of coherence can be subsumed under the general question *Why that now?* (1990:55). In other words, participants in interaction are constantly trying to make sense of talk as recipient-designed, situated action. When they are unable to infer plausible answers to the question *Why that now?*, they have sets of methods which allow them to remedy the situation, one of these sets involving the initiation and execution of 'repair'. Remedial procedures help clarify the misunderstood or the misunderstandable, on occasion they make explicit the unexplicit (see also Schegloff 1996). But remedial procedures also provide analysts with an invaluable instrument of analysis. It is via conversationalists' pursuit of coherence that analysts can learn more about the object from an insider perspective.

Observations such as these on coherence in interaction suggest a way to approach the relation between prosody and reported speech. Coherence in reported speech sequences,⁴ it can be argued, will be manifestly lacking where participants in interaction find repair to be necessary. When 'troubles' in coherence can be plausibly reconstructed as involving some prosodic or paralinguistic factor, insight will be gained into the specific nature of prosody's contribution to reported speech. A subsequent comparison of repaired and repairable reported speech sequences with non-repaired and non-repairable ones will suggest some of the methods which participants employ for the prosodic animation of voices.

Three types of 'trouble' in reported speech sequences

That the use of reported speech places special demands upon conversationalists is evident from the 'trouble' it sometimes occasions for interaction. One type of trouble appears to stem from unclarity as to whether or not a speaker is reporting speech in the first place. These are cases in which there is no clear answer to the question *Is this current speaker's 'voice' or someone else's?* The reason why there might be doubt about this is because — contrary to what grammar books propagate about reported speech — speakers in conversational interaction do not always explicitly introduce different 'voices' with reporting verbs or quotative constructions. Instead figures are often 'brought on stage' for the first time merely by being animated,

⁴ I use the term *reported speech sequence* for any part of a conversational exchange in which the presence of reported speech can be identified.

without, for instance, a prefatory *he said* or *she said*. In order for this device to succeed, however, the figure's 'voice' must be reconstructibly different from the current speaker's own 'voice'. The following sequence demonstrates that participants are aware of this distinction and that prosody may be a means for marking it:

(1) Toxic chocolate (38/22.16)

Ann and a girlfriend are spending the weekend as guests at Joy's house. The following exchange takes place over lunch.

(The small f signals that a reported figure is being voiced somewhere within the line, different indices representing different figures. The hyphen signals a return to the reporting situation.)

1 J: oh and on Sunday,
you can open the first door of the advent calendar.
All: aahh oooh
J: <<p>I never had one before.>
5 A: did you buy it?
J: mhm,
A: you didn't buy the kind with chocolates in it?
J: no.
All: huh huh huh huh
10 J: f₁ I didn't think I needed any <<l> chocolate
(thing).>
f₁ it's a <<l> to:xin you know.>
A: - who're you talking to (.)
talking about.
J: there-
15 f₁ <<l> chocolate is toxic.>
-> A: - yeah but you said that like somebody says that.
J: no that's the (.) cancer pre-
cancer pre- [ventative type.
A: [aaah.
20 J: I honestly cannot fathom; (.)
following that diet;
just to prevent cancer.

Joy presents opening up the first door of the advent calendar as a special treat for her guests (lines 1-2). In fact, even the calendar itself takes on special status when Joy claims that it is her first (line 4). Ann expands on this topic by asking whether Joy bought the calendar. (Buying an advent calendar is to be understood as contrasted with making one, a custom common in Germany, where the conversation takes place.) When Joy acknowledges that she did buy it, Ann expands the sequence again by asking *you didn't buy the kind with chocolates in it* (line 7), subtly implying that this kind of calendar might be more desirable. Thus when Joy now states *I didn't think I needed any chocolate (thing)* and *it's a toxin you know*, she is heard as justifying her

purchase, an activity prompted by Ann's treatment of buying a calendar without chocolates as remarkable and therefore 'accountable'.

Joy's accounts are couched as a warning implying that chocolate is a health risk (*chocolate (thing). it's a toxin you know*, lines 10- 11). This warning, however, is presented in someone else's voice, as Ann herself observes: *you said that like somebody says that* (line 16). Yet Joy's 'voicing' in lines 10ff manifestly creates a problem for Ann, because she initiates repair immediately thereafter: *who're you talking to (.) talking about* (lines 12-13). Note that it is not the content of the utterance which is at issue here. It is true that Joy initially treats the problem for which repair is initiated as referential in nature. She responds with a partial repeat, making the reference of *it* explicit: *chocolate is toxic* (line 15). However, Ann's *yeah but* in next turn (line 16) shows that what was said is not the point. The issue is the way the utterance was said. As Ann puts it, the utterance 'sounded like' someone else. What is her interpretation based on?

Deconstructing Ann's observation leads to a consideration of the nature of vocal deixis. Prosodic and paralinguistic effects are in fact deictic to a certain extent: they involve speaking within a given range of relative loudness, pitch and tempo (Laver 1994) and with a given voice quality (Laver 1990). In the default case, loudness, pitch and tempo ranges together with voice quality are anchored to the prosodic/paralinguistic *habitus* of the speaker. That is, speakers are accustomed to deploy, and their interlocutors are accustomed to expect, certain prosodic and paralinguistic 'reference values'. A noticeable shift of these values - using a pitch, loudness or tempo range or a voice quality which departs from the speaker's *habitus* - will be heard as shifted deixis and can evoke the presence of a second deictic center. It is precisely this kind of shift which is hearable in Joy's *chocolate (thing)* (line 10), *toxin you know* (line 11) and *chocolate is toxic* (line 15): in each phrase she drops into low pitch register shading off into a final 'vocal fry' or glottal creak. Fragment (1) thus provides demonstrable evidence of the fact that a figure can be 'voiced' by the way in which an utterance is configured prosodically and paralinguistically. At the same time it suggests that the question underlying coherence must be expanded to *why that now and in that way?*

Given the deictic nature of prosodic and paralinguistic phenomena, 'troubles' in reported speech sequences are to be expected when prosodic and paralinguistic cues are ambiguous as to whose vocal deictic center they index. This is arguably what creates the necessity for repair initiation in the following sequence:

(2) Political contradiction

Bill is telling his girlfriend Gina about a 'real political' discussion which he had on the beach with someone he had just met

1 B: You know and it's so funny cause he's-
he's a Catholic?
(.) and::
5 you know
like I nailed him on the contradiction;
he's like pro-capital punishment,
but- and pro life,
hhh
10 G: uh huh,
B: f₁ I said try to explain that to me.
f₁ I don't underst(h)a(h)nd(h).
f₁ heh heh
G: - (wait) you said he's:
15 pro:: capital punishment -
and pro life,
G: and pro life.
ri:::ght.
B: that's like
20 to me that's like a hu::ge contradiction?
(.)
and he said
he- he justified it as (.)
?f₂ well see
25 ?f₂ th- those little inf-
?f₂ the little (.) uhm
?f₂ embryos?
(.)
?f₂ they don't have a (.) decis-
30 ?f₂ they can't make a decision.
?f₂ you know
?f₂ ('n) they're gonna die.
-> G: - this is what he was saying.
B: yeah
35 and he said that you know
f₂ people who (.) commit crimes:
?- you know
f₂ that (.) call(s) for the death penalty;
f₂ they brought it upon themselves.
40 G: - oh wait
just
can you hang on just a sec?
B: yeas

In recounting how he 'nailed' his interlocutor on the political contradiction between supporting capital punishment but being against abortion, Bill re-enacts the incident as a dialogic exchange, animating himself as a figure: *try to explain that to me. I don't underst(h)a(h)nd(h). heh heh* (lines 11-13). What his interlocutor's response to this

request is, however, is not immediately clear. A response is projected with the reporting verb *and he said* (line 22) and then with another quotative construction *he-he justified it as* (line 23). Both constructions announce immediately upcoming talk as reported speech.⁵ But Bill's next utterances have no prosodic or paralinguistic shift of voice. Moreover, they deploy the discourse markers *well see* (line 24) and a high rising intonation pattern following *embryos?* (line 27), which characteristically elicits an in situ response from one's interlocutor.⁶ Thus, one interpretation this talk can be given is that it is part of the reporting situation, addressed to Gina as background information or as an aside. But this interpretation stands in direct conflict with the original framing of upcoming talk as reported speech (lines 22-23).

It is arguably the conflicting signals of quotative introduction but lack of vocal deictic shift (or any other marker clearly framing the talk as reconstructed) which prompt Gina's repair initiation *this is what he was saying* (line 32). The fact that *he* is stressed here suggests that it is to be interpreted as standing in contrast with other possible members of a set, in this case with *you*. Thus the format of Gina's repair initiation supports the interpretation that her problem is knowing whether this is some background commentary by Bill the narrator (part of the reporting situation) or his animation of a figure in the story (part of the reported situation).

'Trouble' in reported speech sequences occurs not only when there is no straightforward answer to *Is this current speaker's 'voice' or someone else's?*. It may also occur if the answer to the question *Whose 'other voice' is this?* is unclear, as example (1) reminds us:

(1) Toxic chocolate (38/22.16)

1	J:	oh and on Sunday, you can open the first door of the advent calendar.
	All:	aahh oohh
	J:	<<p>I never had one before.>

⁵ Which canonical form is being projected is unclear. The quotative *he said* is compatible both with the presence of expressive elements (canonical direct reported speech) and with their absence (canonical indirect reported speech), whereas the quotative *he justified it as*, canonically speaking, projects upcoming talk in which expressive elements are absent.

⁶ The noticeable pause following *embryos?* (line 28) is thus attributable to Gina. Her silence here may be a first indication of the problem which Bill's conflicting signals are creating for her.

f₁ heh heh
 -> G: - (wait) you said he's:
 -> 15 pro:: capital punishment -
 B: and pro life,
 -> G: and pro life.
 -> ri:::ght.
 B: that's like
 20 to me that's like a hu::ge contradiction?
 (.)
 and he said
 he- he justified it as (.)
 ?f₂ well see
 25 ?f₂ th- those little inf-
 ?f₂ the little (.) uhm
 ?f₂ embryos?
 (.)
 ?f₂ they don't have a (.) decis-
 30 ?f₂ they can't make a decision.
 ?- you know
 ?f₂ ('n) they're gonna die.
 G: - this is what he was saying.
 B: yeah
 35 and he said that you know
 f₂ people who (.) commit crimes:
 ?- you know
 f₂ that (.) call(s) for the death penalty;
 f₂ they brought it upon themselves.
 40 G: - oh wait
 just
 can you hang on just a sec?
 B: yeas

In re-enacting the exchange he has had with his new acquaintance, Bill first introduces himself as a figure in the story and animates this 'voice': *try to explain that to me. I don't underst(h)a(h)nd(h) heh heh* (lines 12-14). The animation is noticeable first by shifts in personal and temporal deixis: Bill's friend becomes *you*, to whom the request *try to explain that to me* is directed, and the speaker's account for this request *I don't understand* is anchored to the moment of speaking in the reported situation. But there is also a sign of shifted vocal deixis in the breathiness superimposed on the figure's account *I don't underst(h)a(h)nd(h)* (line 12).⁷ The laughter particles in line 13 are compatible with such a shift.

The paralinguistic formatting of Bill the figure's voice is interpretable as Bill the narrator 'doing' something. But Gina manifestly has a problem in determining what exactly he is doing with Bill the figure's voice. Her problem becomes clear when she initiates repair

⁷ Alternatively, the breathiness and laughter particles could be interpreted as indexing the reporting situation, signalling the narrator's commentary on the figure's action (see below).

in next turn: *(wait) you said he's:* (line 14). Her initiation is accompanied by a candidate repair *pro:: capital punishment* (line 15), which Bill acknowledges by completing with *and pro life* (line 16). Gina ratifies this completion *and pro life* (line 17), signalling with a follow-up *ri:::ght* (line 18) that she has got the point. Moreover, Bill now makes explicit in next turn what his point is: *that's like - to me that's like a hu::ge contradiction* (lines 19-20). It is thus an understanding of the contradiction between the two political positions referred to which is treated as being necessary in order for Gina to make sense of Bill the figure's turn. Once it is clear that there is a contradiction and what it is, the paralinguistic overlay of breathiness and laughter fits in as a contextualization cue to how Bill the figure evaluates the contradiction (as well as the person who epouses it) and to how Bill the narrator wishes his addressee to evaluate it.

To summarize the argument so far: by examining reported speech sequences in which there is 'trouble', as evidenced by participants finding it necessary to initiate repair, we have shown that the pursuit of coherence in conversational reported speech involves finding plausible answers to at least three questions: *Is this current speaker's 'voice' or someone else's?; If so, whose 'other voice' is this? and How is this 'other voice' being done? (or What is the speaker doing with this 'other voice'?)*. 'Trouble' may occur when there are no clear answers or only conflicting answers to these questions. Moreover, finding appropriate answers sometimes depends crucially on prosodic and paralinguistic framing devices. Where prosodic and paralinguistic signals are inappropriate or ambiguous (and verbal content does not fully disambiguate), participants may find that repair is required in order to establish coherence in reported speech sequences.

The question which now arises is what counts as appropriate prosodic and paralinguistic framing for conversational reported speech. To address this question we shall examine a selection of 'successful' reported speech sequences and compare them to the repairable ones above.

Prosodic and paralinguistic framing in 'successful' reported speech sequences

'Successful' reported speech sequences are recognized not only negatively by the absence of repair. There are often other clues to the fact that participants are making sense of ongoing talk. For instance, participants may show through the recipient design of their uptake that they are orienting to a bit of talk as the reported speech of a figure rather than as current speaker's own words. One special case of this

involves the phenomenon of ‘chiming in’, when recipients participate in the voicing of a particular figure. On other occasions conversationalists may produce metalinguistic talk about the ‘other voice’ and/or the way it is being done. These phenomena can be thought of as positive signs of coherence in conversational reported speech sequences. They will all be documented in one way or another in the examples we shall now examine of successful ways of handling the three questions of reported speech.

Vocal framing of an ‘other voice’

If ‘trouble’ in reported speech sequences involves problematic answers to the question *Is this current speaker’s ‘voice’ or someone else’s?*, then ‘successful’ sequences should display the use of clear and unambiguous signals that some figure is being animated. Where prosody and paralinguistics are crucially involved, this means using features which are not likely to be taken as indexing current speaker’s own voice. Use of a marked voice quality, intonation or rhythmic pattern which departs from the local, momentary ‘norm’ of talk will thus often suffice to cue some ‘other voice’. Such a departure from local prosodic norms is demonstrated, for instance, in the following episode:

(3) Rainbows (12/739)

Two high school friends Janet and Ann, reunited after several years, are recalling their college experiences. Janet’s husband Steve is present.

1 J: we had a professor from Carleton;
who was an atheist.
A: uh huh ((*off-stage*))
J: and there was this one guy in class.
5 and I can’t remember what his name was;
but he insisted upon::
the Bible being thee truth;
and thee word of Go- [d.
A: [‘of course’
10 J: and argued everything.
A: of course.
J: f₁ </There /were /no /rain=/bo:ws.> <*rhythmic*>
f₁ before the flood.
f₁ [because
-> 15 A: - [<<*h*> WHAT?>
S: ha ha ha
-> J: f₁ </GOD /MADE THE /RAINBOW!> <*rhythmic*>
J: - and he’s like
f₂ <I’m sorry:::> <*stylized*>
20 - you know
f₂ <if you really look at s:ci:entific evide:nce;>
<*stepping down*>

f2 You will see that
 f2 <whenever these things happen:;> <stepping down>
 f2 a rainbow occu:rs.
 25 f2 and as long as those things happened
 f2 <<l> before the flood
 f2 there were rainbows.>
 - and this (.)
 kid just could not (.)
 30 accept [it.
 A: [see you have to say kid too now
 [hhh ()
 J: [yeah he was.
 he was really a kid.
 35 (.)
 and so- yeah
 it would be like
 Every lecture
 he would bring up another: (.)
 40 Bible story; (.)
 and (.)
 defend it because (.)
 -> f1 <<l> /that's what it /said in the /Bible <rhythmic>
 (.)
 45 - it was very frustrating.
 A: wo::w.
 [(you were)
 J: [Amusing; (.)
 but frustrating.

Janet brings two figures on stage here, the 'professor from Carleton' (line 1) and a 'guy in class' (line 4). In line 12, with no further quotative introduction, she animates the figure of the student. Moreover, her interlocutor Ann orients to the presence of this 'other voice' with her high *WHAT?* in line 15.⁸ Although the token *what* functions as a next-turn repair initiator with respect to prior talk in some contexts, Janet does not treat it this way here. She does not respond by recycling her turn but carries on instead with *(because) God made the rainbow*. In fact, Ann's *what* token is cued with loud volume and high pitch - a display of astonishment (Selting 1996), in this case strongly projective of disagreement. Such strong disagreement would hardly be appropriate, were Janet's talk to be understood as being 'in her own voice'. But it is fully appropriate if *there were no rainbows before the flood* is animated as belonging to someone else.

How is it that Ann recognizes lines 12ff as being 'in another voice'? In this case there is no significant change in the voice quality, pitch or loudness of Janet's speech. What does make these lines stick out from surrounding talk, however, is their pronounced rhythmic quality. The accented syllables are timed so as to come at approximately

⁸ Steve's laughter in line 16 can also be thought of as a sign of orientation to Janet's 'voicing' of the student.

equal intervals in time, i.e. they are perceptually isochronous (Couper-Kuhlen 1993). Moreover, in line 12, every word is 'scanded' with a separate accent (the compound *rainbows* being rendered prosodically as two independent words *rain* and *bows*).⁹ It is thus the rhythmic declamation of these lines which contributes to marking them as different from prior talk and serves as an iconic cue to the shift in footing.

Yet the rhythm of lines 12ff is not only a cue to the presence of different 'voice'; in its declamatory style it is also indexical of some specific 'other voice', a 'voice' which could be described in this context as school-boy-like. (See Janet's characterization of him as *a kid* in lines 29 and 34.) In fact, the student's voice is consistently given this rhythmic configuration in Janet's story. Not only is *GOD MADE THE RAINBOW* (line 17) also rhythmically marked, but later in the story, when the student reappears, his voice is once again rhythmically declamatory: *that's what it said in the Bible* (line 43).¹⁰ The choice of prosodic device for cueing the voice of a figure is thus a motivated one. In conversation analytic terms, it is part of the recipient design of a speaker's talk, talk which is geared to accomplish specific goals in specific situations. In reported speech sequences these goals often include presenting figures with a particular 'take', a point we shall return to shortly.

Acknowledging the fact that prosodic voicing in reported speech sequences is recipient-designed allows us (as analysts) to appreciate the significance of 'chiming in'. Conversationalists on occasion display orientation to an 'other voice' by participating in the reporter's animation:

(4) Galileo (12/370)

Same speaker constellation as in (3). Talk here is about the Catholic church and the fact that its doctrines change very slowly

1	J:	They just decided what. like last year, that uhm (.) tsk who was it now.
5	S:	<<I> I'm forgetting the name of the guy.> Oh yeah Copernicus?
	J:	C- no not Copernicus;

⁹ The phenomenon of rhythmic scansion in everyday discourse is discussed at length in Auer, Couper-Kuhlen & Mueller (forthcoming).

¹⁰ In fact, it is partly due to the distinctive rhythmic cueing that we recognize the voice in line 43 as the student's and not, for instance, the professor's.

10 S: [()
 J: [GALILEO.
 A: Galileo.
 A: .haah
 J: f₁ that [<<f> /↑GALI/LEO:: <scanded,
 15 A: - [yes
 J: f₁ [/WAS: stepping down>
 A: - [yes
 J: f₁ (.) /RIGHT!>
 - THEY JUST DE- [CIED THAT.
 20 A: [uhh heh heh
 J: f(h)ive hu(h)ndred years later!
 A: That's right.
 that's right.
 that was uh-
 25 J: f₁ <that /↑ma:ybe he /ha:d a /good i/dea.>
 <rhythmic, stepping down >
 S: - that's the speed at which they uhm (.)
 A: hnn.
 S: yeah.
 J: f₁ <they /↑probably /shouldn't have /tortured him
 30 f₁ as /much.> <rhythmic, stepping down >
 All: - HA HA HA
 -> A: next year they'll decide;
 -> f₁ <↑probably the ro- > <stepping down >
 -> f₁ the world ↑is round.
 35 - heh heh
 J: - Right.
 right.

Janet's anecdote is occasioned by Steve's prior remark *even the Vatican changes its policy - over the millennia*. As part of the anecdote she animates the voice of the Catholic church in a recent proclamation concerning Galileo.¹¹ This voice is given a distinctive prosodic configuration: the statement *Galileo was right* (lines 14, 16 & 18), for instance, is configured as one intonation phrase with high pitch on its first accented syllable and all subsequent accents descending in a stylized fashion throughout the phrase. Every possible lexical stress (including secondary stresses and monosyllabic function words) is given an accent and these accents are timed so as to occur regularly in time. The same distinctive isochronous pattern, with exactly the same high onset and stylized accents which step down, recurs on *maybe he had a good idea* (line 25) and *they probably shouldn't have tortured him as much* (lines 29-30).

¹¹ Interestingly, the grammatical framing is not one canonically associated with 'expressivity' or vocal animation: *they just decided* (line 1) <...> *that* (line 3) <...> *that Galileo was right* (lines 14ff), *that maybe he had a good idea* (line 25). See Guenther 1997 for further discussion of this point.

But this pattern is not only distinctive for analysts. Janet's interlocutor Ann joins in with the animation of the Vatican's voice in lines 32ff and thus displays her orientation to it. And she attempts the same prosodic configuration: her accent on *probably* (line 33) hits exactly the pitch level that Janet's prior animated onsets had and the next accent on *ro-* is targeted to step down from there. The pitch replication now breaks down with what is manifestly a slip of the tongue: *ro(und)* instead of *world* (cf. line 33) but the rhythmic replication persists with the accumulation of accents in *the world is round* (line 34). With her prosodic 'chiming in', Ann thus implicitly signals her understanding of the presence of some 'other voice' and indeed of some particular 'other voice'. Moreover, she can also be heard to co-align with the 'take' which current speaker has on this voice (see below).

Vocal framing of several 'other voices'

When more than one figure is 'on stage' in a reported speech sequence, the task which conversationalists face is all the more complex. If the figures' voices are prosodically and paralinguistically animated, they must be done in such a way that they are not only recognizably distinct from the reporter's own voice but also distinct from one another. Fragment (3) (repeated here) demonstrates one way conversationalists have of accomplishing this:

(3) Rainbows (12/739)

```

1   J:      we had a professor from Carleton;
      who was an atheist.
      A:      uh huh ((off-stage))
      J:      and there was this one guy in class.
5   J:      and I can't remember what his name was;
      but he insisted upon::
      the Bible being thee truth;
      and thee word of Go-   [d.
      A:                                 ['of course']
10  J:      and argued everything.
      A:      of course.
      J:      f1  </There /were /no /rain=/bo:ws.> <rhythmic>
      f1  before the flood.
      f1  [because
15  A:      -   [<<h> WHAT?>
      S:      ha ha ha
      J:      f1  </GOD /MADE THE /RAINBOW!>           <rhythmic>
      J:      -   and he's like
->  J:      f2  <I'm sorry:::>                               <stylized>
20  J:      -   you know
      f2  <if you really look at s:ci:entific evide:nce;>
      f2  <stepping down>
      f2  You will see that
      f2  <whenever these things happen::>

```

<stepping down>

25 f₂ a rainbow occu:rs.
 f₂ and as long as those things happened
 f₂ <<l> before the flood
 f₂ there were rainbows.>
 - and this (.)
 kid just could not (.)
 30 accept [it.
 A: [see you have to say kid too now
 [hhh ()
 J: [yeah he was.
 he was really a kid.
 35 (.)
 and so- yeah
 it would be like
 Every lecture
 he would bring up another: (.)
 40 Bible story; (.)
 and (.)
 defend it because (.)
 -> f₁ <<l> /that's what it /said in the /Bible.>
 <rhythmic>
 (.)
 45 it was very frustrating.
 A: wo::w.
 [(you were)
 J: [Amusing; (.)
 but frustrating.

Following her animation of the student in lines 12, 14 and 16, Janet begins to animate the professor in line 19 with *I'm sorry*. Significantly, the fact that this line is the professor's voice and not the student's is inferrable only from its distinctly different prosodic delivery. (The anaphoric expression in Janet's quotative *and he's like* (line 18) could refer to either the student or the professor.) Instead of the rhythmic and loud configuration which characterizes the student's *there were no rainbows* (line 12) and *God made the rainbow* (line 17), line 19 —*I'm sorry* — has lower volume, stylized pitch and syllable stretching. It is because of this markedly different prosody that we infer that it is not the student who is 'apologizing' but some other figure.

Just as with the prosodic framing of a single figure, the voicing of two or more figures also displays individualized recipient-design. That is, different figures receive not only distinct voicing but also specific different voicing. Consider now another case in which this is successfully accomplished:

(5) Rollerblades (12a/40)

Bret and Wanda are brother and sister. Ricky is Wanda's young son. The topic here is rollerblading.

1 B: well I mean
Here you see
you know five-year-old kids doing -
W: well Ricky
5 some of Ricky's friends
and they do wonderful -
but I mean
°gosh°
They're dangerous too
10 and kids don't real-
His one friend
uhm
gosh he-
I said
-> 15 f1 </wha:t /↑h:appened. < slow>
f1 did a /tr:u:ck hit him?
?f1 <the kid ('s got) his f:a:ce, <listing>
?f1 his who:le l:e:g,>>
-> f2 <<h+p> ↑well I was rollerblading;>
<lax, breathy>
20 - <and I'm going> <whispered>
f1 <wooh!> <whispered>
A: - <yeah <whispered>
I mean>
B: [WELL ACTUALLY YOU SHOULD HAVE-
25 A: [you can't stop!
B: YOU SHOULD HAVE ELBOW PADS
KNEE PADS [AND A
W: [<<h> oh yeah!>
B: AND A HELMET ON.
30 A: you can't stop!

Wanda's anecdote involves three figures, two of which are animated: herself and *the kid*, a friend of her son Ricky. The first animation begins in line 15, where Wanda the figure asks *what happened*, the type of question which — as an opener and accompanied by the appropriate prosody — tends to be occasioned by some remarkable feature in the addressee's external appearance. As the anaphora in line 16 makes clear, however, this question is not addressed to the kid himself but to a third party, who is thereby presented as knowledgeable about the kid and his affairs; this is presumably Ricky, the kid's friend. The next line *did a truck hit him* (line 16) proffers a candidate answer to Wanda the figure's own question, couched itself as a question presumably to the same third party. Both of these lines are animated with distinct prosody: the accented syllables are given a slow, rhythmic delivery, with marked lengthening of syllables and initial consonants. The pitch movements have a relatively wide range. This distinct delivery pattern is continued in the next two lines: *the kid('s got) his face, his whole leg*, with noticeable lengthenings and slow tempo. These lines (17-18) are thus interpretable on one reading as Wanda the figure making

explicit to her interlocutor what the visual signs were which occasioned the concern she expressed in lines 15-16.¹²

In line 19 a new figure enters.¹³ Despite the lack of any reportative introduction, the presence of a different figure is clearly marked by a shift in prosody: the line *well I was rollerblading* is configured with overall high pitch and low volume. The voice quality is breathy and the articulation lax. It is the prosodic contrast between this line and prior talk which cues the new figure.¹⁴

Yet not only the contrast between this line and prior talk is noteworthy, but also the way the contrast is constructed. The contrastive figure being enacted is a particular *kid*, with particular characteristics, ones which the recipient-designed prosody of his voice cues: rather than the slow, weighty delivery of lines 15-18 — cued as belonging to Wanda the figure — line 19 comes across as light and airy. The kid is thus not only distinguished from Wanda, his voice is also designed to suggest an opposing stance. Both sets of prosodic features can be heard as cueing stances which are hinted at elsewhere in talk: Wanda the figure's 'weighty' prosody becomes an index of *they're dangerous too* (line 9), while the kid's 'light and airy' prosody indexes *kids don't real(ize)* (line 10).

Notice now that the next two lines enact Wanda the figure's reaction to the kid's response and its stance. Her *wooh* (line 21) is delivered in a whispered voice, which 'leaks' into the prior reporting construction *and I'm going* (line 20). The whispered quality of Wanda the figure's turn here is significant in two ways. First, it forces us to refine the statement made earlier that figures' voices are animated consistently in conversation. If this were true here, Wanda the figure should say *w:ooh:* and use normal volume just as she does in lines 15-18. Instead, the whispery prosody of this line is designed to cue Wanda the figure's reaction to the kid's response, and this stance is not — for want of a better label — 'serious concern' but something closer to

¹² Alternatively lines 17-18 could be attributed to Wanda the narrator, with their marked prosody contextualizing her 'take' on the events she is recounting (see below). In this case the switch to Conversational Historical Present (Wolfson 1979, 1982, Schiffrin 1981) would begin here rather than in line 20.

¹³ Although we might expect Ricky to enter the scene at this point, the wording of the line makes clear that Ricky's friend, *the kid*, is answering Wanda's question.

¹⁴ The fact that prior talk contained a first pair-part (*what happened?*) which makes a second from some other (addressed) party conditionally relevant may also contribute to the fact that we hear line 19 as belonging to a different figure.

'speechless (or voiceless) amazement' (cf. her *gosh* in lines 8 and 13, the first time also done with whispery prosody). Thus, to have used the slow, rhythmic prosody of lines 15-18 would have cued the wrong message. Yet although Wanda's voice is not consistently done, her 'voice' is, in the sense that she is constructed as a figure who is concerned by young rollerbladers' accidents and amazed by their carefree attitude.¹⁵

Vocal characterization of 'other voices'

Figures, we have stressed, are voiced as specific 'other voices': not only have their 'lines' been constructed for them to 'say', their prosody hints at stances which they are constructed as assuming — and/or which their constructor, the current speaker, wishes to take towards them. The speaker, in other words, is not only 'doing' voices but also doing something with those voices which it is incumbent upon recipients to infer. Every detail of the way figures in reported speech sequences are constructed is thus inspectable for some clue as to what the speaker is doing with these 'voices'. In particular, the way a given 'voice' is configured prosodically and paralinguistically will serve as a hint to the 'take' the speaker has on that figure, suggesting how it, or the talk of which it is part, is to be received and evaluated by interlocutors.

How do interlocutors make sense of the prosodic and paralinguistic cues of reported speech? Or better: how do we as analysts know what sense they have made of them? What signs of participant 'coherence-making' are visible in the interaction itself? One type of sign is verbal in nature: recipients may engage in meta-talk which displays how they have interpreted the prosodic cueing of a particular figure:

(6) Breathing in smoke (12/106)

Same speaker constellation as in (3). The topic of conversation here is the ban on smoking in public buildings which has been instated since Ann left the country to study abroad.

1 A: it's funny (.)
 cause I haven't- (.)
 I mean I- (.)
 I remember being kinda for it?
5 when- (.)
 when I left

¹⁵ In other words, as Tom Luckmann and Susanne Guenther have pointed out, these characteristics do not contradict each other but are quite compatible.

it was just kinda like (.)
 f₁ ↑Actually you know - (.)
 f₁ ↑I don't like to smoke all this- er
 10 f₁ <<l> /breathe in all this /smoke all the /time,>
 Now I come back
 and I'm just kinda like (.)
 f₂ <<p> /they should be able to /sm(h)oke,>
 f₂ I mean /what are you guys DOING! heh
 15 S: yeah=
 J: =right
 -> it's a little overboard.
 -> A: yeah

Ann animates two figures here, both of which are herself at different moments in time. The 'early' Ann figure is made to say lines which are hearable as supportive of the anti-smoking ban *actually you know I don't like to (..) breathe in all this smoke all the time* (lines 8-10), whereas the 'later' Ann figure claims that those who are prevented from smoking should be able to do so (line 13), thereby implying that she no longer supports the ban. Both 'voices' are set off prosodically from Ann's framing talk (*it was just kinda like*, line 7; and *I'm just kinda like*, line 12) by pauses, pitch shifts and volume shifts. But the way the 'later' Ann figure is voiced hints at something more: *smoke* in line 13 has the beginning of an overlaid 'snort' and the line *what are you guys doing*, delivered with raised larynx, is followed by a laughter particle. These delivery features are signs that Ann the narrator is not just contrasting an earlier state of mind with a later state of mind on smoking bans but that she is also evaluating these positions, aligning herself with the latter as opposed to the former.¹⁶ Ann's interlocutors respond in ways which are hearable as aligning with this critical stance. In fact Janet's next turn actually formulates verbally what Ann was alluding to with her 'snort' and laugh particle: *it's a little overboard* (line 17).¹⁷ Moreover, Ann ratifies this understanding with her follow-up *yeah* (line 18). Thus, Janet's turn amounts to

¹⁶ The overlaid 'snort' and the laugh particle could be thought of as cueing the talk of either the 'later' Ann figure or the narrator Ann. Since Ann the narrator's stance is congruent with that of Ann the 'later' figure, the net outcome is the same as far as the inferencing here is concerned. See, however, fragment (7) below and its analysis.

¹⁷ In this sense it is the same practice as that documented in Schegloff's (1996) collection, namely formulating a candidate observation, interpretation, or understanding of something which a prior speaker has conveyed without saying: "...some telling may be constructed by its teller, and/or be taken by its recipient, to embody and/or to reveal a tack that the teller is taking to the tale, some stance being taken up, or some action being done. When a recipient makes that explicit in the uptake, the teller can confirm both the particulars of the uptake, its 'propositional content' so to speak, and that he or she was engaged in such a 'project'" (1996:188).

'putting into so many words' the effect of the work which prior speaker's vocal animation was designed to do. It is metapragmatic discourse (Silverstein 1993) *par excellence*.

The Galileo fragment (repeated here) provides another example of metapragmatic reference to the work which prosody is doing in a reported speech sequence:

(4) Galileo (12/370)

1 J: They just decided what.
like last year,
that uhm (.) tsk
who was it now.
5 <<l> I'm forgetting the name of the guy.>
S: Oh yeah
Copernicus?
J: C-
no not Copernicus;
10 [()
S: [GALILEO.
J: Galileo.
A: .haah
J: f₁ that [<<f> /[↑]GALI/LEO:: <scanded,
15 A: - [yes
J: f₁ [/WAS: stepping down>
A: - [yes
J: f₁ (.) /RIGHT!>
- THEY JUST DE- [CIDED THAT.
20 A: [uhh heh heh
J: f(h)ive hu(h)ndred years later!
A: That's right.
that's right.
that was uh-
25 J: f₁ <that /[↑]ma:ybe he /ha:d a /good i/dea.>
<rhythmic, stepping down>
-> S: - that's the speed at which they uhm (.)
A: hnn.
S: yeah.
J: f₁ <they /[↑]probably /shouldn't have /tortured him
30 f₁ as /much.> <rhythmic, stepping down>
All: - HA HA HA
A: next year they'll decide;
f₁ <[↑]probably the ro- > <stepping down >
f₁ the world [↑]is round.
35 - heh heh
J: - Right.
right.

In line 26 Steve responds to Janet's enactment of the Vatican proclamation with *that's the speed at which they uhm*. On one level this utterance is hearable as an explicit reference to — or formulation of — the point which Janet is making with *they just decided what, like last year* (line 1f), *they just decided that* (line 19) and *five hundred*

years later (line 21). On another level, however, Steve's remark can be heard as putting into so many words what Janet's (slow and laborious) prosody is cueing in her animation of the Vatican's voice. On this reading, it 'verbalizes' the coherence which Steve is attributing to the way the reported speech is done.

On other occasions, rather than verbalizing the effect of the vocal framing of a figure or figures, recipients will instead show their understanding of what the speaker is doing by making responses tailored specifically to the reported speech sequence. The Rollerblades fragment (repeated here) provides an example of this:

(5) Rollerblades (12a/40)

```

1   B:      well I mean
      Here you see
      you know five-year-old kids doing -
5   W:      well Ricky
      some of Ricky's friends
      and they do wonderful -
      but I mean
      °gosh°
      They're dangerous too
10  and kids don't real-
      His one friend
      uhm
      gosh he-
      I said
15  f1 </wha:t /↑h:appened.           < slow>
      f1 did a /tr:u:ck hit him?
      ?f1 <the kid (`s got) his f:a:ce,       <listing>
      ?f1 his who:le l:e:g,>>
-> f2 <<h+p> ↑well I was rollerblading;>
      <lax, breathy>
20  - <and I'm going>                 <whispered>
      f1 <wooh!>                       <whispered>
-> A: - <yeah                          <whispered>
-> I mean>
-> B: [WELL ACTUALLY YOU SHOULD HAVE-
-> 25 A: [you can't stop!
-> B: YOU SHOULD HAVE ELBOW PADS
-> KNEE PADS [AND A
      W: [<<h> oh yeah!>
-> B: AND A HELMET ON.
-> 30 A: you can't stop!

```

Recall that Wanda the figure is animated with 'weighty' prosody, indexing a stance which Wanda the narrator has articulated verbally with *but I mean gosh they're dangerous too* (lines 7-9), while *the kid* is animated with a breathy, light voice, indexical of the stance implicit in Wanda the narrator's *kids don't real-(ize)* in line 10. In a sense then Wanda the narrator has 'primed' the vocal characterizations of her figures and in doing so, set up two models for co-alignment, one roughly associated

with the danger motif, the other with not realizing the danger. Therefore, it is not wholly coincidental that her interlocutors orient to these models in subsequent talk. Ann picks up the danger motif by specifying why rollerblades are dangerous - *you can't stop* (line 2), while Bret addresses the kid's non-realization by detailing what protective equipment rollerbladers should wear: *actually you should have elbow pads, knee pads and a helmet on* (lines 24, 26f and 29).¹⁸ Thus here participants show through responses specifically designed with respect to the perceived stances what understanding they have of the way the figures' voices are being done. In this case, the inferencing job is made easier by the verbal hints which current speaker has provided in prior talk.

Yet what if a current speaker's 'priming' talk is ambiguous? A final fragment demonstrates that this can and does happen in conversational reported speech sequences:

(7) Rented cars with phones (12a/200)

Same speaker constellation as (5). Reference has just been made to cellular phones.

1 B: Oh I got a kick outa- (.)
well you had a rented car;
and Ricky was so worried;
cause you were gone somewhere.
5 (it's just)
f1 <<l+p> what if she has a flat ↑tire.>
f1 <<cresc> she /hasn't /got a /↑phone;
f1 in her /↑car!> <rhythmic>
-> W: - heh heh heh heh (.) heh
10 B: f1 <CAN YOU IMAGINE /RENTING A /CAR
f1 THAT /DOESN'T HAVE A /PHONE?>
<rhythmic>
-> All: - ha ha
W: I didn't know that they rented cars with phones.
15 although I'm sure they do;
-> isn't that fun(ny)?
B: he was so worried cuz-
W: <<h> when was this;>
B: when you were up here,
20 sometime,
W: <<p> hn>
last time,
-> W: <<p> isn't that funny?>
OHHH
25 at- at Christmastime.
B: yeah,

¹⁸ Note too that Ann 'chimes in' with Janet's vocal framing when she whispers her uptake *yeah I mean* (line 22f).

W: right after Christmas.
 yeah
 B: he was so:: worried;
 30 f₁ that his /mom was out in the /car,
 f₁ and she had /rented one;
 f₁ <that didn't have a /phone;> <rhythmic,
 stylized melody>
 -> C: - ha ha ha
 B: f₁ <I hope she's all right; <rhythmic,
 35 f₁ she can't /call us; same stylized
 f₁ she /hasn't got a /phone;> melody>
 -> W: - heh heh
 (he) takes after his father.
 I don't worry about stuff like that.
 40 they do though.

Bret's priming for the figure of Ricky, whom he animates in line 6ff, becomes apparent for the first time in line 3: *Ricky was so worried*. This verbal characterization is repeated in line 17: *he was so worried cuz-* and once again in line 29: *he was so:: worried*. Yet the way Bret 'does' Ricky's voice is not indexical of worry throughout the animation. Line 6, *what if she has a flat tire*, is configured with somewhat softer volume and lower pitch than prior talk, a style of delivery which is indeed suggestive of worry. Yet in lines 7-8, although Bret continues to animate Ricky's voice, the prosody changes gradually to increasingly louder volume and marked rhythm; a high point is reached in lines 10-12, where the volume is very loud and the rhythm heavily marked. This prosodic configuration suggests something more than worry, and indeed recipients do not respond with, say, co-aligning expressions of concern and sympathy but rather with giggling (line 9) and outright laughter (line 12). Wanda even verbalizes her understanding of Bret's voicing twice *with isn't that funny?* (lines 16 and 23). Bret now animates Ricky once again, this time employing, in addition to marked rhythm, a stylized melody ending with a call contour (Ladd 1978) on each of the animated lines. This animation too is receipted with laughter from recipients in lines 33 and 37. Thus, Bret has verbally cued Ricky as being worried, while his voicing of Ricky is suggestive of something which provokes giggling and laughter in recipients.¹⁹

This example is telling because it reminds us that recipients do not rely blindly on verbal 'priming' by the speaker in trying to make sense of the prosodic and paralinguistic cueing of an animated figure in a reported speech sequence. Instead, in cases of non-congruence, the verbal hints must be weighed up against the prosodic hints and a more or less warranted 'guess' must be hazarded as to what the speaker

¹⁹ Note that the broken off *I got a kick outa* (line 1) and the expressive lengthening on *so:: worried* (line 29) both serve as cues to Bret's overall framing.

is doing. This is of course risky business, but example (7) shows that it need not lead to a break-down of coherence.

In conclusion, the above discussion has shown that not only when explicit cues to reported speech are lacking but also when explicit cues about reported speech are misleading, participants make sense of conversational reported speech in part by relying on the prosodic and paralinguistic details of 'voicing'. As Schegloff reminds us, "talk is laced through and through with inexplicitness and indexicality" and this inexplicitness is constantly being "solved" by hearers. Moreover, "its results are displayed (even when not formulated) in the ensuing talk and action and are subject to repair there if found problematic" (1996: 219f). This paper has attempted to spell out exactly what might be problematic about indexically (or prosodically) cued reported speech, thereby making it repairable, and how in more 'successful' instances, recipients display their 'solutions' of the inexplicitness involved to each other and to analysts.

References

- Auer, P., & Luzio, A. d. (Eds.). (1992). *The Contextualization of Language*. Amsterdam: Benjamins.
- Auer, P., Couper-Kuhlen, E., & Mueller, F. (forthcoming). *Language in Time: The Rhythm and Tempo of Spoken Interaction*. New York: Oxford University Press.
- Couper-Kuhlen, E. (1993). *English Speech Rhythm: Form and function in everyday verbal interaction*. Amsterdam: Benjamins
- Couper-Kuhlen, E., & Selting, M. (1996). Towards an interactional perspective on prosody and a prosodic perspective on interaction. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation* (pp. 11-56). Cambridge: University Press.
- Goffman, E. (1981). *Forms of Talk*. Oxford: Blackwell.
- Guenther, S. (1997). Stilisierungsverfahren in der Redewiedergabe. Die 'Ueberlagerung von Stimmen' als Mittel der moralischen Verurteilung in Vorwurfsrekonstruktionen. In B. Sandig & M. Selting (Eds.), *Stil und Stilisierung*. Berlin: de Gruyter.
- Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In M. J. Atkinson & J. Heritage (Eds.), *Structures of Social Action. Studies in Conversation Analysis*. Cambridge: University Press.
- Ladd, D. R. (1978). Stylized intonation. *Language*(54), 517 - 39.
- Laver, J. (1980). *The Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.
- Laver, J. (1994). *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Levinson, S. (1988). Putting linguistics on a proper footing: Explorations in Goffman's concepts of participation. In P. Drew & A. J. Wootton (Eds.), *Erving Goffman: Exploring the interaction order* (pp. 161-227). Cambridge: Polity Press.
- Schegloff, E. A. (1990). On the organization of sequences as a source of "coherence" in talk-in-interaction. In B. Dorval (Ed.), *Conversational Organization and its Development* (pp. 51-77). Norwood, NJ: Ablex.
- Schegloff, E. A. (1996). Confirming allusions: Towards an empirical account of actions. *American Journal of Sociology*, 102(1), 161-216.
- Schiffrin, D. (1981). Tense variation in narrative. *Language*, 57, 45-62.
- Selting, M. (1996). Prosody as an activity-type distinctive cue in conversation: The case of so-called 'astonished' questions in repair initiation. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation* (pp. 231-270). Cambridge: Cambridge University Press.
- Silverstein, M. (1993). Metapragmatic discourse and metapragmatic function. In J. A. Lucy (Ed.), *Reflexive Language. Reported Speech and Metapragmatics* (pp. 33-58). Cambridge: University Press.
- Tannen, D. (Ed.). (1993). *Framing in Discourse*. New York & Oxford: Oxford University Press.
- Wolfson, N. (1979). The conversational historical present alternation. *Language*, 55, 168-182.
- Wolfson, N. (1982). *The Conversational Historical Present in American English Narrative*. Dordrecht: Foris.

Transcription conventions

One line	One intonation phrase
First word capitalized	High onset (=declination reset)
[Line	
[Line	Overlapped utterances
Line=	
=Line	Latched utterances
Line.	Final pitch falling to low
Line!	Final pitch falling to low, emphatic
Line;	Final pitch falling slightly
Line -	Final level pitch
Line,	Final pitch rising slightly
Line?	Final pitch rising to high
<<p> Line>	Piano
<<f> Line>	Forte
<<l> Line>	Low register
<<h> Line>	High register
<<cresc> Line>	Crescendo
<<decresc> Line>	Decrescendo
↑Word	Noticeable step-up in pitch
Wo::rd	Lengthened sound or syllable
Word-	Cut-off sound or syllable
WORD	Loud volume
°word°	Soft volume
<u>word</u>	(Extra)stress
/word /word /word	Rhythmic delivery
(h)	Breathiness
.hhh	Inbreath
hhh	Outbreath
(word)	Unsure transcription
(.)	Brief pause
(1.0)	Measured pause