PART TWO: NOTES ON PROSODY

This section details some of the current thinking about prosody and uses a range of data examples to illustrate it. You can go straight to Part Three if you want to practise transcribing and do not need the theories about prosody presented here. Or you may want to come back to this part once you have tried transcribing from the CDROM.

Many people think of transcription as writing down the words of speakers as if it was a play script. This kind of transcription only tells you what words they used and who spoke them. But this is only the beginning of trying to understand what was going on in the interaction you recorded. Some of the other issues have been referred to in Part One. But the most technically challenging aspects of transcription relate to prosodic transcription and analysis.

What is prosody?

Prosody has been called the musical attributes of speech. Indeed, the word ‘prosody’ has its origins in the Greek word for melody. It consists of those elements of talk which can be heard and which are not words or sounds (such as hmms, laughter, false starts and so on). These hearable elements (or, more technically, auditory effects) include: melody (tunes and other intonation patterns), rhythm, dynamics, tempo and pausing.

Prosody is absolutely central to discourse analysis because when we talk we do more than produce words. We convey emotions, we show what is important and what is not, we tie our information together into a line of argument, we create a comfortable interactional climate and so on: ‘Prosody plays a key role in discourse level interpretation, in fact without it there can be no conversing’ (Gumperz 1996:x). We only have to listen to computer generated speech for a few seconds to be aware of how unmusical it is and therefore hard to listen to. But even this kind of speech has some prosodic features such as a fixed rhythm. We (or machines) simply could not talk at all without prosody.

The Daily Mail has an advert which goes as follows:

‘You’re better looking in the Mail’.

The double meaning here depends on prosody. The message the newspaper wants to sell is that ‘you are better off looking in the mail’ - better off in terms of information, financial advice etc. But the meaning which is meant to catch your eye, and make you think about buying the mail, has a more personal message: ‘You will look better if you read the Mail’. As you read the advert, the first meaning depends upon taking a slight pause after ‘better’ (where the ‘off’ is implied) and the rest of the words form a unit with the first syllable of ‘looking’ which forms part of a smooth intonation contour, gliding down to a fall at the end. The second meaning ‘reads’ differently.
Here the two words ‘better looking’ are elided together with an equal stress on the first syllable of each word and the slight pause is after ‘looking’. ‘In the mail’ then forms a unit with falling intonation. It is only the prosody which disambiguates the two meanings. So prosody plays a central role in helping us to make meaning and choose between different meanings.

Prosody is not just an extra layer on top of grammar and lexis. It is not a question of laminating on a ‘tune’ to a word. It is distinct from the lexico-grammatical system but often functions in harmony with it as a ‘quasi independent signalling system’ (Couper-Kuhlen and Selting 1996). And it comes as no surprise to find that different languages and varieties of English use prosody and syntax in different combinations to convey information, attitude and conversational involvement as we shall see below.

Prosody in Discourse Analysis has not been given the attention it deserves. This is mainly because language has been equated with what Abercrombie (1965) calls ‘spoken prose’ - i.e. language organised for reading but read out aloud. But also because, as Bolinger (1972) says, ‘prosody is an untamed horse’. In other words, there is still no agreed system of how to transcribe it. The reason it is difficult to ‘tame’ is that it cuts across words and therefore is hard to segment. Also, prosodic phenomena are a matter of degree (what is called gradient) rather than discrete items. So, for example, a raise in pitch or an increase in loudness are relative to what has gone before and what is to come. More generally, it is very difficult to hear pitch movement in naturally occurring speech and then to hear, simultaneously, several other different aspects of prosody as they work together to create meaning. For example, to make sense of a speaker’s utterance, you may have to listen to the combination of: pausing, pitch at the onset of speech and then any pitch movement, the rhythm and pace of the utterance and degree of loudness.

**Prosody: features and functions**

**Basic notes on prosody**

These notes are based on Couper-Kuhlen 1986, Gumperz 1982, Halliday 1976. The following features are included in prosody:

1. Intonation i.e. pitch levels on syllables and their combination into contours - which gives speech its melody
2. Change in loudness
3. Stress - perceived in terms of pitch, loudness and duration
4. Other variations in vowel length
5. Tempo and pausing
6. Overall shifts in pitch register i.e. changing and maintaining a certain pitch of voice over an utterance or part of an utterance.
7. Rhythm

Intonation and pitch register have been analysed quite extensively and the following brief description gives a little detail of some of the thinking in the 1970s and early 1980s.
Some traditional categories for prosodic analysis

1. The tone group or information unit. This is the basic planning and process unit in speech production and analysis. There are:
   a) major tone groups. These indicate some sense of semantic completion (what might be a sentence in written prose) with a final or concluding fall. These are commonly marked with a double slash //       // or with a full stop at the end of the group.
   b) minor tone groups. These indicate that there is likely to be more to come but this particular part of the utterance has some sense of being complete in the sense that a clause in grammar is complete in itself - i.e. temporary closure. These are marked with a single slash /       /

Example:
1. what i said was/
2. .. that it is not a suitable course/ ..for you to apply for//
3. .. now if you want to apply for it/
4. .. of course you can do what you want//

(Gumperz and Berenz 1991)

2. The tonic or accent. This gives the clue as to what the listener should focus on. The tonic or accent is the major stressed syllable. In the notation system in these materials, the accent is marked with an asterisk * if there is extra prominence.
3. The tune or pitch accent movements. This is the rising or falling tune on the tonic or accented syllable. In the notation system used here, detailed pitch movements are not included because they require considerable ear training. For applied linguists interested in developing their prosodic analytical skills much further, the GAT system (Seltig et al 1998) is recommended.
4. Shift in pitch register. This marks where the speaker’s voice shifts up or down over several words and is marked as (hi) or (lo).
5. Loudness and speed. These are marked with abbreviations: (acc), (p), (f).

The idea of a tone group boundary and of a tonic is fundamental to much intonation research and is an important ingredient in the system of transcription introduced in these materials and used as a model in the CDROM. However, recent thinking (see below) suggests that conversation does not necessarily fit into neat tone group boundaries and that, often, intonation researchers have to fall back on analysis of syntax to decide where the boundaries fall. In the next two sections, we will look in a little more detail at how prosody functions. Instead of focusing on the tone group or the tonic, we will look at how prosody functions in interaction. But, first, we will discuss some of the issues surrounding prosodic analysis more generally.
Why analyse conversation prosodically?

Interaction involves planning and co-operation by speaker and listener. The listener has to actively monitor and respond. Prosodic signals are as important as syntactic and lexical signals in helping to make sense of talk. Prosody plays a vital role in enabling speakers and listeners to chunk the stream of talk into basic message units so that talk can be interpreted.

It is part of the constant process of making inferences, tying these into past experiences and generating expectations about what is to come. It is a key feature in the process of contextualisation (Duranti and Goodwin 1992, Auer and di Luzio 1992).

So prosody, by showing us the point of prominence in an utterance helps us to:

1. scan the stream of talk so that we can group words and phrases
2. distinguish main from qualifying phrases
3. signal relations between utterances and parts of utterances e.g. subordination, finality
4. select between a range of possible meanings
5. convey attitude and affect
6. know when to take a turn
7. pick out the line of argument
8. establish the particular shade of meaning or perspective which the speaker wishes to convey
9. establish and maintain conversational synchrony and involvement e.g. showing how a particular utterance is designed for that context.

In this way, we can fit what we hear into a consistent theme or line of argument and interpret speakers’ intent. Part of this process involves categorising a particular activity, and prosody helps us to decide what the activity is: for example, is it a joke or a warning? And to show when the speaker is shifting from one activity to another.

How prosody relates to syntax in order to signal meaning may differ systematically between languages and/or between one cultural area and another. Since prosody is so important as a basic planning and processing mechanism, systematic differences transferred from the first language to other languages can fundamentally affect how the speaker of a second or other language is interpreted. Judgements of competence, co-operation, effectiveness and so on may all be based on the speaker’s communicative style which is influenced by the prosodic features from the first language (Erickson and Shultz 1982, Gumperz 1982a, 1982b, 1992a and b)
Interviewing in Intercultural Situations

Now read through John Gumperz’s chapter ‘Interviewing in intercultural situations’ (1992c).

**NB NELLY I HAVE TO FIND OUT WHETHER WE CAN PUT THIS ON THE WEB.**

Concentrate, for the moment, on the first four data examples from this chapter. Study the transcription key first so that you can read the transcriptions with close attention to detail. The first four data examples will be used to illustrate the nine functions of prosody listed above and show how prosody functions in interaction. The numbers in brackets refer to the line numbers of the data extracts in the chapter.

Transcription key

- Final fall
- ? Final rise
- , Slight rise
- - Truncation
- . . Short pause
- <2> Timed pauses (e.g. 2 seconds)
= Overlap
== Latching
* Extra prominence
{ [ ] } Non-lexical phenomena which overlay lexical stretch e.g. [hi]
indicates raised pitch
[ ] Non-lexical phenomenon which interrupt lexical stretch
( ) Unintelligible speech

1. Scan the stream of talk. See Ex. 1 turn 6:

(6) .. I’m RD from the training services.
    And you understand that,
    ..the panel you’re here, ..ehm, * at today,
    the purpose of it is to confirm,
    .. *finally that eh, you’ve chosen the right, *course.

The final fall after ‘training services’, indicated as a full stop, indicates that the next utterance will be on a different topic. The slight rises at the end of the next three lines (indicated by commas) help to chunk the next three phases and separate off the contextualising phrase about the panel from the statements about the purpose of the interview.

2. Distinguish main from qualifying phrases. See Ex. 1 turn 3:

(3) ………and I’ll introduce mr C,
    an instructor at the skills centre,

Here Mr C is introduced and, in parenthesis, his role is identified by the slight rises, indicated by the commas after ‘C’ and ‘centre’ which serve to bracket it out as a qualifying phrase.
3. **Signal relations between utterances. See Ex. 3 turn 8:**

(8)  \{[hi] oh yeah. I understand them.\} =

Here the trainees emphasis on ‘stand’ and the final fall after ‘them’ indicates finality and a full and complete answer in which T’s relatively raised pitch caps R’s preceding low one.

4. **Select between a range of possible meanings. See Ex. 1 turns 7-8:**

(7)  \{[lo] yeah\}
(8)  \{[lo] and to give you the opportunity,
    to ask any questions that you want to ask.\}

T’s low key ‘yeah’ and the statement from R ending with a final fall, suggests that the opening stage of the interview in which its purpose is formally conveyed is now over. Both T and R orientate to this message rather than any other i.e. that this opening stage is still continuing.

The latching here seems to indicate shared understanding of what is going on in this stage of the interview.

5. **Convey attitude and affect. See Ex. 2 turns 2-3:**

(2)  \{[lo] no..no.no.\}
(3)  \{[lo] no. I think you just crossed the wrong ones there.\}

T here puts prominence on the first ‘no’ presumably to be emphatic about it but with a low tone matching R’s earlier tone In previous sequences he has used a low tone to confirm ‘old’ information and by continuing to use this tone he turns this exchange into something routine rather than potentially damaging. If he had combined prominence with a high tone this could have sounded quite challenging as if he was contradicting R.

6. **Know when to take a turn. See Ex. 2 - 3**

The same data as in number 5 also illustrates the crucial role of prosody in turn-taking. T gives three ‘nos’ each with a final fall. He might have expected R to come in after the first ‘no’ but after waiting about half a second, he continues. After the third ‘no’ R takes the turn, echoing his word and the low pitch. Here the overall rhythm of the interaction as well as the length of the pausing is important. The half second pause does not seem to be read by other parties as a ‘long’ pause and although R could have come in, the pause is not ‘read’ by either side as an uncomfortable silence.

7. **Pick out the line of argument. See Example 4 turn 4:**

(4)  ........

    \{[lo] I worked in a yard for some time,\}
    \{[lo] I worked for a builder as well.\}
    And, you know, I did some brick laying, over there.
.. [in-breath] {[ac] I suppose that’s what got me interested, you know.}

The line of argument conveyed here is a subtle one. On the one hand he shows that he has had a considerable period working abroad - not just twelve months but two years. The work in the boatyard is in parenthesis, indicating that it is less important than the brick laying which receives some emphasis from the pause as he clears his throat. On the other hand he does not overstate his experience and his evaluation of it is speeded up so that he does not come across as overselling himself.

8. Establish the particular perspective and shade of meaning. See example 1 turns 1-5 and also the previous example:

(1) come in. <2> hello mr T
(2)  = = hello.
(3)  = = take a seat [sigh] <1> and I’ll introduce mr C, an instructor at the skills centre,
(4)  = = mhm.
(5)  = = how do you do.

These opening lines are in fast tempo and with frequent latchings, indicating an air of informality.

9. Establish and maintain conversational synchrony and involvement. See example 1 turns 1-5 and also Ex. 3 turns 1-6.

In both these examples, the tempo of individual’s speakers and the tempo of turn taking indicates a high level of synchrony and speaker involvement. There are no noticeable examples of what Erickson calls ‘arhythmia’ where the speaking rhythm is uneven and uncomfortable.

Now look again at the transcription notation given above. Spend some time familiarising yourself with the notation scheme used here as it is very similar to the one that is used in the CD-ROM.

Recent critique of traditional analysis of prosody

The examples that have just been given demonstrate how prosody functions in interaction. Bearing this approach in mind, it should now be easier to see why the more traditional ways of analysing prosody have their limitations.

Linguists such as Halliday, Chafe, Bolinger and Ladd have done ground breaking work on the empirical analysis of intonation and stress at sentence level, and how they are used to convey information, and pitch and tempo as they convey feeling and attitude. The traditional approaches to prosody research provide some insights into how prosody functions as a means of conveying (or misconveying information or attitude). However, there are limitations to this type of general analysis. It tends to work with constructed examples at sentence level rather than with chunks of naturally occurring discourse.
It does not pay sufficient attention to how the local context of talk comes to be produced and it is over reliant on grammar as a way of understanding prosodic meaning. Over the past twenty years, there have been several different approaches to the empirical analysis of real talk. Three of these are considered here.

**Birmingham School**

One approach was developed by discourse analysts in the Birmingham School, notably Brazil, and was primarily concerned with assigning meaning to pitch contours in naturally occurring discourse. As Seltig and Couper-Kuhlen point out (1996), this approach was very reliant on a structuralist approach in which meaning was attributed on the basis of contrastive pairs - so, for example contrasting a marked word with an unmarked word. However, in real talk, no such simple contrasts are often possible. Instead, prosody functions to give clues to the listener as to how to interpret what the speaker says by choosing between a closed set of options e.g. the listener infers from the prosodic cues what is being highlighted, what has already been said that this highlighted item relates to and how it may be related. So it is the preceding talk in its particular context which determines how to make on-line interpretations, as we have seen in the examples from the bricklaying interview.

**Conversation Analysis**

This focus on preceding talk is central to the second approach - that of Conversation Analysis. It is the best known and most widely used notation system at the moment. The prosodic features analysed by Jefferson whose transcription system is foundational for CA include pausing, loudness, pitch movements, sound stretching and hesitation phenomena. A brief general introduction to CA is included in the *Introduction* to these materials and the transcription notations used by CA are described in *Part 3 (Step 3)* in more detail (See *Suggested Reading* for several good introductions to Conversation Analysis).

It is ironical that sociologists within the ethnomethodological tradition, rather than linguists, have taken the lead in designing and developing a system which takes account of paralinguistic and some prosodic features as they are managed and interpreted in on-going talk. And indeed many discourse analysts pay less attention to the fine-grained detail of talk than their CA colleagues. Although there are limitations to the CA system, it has been enormously influential among discourse analysts and it would be difficult now for any applied linguist to ignore the massive contribution CA has made to the transcribing and analysing of talk.

But CA is not a completely adequate system. As Seltig and Couper - Kuhlen point out, CA’s main concern is with sequencing rather than prosody and prosody tends to be taken for granted. Although a number of prosodic features are transcribed in CA, they do not figure very strongly in the analysis of talk. In addition, turn-taking and pausing are more systematically transcribed than the features of rhythm, pitch and so on which are so central to the inferential processes which allow conversationalists to make sense to each other.

There are specific criticisms of the CA notation system which Couper - Kuhlen and Seltig have highlighted.
They have outlined two major limitations:

(i) The use of punctuation marks such as the full stop for a stopping fall in tone, the comma for continuing intonation and the question mark for rising intonation is ambiguous. It assumes that sentences and intonation go together. However, a falling tone is not necessarily the end of a sentence and a question does not necessarily have a rising tone. Because the meaning of punctuation marks is so familiar, it is difficult for the transcriber (and even harder for the reader) to resist assuming there is a one to one connection between rising intonation and questions, for example. Similarly, with underlining of syllables to mark stress and with punctuation marks at the end of an utterance (or part of an utterance) it is difficult to tell exactly where the pitch change is. You will have noticed that Gumperz uses some of the features of CA notation, despite these criticisms. Their virtue is that they are easy to read and providing the transcriber is aware of their limitations, they remain useful.

(ii) The conventions are not always detailed enough to be relevant to all aspects of conversational maintenance. For example, CA gives two categories for rise but only one for fall. However, in conversation there are at least two kinds of falling tone. One suggests ‘paragraph’ non-finality and one ‘paragraph’ finality. The difference between these two may affect the listener’s decision about when to take a turn. Similarly, there are different types of pitch at the beginning of a speaker’s utterance which show whether it is a new topic or the beginning of an old one and there is no provision for this in current CA transcriptions. On the other hand, the timing of pauses may be too exact (see notes on transcription). What is important is the timing and the rhythm set up between participants rather than any etic (outsider) counting of micro seconds.

Although there are limitations to CA transcription conventions and they need some modification and elaboration, CA principles and many of its transcription design features, remain central to recent developments in putting interaction and prosody together. In particular their user-friendliness for non-linguists have made them a popular choice for non-specialist discourse analysts.

The set of conventions developed by interactional sociolinguists combines CA techniques with more attention to prosodic details and this constitutes the third approach outlined here.

This approach is the one illustrated by John Gumperz’s chapter, ‘Interviewing in intercultural situations’ (1992c). Gumperz sees prosody as a central to conversational involvement and the negotiation of intent. The interactional features of CA are enriched by the theories of prosodic analysis, both some of the more traditional approaches outlined above and by more recent interactional perspectives on prosody.
Gumperz, contextualisation and prosodic features.

Prosodic features, as outlined above, help to work out the particular meaning and perspective of the speaker’s words. As such they are a key way in which participants make inferences about the other’s meaning and create the local context within which interpretations are made. Prosodic cues help to construct both the immediate local contexts and wider interpretive frames which help speakers draw inferences. Together with such features as code-switching and non-verbal communication, they function as contextualisation cues to channel the interpretive process (Gumperz 1982a, 1992a, 1992b, 1996). They give messages about what kind of activity is going on e.g. is this a joke or a serious comment?, they signal connections to preceding talk, as we have mentioned above, and trigger expectations about what is likely to come. They cue:

… what is to be expected in the exchange, what should be lexically expressed, what can only be conveyed indirectly, how moves are to be positioned in an exchange, what interpersonal relations are involved and what rights to speaking apply.

(Gumperz 1996: 396-7)

Contextualisation cues act as prompts and guides in the interaction, helping participants to negotiate together what they are actually doing, what the topic is, what kind of social relations they have and how they are getting along (or not). The examples given above from ‘Interviewing in intercultural situations’ illustrate how prosodic features function as contextualisation cues. For example, shifts in tempo and pitch cue participants to interpret certain phases of the interview as routine and others as salient. Final phrase intonation is significant in cueing the next speaker’s turn and low pitch register and acceleration help to indicate to the interviewers that the candidate is not being boastful.

The following is an example taken from a gatekeeping encounter which is rather different from the selection interviews in the examples above and the CDROM. A doctor is conducting an oral exam in which the candidate is another doctor who has applied for membership of the Royal College of General Practitioners (Roberts 2000). See Section One for another example from the same data set. The exam is not going very well as the candidate has not cued into the contextualising work the examiner is doing to tell him, indirectly, that he is not answering the question appropriately. A further complication is that the examiner does not cue into the contextualisation cues of the candidate. The examiner’s question concerns a mother’s anxiety that her young son might have caught AIDS from a needle stick injury. The candidate is asked how he would deal with her concerns.
C. is the candidate and E. the examiner:

1. C:  **SO** it's really to get history

2. and to know MORE about her **con**<cems>

3. <as part of (. . )>

4. E: < ye : s.>

5. C: again ==

6. E: = = so(.) shes immed- (..) <its AIDS ( . )

7. whats on her mi : nd. >

8. C: AIDS < on her mind ok > <2>

9. sorry (.)< you want me to see>

Transcription Code

This is based on the GAT system developed at Konstanz and Potsdam Universities (Seltig et al 1998): Gesprächsanalytisches Trankscriptionsystem. There is no need to learn to use this very detailed system unless you are specialising in prosody. But use the code to ‘read’ the transcription in depth so that you can then read the analysis with a critical eye.

= overlapping talk
== latching
( . ) micro pause
( . ), ( . . ), ( . . . ) brief, mid, longer pauses of about 0.25 second up to 1 second
(2), (3) estimated pause of more than one second
; ; ; ; segmental lengthening according to duration
- truncation
acCENT strong primary accent
acCENT weaker secondary accent

*Pitch at end of turn transition unit*

?  Rising to high level
_  level

*Conspicuous pitch jumps*
↑  to higher pitch
↓  to lower pitch

*Changed register*
<<l>> low register
<<H>> high register

*Pitch accent movements*
\  falling to mid
\_ falling to low
/  rising to mid
/\ rising-falling
/\ rising-falling-rising

*Changes in loudness and speech rate*
<<f>> loud
<<p>> soft
<<len>> slow

*Other conventions*
< > non-lexical phenomena which occur between lexical stretches
<< >> non-lexical phenomena which overlap lexical stretches
( ) unclear words

The focus for the analysis will be on lines 1 and 6. In line 1 ‘SO its really to get history’ the candidate appears to be closing this part of his answer. The formulaic and strongly accented ‘so’ with the emphatic ‘really’ suggest a kind of summing up. However, at line three, C. pauses and does not complete. E., at line 6, follows up a slow and strongly contoured ‘yes’ (line 4) with an attempt to refocus C. on the task of how to advise the mother on the HIV/AIDS question. The latching on to line 5 and the accented ‘AIDS’ are used by the examiner as steering cues. Bearing in mind that this is a peer assessment, E. may rely particularly heavily on contextualisation cues in order to convey his intent rather than bring out in the open, so to speak, that he is not getting a preferred answer.

So, in this short extract, we see a number of examples of typical, prosodic contextualisation cues as defined by Gumperz (1982a). We can also see how these cues function at three levels (Gumperz 1992a: 232-3): (1) the perceptual plane in which speech is chunked into manageable units and assessed for coherence and relevance. For example, in line 6 C. has to process the false start made by the examiner and switch his attention to the thematic focus E. gives to ‘AIDS’. (2) the level of communicative intent in the sense of what is going on right now. Here C. needs to be able to make sense of the latching at line 6, as well as the metamessage given off by the accent on ‘AIDS’. 
Both seem to be cueing some kind of request for repair work on the part of C. Similarly, E. interprets C’s utterance at line 1 as a summing up and closing. (3) at the level at which more general framing work is going on in which the candidate is expected to give an institutionally appropriate response according to the oral examination criteria to produce a coherent and consistent response.

So both sides bring interpretative resources to manage the particular and local work that must be done to process the stream of talk, as it happens, and make a situated assessment of intent. The problem is, as line 9 indicates, they do not seem to share the same interpretive conventions.

**Linguistic Varieties and misunderstandings**

Gumperz and his associates have studied how the influence of first language prosodic systems on a speaker’s second language can lead to misunderstandings and uncomfortable moments as in the oral exam example just given. This is the theme of the CDROM material and the accompanying paper by John Gumperz, as we have already seen. The most detailed studies have been carried out comparing the prosody of standard variety speakers of English with speakers of English influenced by North Indian languages (Bhardwaj 1982, Gumperz 1982a, 1982b, 1996, Mishra 1982, Roberts, Davies and Jupp 1992).

The following are some of the contrasts which can lead to misunderstandings and/or negative judgements:

1. **Different ways of distinguishing between normal information flow and contrastiveness.** A Panjabi speaker of English, for example, is more likely to shift the pitch register over a whole utterance or part of an utterance to show contrast whereas a standard variety of English speaker tends to put the accent on the particular word contrasted:
   
   e.g. Panjabi: Do you want a cup of tea/ or/ {[lo] do you want a cup of coffee ?}
   
   e.g. English  Do you want a cup of *tea/ or do you want a cup of *coffee ?

   The Panjabi speaker shifts down to a lower pitch over the whole utterance to show contrastiveness whereas the English speaker picks out ‘tea’ and ‘coffee’ to stress. To the English speaker, the contrast between the tea or coffee is not clear. To the Panjabi speaker, the English speaker is being unnecessarily emphatic.

2. **Different ways of stressing important points.** In the English of speakers of North Indian languages, it is quite common for stress to be placed on a particular word such as a relative pronoun or auxiliary verb to stress its literal meaning or function where speakers of the standard variety of English would put stress on individual words for contrast or emphasis (see 1. above).

   e.g. Speaker of a North Indian language: In the third school/ in which I *had been transferred
   e.g. English speaker: In the third school in which I had been *transferred or trans*ferred

   The first speaker puts the stress on ‘had’ to highlight the passiveness of the transfer (i.e. against my will) but the English speaker might interpret this as ‘in spite of what you might think’. The latter is more likely to put the stress on the first or second syllable of ‘transferred’ which would convey normal information flow in the standard variety of English.
3. Prosody used to convey attitude in one system but information or speaker’s perspective on knowledge in another. Whereas the low fall pitch movement in English routinely indicates a statement, new information or proclaiming, in Panjabi, it would have associations of lack of interest, boredom or rudeness. The English high fall which indicates contrast or emphasis can convey to a Panjabi speaker irritation or some kind of rebuke. In Panjabi, the low rising pitch movement is common to show respect but if transferred to English it can suggest uncertainty.

4. Different ways of conveying importance. North Indian speakers of English may use slow rhythm and highly contoured intonation to show that something is important or to show a lot of personal concern. English speakers tend to ignore this or find it too mannered or emotional (see for example the interview in the Introduction where Panjabi and English viewers interpreted the Panjabi speaking candidate in very different ways).

Other examples of misunderstandings and miscommunication where speakers to not share the same prosodic system:

- Frederick Erickson and Jeffrey Shultz, in their study of intercultural counselling interviews (1982), explored the phenomenon of what they called ‘talking down’. This frequently occurred when white speakers in some position of power and authority used a style of communicating which prevent the black American speakers from putting their case. White speakers tended to take long turns when explaining a point and when they did not receive the expected listening response or turn taking from the black speaker would re-explain the point in a simpler, less abstract or less structurally complex way. This re-explaining Erickson calls hyper-explanation.

At least one reason for this phenomenon, Erickson argues, is because of different prosodic cues used by the white and black speakers. White speakers used less obvious ways of signalling they had finished their turn: they tended to finish their turn with a sustained contour over a chunk of language and then give a pause. The black speakers tended to use a steeply falling intonation contour to finish their utterances, although it did not always come right at the end. Sometimes a filler like ‘you know’ came after the steep fall. So they tended not to interpret the white speakers’ less marked intonation to show completion of their turn.

- Ron and Suzanne Scollon have studied the language use of the Athabaskan people of Alaska. Where standard variety speakers of English would use prosody to show emphasis, Athabaskan speakers uses emphatic particles such as k’é or ‘eku. When Athabaskan speakers use English they may use these particles to indicate that one section of the narrative is finished and another is about to begin. ‘Eku’ is translated as ‘and’ or ‘and then’. To English speakers this often sounds like hesitation and they wait for something more when the Athabaskan speaker has concluded.
There are two other ways in which turn taking can be a problem in intercultural communication between Athabaskan and English speakers. In Chipewyan, a high tone is often used on the last syllable of a clause to show it is subordinated. English speakers hear this raised pitch as a question, take over the floor and answer when the Chipweyan speaker is in mid-flow. In addition, the Athabaskan speakers tend to pause longer than English speakers between tone groups or information units and so, again, English speakers tend to jump in.

- Sarah Michael’s study of white and black American children in first and second grade classrooms showed differences in narrative style which led to indirect discrimination against them. Black children’s narrative style was frequently misinterpreted by white teachers. The black children were interrupted more and given less ‘scaffolding’ support than white children. This had a knock-on effect on how the children were assessed as potentially good readers and the black children tended to be placed in the lowest reading group.

A summary of misunderstandings

- misinterpretation of attitude e.g. ‘over-emotional’, ‘passive’, sincerity perceived as uncertainty etc.
- apparent lack of coherence because of difficulty in distinguishing between main and subordinate points, contrastiveness, emphasis and so on.
- misinterpretation of intent e.g. where the speaker is trying to convey their perspective but it is perceived as contradicting the previous speaker.
- turn-taking problems where speakers find they cannot get a word in edgeways.
- general confusion because the line of argument is not clear to the listener and the relationship between ideas cannot be disambiguated.
- uncomfortable moments where the conversation lacks rhythmic co-ordination.

Recent approaches

More recent work, very much influenced by Gumperz, has taken the issues of prosody in interaction further. This recent approach combines Conversation Analysis methods with the insights of intonation specialists such as Elizabeth Couper-Kuhlen, Margaret Seltig and Susanne Günthner in Germany (Cooper-Kuhlen and Seltig 1996, Gunthner 1996) and Local, Kelly and Wells in Britain (Local 1992, Local and Kelly 1986 and Local, Wells and Sebba 1986). This section will concentrate on the work of the intonation specialists in Germany since they have been particularly influenced by John Gumperz and so their approach is especially relevant to the CDROM and related discussion. This method can be summed up as follows:

- Focus on interactional analysis of prosody in talk i.e. Be aware of all aspects of interaction especially how speakers take and compete for turns, manage repairs, negotiate meaning etc. and not just on how they convey information or attitude. In other words how prosody is used to accomplish interactional goals.
- Prosody refers to pragmatic meaning rather than grammatical meaning i.e. the use of prosody to cue inferences which lead to a particular interpretation at a particular moment in the interaction. So prosody is an important element in contextualisation. Rather than tone units which are associated with clauses in the grammatical system, it is more useful to talk about prosodic units which are associated with speaker’s turns.
• Prosodic features are sensitive to the particular context in which they are used and which they help to create - in particular to where they are placed in the sequence of turns.
• Prosodic cues are important only in as far as the participants attend to them and make inferences from them.
• Prosodic analysis does not rely on the analysts’ intuitions of what a particular pitch movement or pause means. Instead, they attend to how the participants’ handle prosodic cues. In this way, prosodic analysis is attuned to the assumptions of Conversation Analysis. In other words, how the participants orientate to each other’s prosodic cues is a warrant for the analyst’s interpretation.
• Prosody can be seen as one of the orderly details of interaction just as CA has described such orderly details as preference organisation and repair sequences etc. So they are members’ devices just as grammar, bodily movement etc. are devices for the organisation and management of talk.
• The basic unit of conversational prosody must be understood in relation to a turn-constructional unit. So, instead of looking at sentence level prosody, the analyst must look at how a speaker constructs their turn and how this turn relates to/ compares with their other turns and other speaker’s turns. For example, a low pitch and fast tempo may mean this is just an aside and not important or it may, when compared to the speaker’s other turns and understood within the particular context in which it occurred, suggest grave concern, anxiety, rejection and so on. Or, if a speaker’s turn is interrupted by another, their prosody when they get the turn back again will show if what they are saying is a continuation from before (what Langford 1994 calls ‘re-starts’). Understanding what meanings a particular set of prosodic features have depends upon the particular local context.
• Similarly, prosody does not just map onto a sentence-like utterance in a finished way e.g. a smooth intonational contour from the beginning to a low fall at the end. Instead it works dynamically, depending on each moment in the utterance where the listener(s) may come in with their turn (what CA calls a transition relevance place - TRP). So prosodic units work contingently. For example, a speaker may take a turn and make a point and when the listener does not come in after a pause, the first speaker continues. But it is not just a continuation of the last utterance, the speaker designs her next utterance, prosodically, as a new contribution and so shows, retrospectively, that her last utterance was complete (see Couper-Kuhlen and Seltig :29-30).

Some of the examples from ‘Interviewing in intercultural situations’ illustrate a number of these points well. For example, in point number 5, on showing attitude and affect, in the section above Why analyse conversation prosodically?, T’s low pitch is not an absolute pitch which always conveys a certain perspective but is related to his earlier low pitch to convey old information.

The following short data extract from Susanne Günthner’s work on how prosodic features help speakers to distinguish between different activity types will be used to illustrate these points further. She has looked at the role of prosody, in conjunction with other features, in cueing the listener in to ‘reproachful talk’ (Günthner 1996). In particular, she has looked at reproaches packaged as ‘why’ formats: For example, ‘why do you always let her in’.
Her concern is to find out how interactants differentiate between real ‘why’ questions and reproach formats which begin with ‘why’. In the following example, a caller has asked telephone information for a the number of a family in Konstanz. When they cannot be traced, the caller reformulates her request:

14 s: ja die wohnen glaub ich auf der * Reichenau.
    I think they live on the Reichenau
15      und gar nicht direkt * in Konstanz
    and actually not directly in Konstanz
16 a:   WARUM = * SA: ↑↓ GEN = SIE = DANN = KONSTANZ
    then why did you say Konstanz
17 s:  tut mir leid, ich dachte die Reichenau fällt unter * Konstanz.
    I am sorry I thought Reichenau belonged to Konstanz
18   (2.5)
19 a:  Also die * Nummer ist
    okay the number is

(p284)

[transcription conventions - see GAT notation]

(0.5)         pauses of 0.5 seconds or longer
word=word=word fast tempo
.             intonation phrase-final: falling
,             intonation phrase-final: slightly falling
word ↑↓      rise-fall
* WORD       primary accent of the intonational phase
a:           lengthening
no            strong accent, emphasised ]

Günthener argues that it is the prosody here which is the major cue to the caller that she should interpret the ‘why’ format as a reproach. First of all the ‘why’ utterance has a falling terminal pitch. Secondly, the turn as a whole is spoken with increased loudness and tempo. Thirdly the ‘why’ utterance has an extra strong and loud accent on the verb * SA  ↑↓ GEN with as extreme rising and falling glide on the first lengthened syllable.

Günthener argues that there is no single prosodic feature which, in a context free way, constitutes a ‘reproachful voice’. Rather it is the interplay of prosodic cues in combination with other linguistic and rhetorical features which produces a context in which ‘why’ formats are interpreted as reproachful. Although falling terminal pitch and rising - falling pitch movements are activated in all ‘why’ formats, they do not automatically mean reproachfulness. But they can mean this when combined with other features.
As well as illustrating the reproach theme, this example can also be used to illustrate the more general points made above:

*Focus on the interactional analysis of prosody in talk:* here the telephone operator is not only showing her displeasure in some general way, but is negotiating a particular meaning at this precise moment in the interaction.

*Pragmatic meaning:* the falling pitch, loud accent and so on are not associated with the grammar but with the overall turn i.e. the way in which reproachfulness is contextualised.

*Sensitive to sequence:* Line 16 coming as it does after the caller’s clarification and before her explanation, opens up a space for the telephone employer to be reproachful.

*Prosodic cues are important only in as far as participants attend to them:* The caller’s apology at line 17 shows that she attended to the reproachful message of line 16.

*How participants handle prosodic cues:* It is not the analyst’s intuitions that count but, in the CA tradition, how members i.e. participants orientate to the prosody that counts. So, the analysis of line 16 depends upon how the caller reacts in line 17.

*Prosodic cues as orderly devices:* Prosody functions in an orderly way like adjacency pairs etc. This does not mean that, for example, raised tempo always means the same thing, whatever the context, but that raised tempo will produce a reproachful response in a particular local context and will in turn produce the kind of expected response that we see in line 17.

*Prosody in relation to turn construction units:* The ‘reproachful’ utterance is spoken with raised volume and increased tempo which, coming immediately after the caller’s quieter explanation, clearly marks it as exceptional. The lengthened syllable and rise fall on ‘SA’ (the first syllable of ‘say’) foregrounds the action of ‘saying Konstanz’ and shows it up to be a mistake on the part of the caller.

So, this prosodic approach to interaction:
- focuses on how interactants interpret each other and attends to those prosodic features which they display as relevant to them
- requires the analyst to look *holistically* at the whole interaction to see how each speaker uses tempo, rhythm, pitch shifts etc. and so can compare a particular local production of speech with speakers’ previous performance
- looks at on the spot, emergent decisions by speakers e.g. a speaker may use the same pitch level, loudness and tempo to show a continuation of their talk despite an apparent interruption.
- puts greater emphasis on rhythm and tempo in describing how interactions work e.g. the length of a perceived pause will depend upon the rhythm and tempo of the surrounding talk
is cautious about generalising about variation in styles of speaking. It can be misleading to identify features in the talk of a particular group of speakers and assume that these features will always be present no matter what the local context. It is possible to talk about mid-level pitch to show finality in standard variety of German or to discuss Punjabi influenced English or pausing in Black American vernacular English (see the examples above) but only if these features have been analysed within their local context and if speakers’ orientate to them.

When working on the CDROM, it is useful to bear these general points in mind since they will help you defend the decisions you have made about how to interpret the data.

Summary

1. Any analysis of talk which is concerned with its local interactional features must include some element of prosodic analysis.

2. Prosody functions to convey perspective, to help the listener select between meanings, to pick out a line of argument and to interpret the speaker’s role and maintain conversational turn-taking and involvement.

3. Transcribers must start by notating what they can hear, whether they are important to the participants or not and for this some guidance on how to chunk the stream of speech, indicate pauses, pitch changes etc. is important. However, the recent approaches described here argue that after a preliminary transcription, the transcriber needs to focus on the participants’ perspective - on those prosodic features which are relevant to them and not on some outsider set of categories.

4. The principles and the notation system of CA, despite some limitations, are an important resource for undertaking prosodic analysis.

5. The most recent approaches to prosodic analysis of conversation combine CA and studies of contextualisation with a more fine-grained analysis of intonation.

6. In real life situations, particularly in gatekeeping ones, clients and candidates are judged on the basis of their communicative style. A significant feature of this style is the range of prosodic features discussed above. Where speakers bring different communicative styles to the interaction, miscommunication and misjudgements may result. For applied linguists concerned with contributing to the solving of real world problems, an understanding of how prosody feeds into the negotiation of meaning and judgement of others is, by no means, a trivial matter.
References


Auer, P and di Luzio A (eds) 1992 *The Contextualisation of Language*. Amsterdam: John Benjamins


Chafe, W 1993 Prosodic and Functional Units of Language. In Edwards and Lampert (eds.)


Erickson, F and Shultz, J 1982 *The Counselor as Gatekeeper*. Academic Press

Gumperz, J 1982a *Discourse Strategies*. Cambridge: CUP


Gumperz, J 1992a Contextualisation and understanding. In Duranti and Goodwin (eds) 229 - 252

Gumperz, J 1992b Contextualisation Revisited. In Auer and di Luzio (eds) 39 - 54

Gumperz, J 1992c Interviewing in intercultural situations. In Drew and Heritage (eds) 302 - 327


Mishra, A 1982 Discovering connections. In Gumperz (ed.) 57-71


