An impertinent question: what happens in spontaneous speech?

Richard Cauldwell

In the early 1990s at the University of Birmingham I had the privilege of studying both with David Brazil and with the many visiting scholars who wanted to work on Discourse Intonation. David would often lead small seminars in which we would work on each other's transcriptions, and talk about our research. He was a great teacher, and a beautifully tactful person when people asked him difficult or impertinent questions. Below, I shall relate my experience of asking an impertinent question which arose from my experience of transcribing a lot of spontaneous speech. But before that, a reminder and a couple of definitions.

Reminder: Discourse Intonation and tones

Discourse Intonation (DI) holds that intonational meanings are discoursal, not grammatical nor attitudinal. For DI, English intonation has five tones: two rising tones, two falling tones and an 'opt-out' level tone. The view is that if a speaker chooses a falling tone, he is projecting the contents of the tone unit in which it occurs as 'news' for the hearer – it is a **proclaiming** tone. On the other hand, a choice of rising tone projects the contents of the tone unit as part of the 'shared world' of both the speaker and hearer – it is a **referring** tone.

Definitions: Thick and thin meanings

The term 'thick meaning' refers to a concrete specific meaning, which people can readily identify as being communicated in speech – meanings such as 'surprise' or 'anger'. In contrast the term 'thin meaning' refers to meanings which are more abstract, less easily identified – meanings such as 'proclaiming' and 'referring'. It seems to me that if tones do indeed carry meanings around with them wherever they go (i.e. they mean 'x' on every occurrence of the tone) then it is likely that those meanings are going to be thin, and not thick meanings. And this is what made DI's approach attractive to me, the meanings it ascribed to tones seemed to me insightful and plausible because of the very fact that they were thin.

An uncomfortable feeling

Nevertheless, after doing a lot of transcription of different types of speech, I developed the uncomfortable feeling that even DI's meanings were too thick. In particular, this statement from Brazil (1985/1997: 67) worried me (I will refer to it below as 'Quote 1') '...the meaning increment that any one of the five tones contributes is the same regardless of the environment.'

The inconveniences of spontaneous speech

Quote 1 worried me because of the sheer number of tones in the recordings of all the types of speech (poetry, interview, spontaneous speech) that I was transcribing. They occur in huge numbers and with a wide range of realisations. Occasionally they occur as clear, steep-contoured shapes, but most often they occur in shallow-contoured versions, which only trained transcribers can hear. And then there will be a large number of indeterminate tones, where even expert transcribers find it difficult to agree what tone was present.

Tones also occur in places which are very inconvenient for the purposes of explanation. For example, if you hold a thick-meaning view of tones and believe that falling tones signal 'completion', there is the inconvenient truth that there are far more occurrences of falling tone than there are of 'points of completion'. Or if you hold that rising tones mean 'uncertainty', there are far more rising tones than there are 'moments of uncertainty'. So the number of occurrences of a tone outweigh (hugely) the number of occurrences of the meanings they are supposed to convey. There are simply too many tones in too many different places for the meaning of a tone to apply to every instance on which it occurs.

Putting the question

So one day I plucked up courage and put the question to David – saying (something like) 'Do you really think that the proclaiming/referring meanings apply on every instance of a tone?' He smiled and licked his lips and said (brilliantly manoeuvring my question into a form he was happy to answer), 'Hmm, yes, I think the question you meant to ask was this: "Why is there this inconsistency in Discourse Intonation between the system of prominence – in which the meanings are all contextual – and the system of tone, in which the meanings are general and are held to apply in all instances?"'

He went on: 'What tones really mean is to be found in the immediate context of interaction: a falling tone means *not what a rising tone would have meant in this context* and a rising tone means *not what a falling tone would have meant in this context*. In other words tones take their meaning from the context in which they occur – they are all local, there are no general meanings. And very often, contextual circumstances are such that it does not matter which tone you use. The reason for my presenting the 'proclaiming' and

the 'referring' meanings is that teachers need something to hang on to, something to teach'.

This was a great reply, and I have been dwelling on the consequences of it ever since. I went back to the same page (Brazil, 1997: 67) where had found quote 1 and I found that David himself had written a question which I shall refer to as quote 2: 'What are the consequences of choosing tone x *in preference to another?'* [Emphasis in the original]. So his answer to my impertinent question (or the route to answering my question) was sitting there in his book – itself in the form of a question. And it was in very close proximity to the very statement (quote 1) which had worried me. So, for me the answer to the issue of the meaning of tones is to be found in considering the question in quote 2.

Meanings of tones

Most often, tones are unremarkable, swiftly passing shallow slopes in the rapidly moving contours of the stream of speech. They are simply one of the range of devices that speakers use to make their speech non-monotonic, non-predictably variable and therefore listenable-to. Tones do not carry meanings: they are a meaning-neutral phenomenon with the potential (shared with other prosodic phenomena) to highlight the presence of meanings which reside in the context.

For a tone to be associated with a meaning, there has to be a constellation of other phenomena coming together (contextual/prosodic) at a given moment. And when this constellation happens the tone may highlight their presence, *but it is not the tone itself that signals the meaning:* it is the in-the-moment constellation of features that is meaning bearing (cf. Cauldwell, 2013: Chapter 15). So, very often, the answer to the question in quote 2 is, 'There are no consequences for choosing one tone rather than another'.

Why then do course books and teacher-training books contain such (apparently) plausible and convincing demonstrations of the fact that tones carry meanings? Let us consider two examples, one from Halliday, one from Brazil. Halliday (1970: 27) gives the following examples of attitudinal meanings on a wh- question:

01 || ↘ WHAT'S the <u>TIME</u> || neutral

02 || ↗ WHAT'S the <u>TIME</u> || tentative or deferential

[This is not Halliday's notation. Upper case letters indicate prominent syllables, the arrow indicates the tone which starts on the underlined prominence.]

For Halliday, the falling tone on a wh- question has a 'neutral' meaning, whereas a rising tone it is 'tentative' or 'deferential'. Now this seems an entirely reasonable explanation, but only as long as you don't probe the assumptions embedded in this example by asking the question in quote 2.

This is a script, which occurs with two dimensions (at least) of context: first, it occurs in a language education context, where we are having the meanings of tones explained by a great linguist; second, we are given a strong contextual colouring in the labels 'neutral', 'tentative' and 'deferential'. Note that (and this is very important) these labels prime us to agree with the assertions about meaning that are being made. It is like an experimental demonstration where we are told the result, and then we are asked to perform the experiment in just such a way as to produce the desired result.

Such demonstrations utilise acted speech which we perform for ourselves (silently), or for our students (out loud), or play the associated (acted) recording. Both dimensions of context in this example predispose us to perform or hear these examples in the way the author wants. But if we probe the example by asking 'What are the consequences ...' then we will see that this simple relationship between tone and meaning does not survive.

Instead of the labels 'neutral', 'tentative' and 'deferential' we can substitute a set of labels with opposite meanings – for the opposite of 'neutral' we will use the labels 'loving' and 'aggressive'. You can say 01 with a 'loving' falling tone, an 'aggressive' falling tone and you have then demonstrated that the falling tone 'means' these two labels, neither of which are neutral. You can say 02 with an 'aggressive' rising tone, or an 'insistent' rising tone both of which are close to being the opposite of 'tentative' or 'deferential', and claim that rising tone means these two labels. This done, you have demonstrated that it is not the tone that is conveying any meaning, because you have kept constant the presence of a particular tone, and varied the meanings. And you may well have used other prosodic phenomena (cf. Crystal 1969: 177) associated with both the tone and the tone unit as a whole which contribute to the creation and perception of these other meanings. But the biggest creator of the meanings in each case is the fact that each contextual label primes you both to create and perceive the tone in a certain way. The explanatory power is in the suggestive priming of the label.

So I would urge anyone to evaluate any such statement by considering (perhaps impertinently) whether the opposite of any assertion is true 'Can I do a non-neutral version with falling tone?' or 'Can I do a non-tentative version with rising tone?'

If we turn to an example from Brazil (1997: 68-9) we will see – again – that we are primed to assent to the demonstration, which requires our participation in acting out a contextual role. The tone units below demonstrate the proclaiming and referring meanings of falling and fall-rise tones respectively.

01 || ↘↗ MAry BROWN || ↘ is a TEACHer ||

02 || ↘ MAry BROWN || ↘↗ is a TEACHer ||

Example 1 has a fall-rise tone followed by a falling tone, 'Talking of Mary Brown (fall-rise = referring), she's a teacher (fall = proclaiming)' and the meanings of the tones are glossed as 'we both know we are talking about Mary Brown, and I am telling you she is a teacher'. Example 2 keeps the same wording, but reverses the order of tones, resulting in 'I am telling you that it is Mary Brown who is (what we have both been talking about) a teacher'.

Just as with the example from Halliday, we are primed to assent to the explanation of the relationship between tone and meaning. Such explanations are useful, because they give us something to teach, and (as I will mention below) they comprise the first step in a process to accommodate the reality of everyday speech. But when we move on through the process of accommodating to reality, we come across the number problem which I mentioned earlier.

Big numbers

I would argue that the number of falling tones seriously outnumbers the occurrences of the 'proclaiming' meaning, and the number of rising tones – similarly – seriously outnumbers the occurrences of 'referring' meanings. And I would add that because of the overwhelming numbers of falling tones and rising tones that are not associated with any meaning whatsoever, it is safest to say that tones don't mean anything in themselves; they are simply one of the prosodic features of speech (Crystal, 1969: 177) which constellate with other prosodic features to signal the presence of a contextually available meaning. The meaning does not reside in the tone.

So, for me, Brazil's proclaiming/referring meanings have the same status as all other meanings: their presence can be signalled by the use of tones and other prosodic devices, but they reside in the context of interaction. And the same will apply to every attempt to state the relationship between tone and meaning – whether you take a discoursal, grammatical, accentual, or attitudinal approach.

Something for teachers

The second part of David's reply was that teachers needed something they could get hold of, understand and teach.

Of course it is important to give something pedagogically useful to teachers – a simplification which 'works' in that it gives teacher-trainers and teachers something to teach, and learners something to be taught. And it is a sensible procedure to begin with a simplified model of how things are, then to demonstrate how the model fits with a small set of carefully chosen data, and subsequently to add further layers of explanation to the model which make it a closer fit to real-life data. This process is sometimes insightfully described as the process of decreasing deception. But that does not happen to the model of speech that we use in language teaching: we remain with the simple model, which matches our textbook language, and we remain in a state of deception about how speech really is.

And we are happy to be deceived; (Swift, 1710: 108) defines happiness as 'a perpetual Possession of being well deceived') because one of the advantages of the deception is that we are given things to do and say in classrooms and teacher-training sessions. We certainly seem to be content with fictions such as 'English is stress-timed' and 'Rising tones mean uncertainty' and 'Nuclear stress falls on the last lexical item' – and we invent act-out pseudo proofs (of the type we saw with the Halliday and Brazil examples) to 'demonstrate' them to ourselves, our teachers in training, and our students.

The problem is that we never get beyond these deceptions, and we need to get beyond them if we are going to teach learners to cope with the realities of spontaneous speech as they learn to listen. These pedagogic simplifications get in the way of effective learning – the simplifications become beliefs, which predisposes you to hear something a certain way – and becomes an obstacle, particularly in the teaching of listening.

The way forward is pointed already

Interestingly, just as Brazil's explanation of tones contained both the pedagogic simplification (quote 1) and a clue to his view of how things really are (quote 2), other authors do the same. John Wells (Wells, 2006: 91, and footnote) presents a table showing the relationship between 'default tones' and sentence types, listing meanings such as 'definitive' for a falling tone on a statement and 'encouraging' for a rising tone on a statement. However in a footnote he adds:

'It is not necessarily the case that the default tones, as described here, are statistically the most frequent. Nevertheless, it is at the very least pedagogically useful that there are default tones, and to regard any deviation from them as necessarily due to a reason that can in principle be made explicit' (footnote 91–92).

Here again is a hint that what is actually true about tones might be quite different from what we are given to teach.

Similarly, Roach (2009: 107–110) provides an explanation of the rhythms of English which succeeds in both presenting stress-timing theory and expressing doubts about its validity as a theory of language rhythms. The fact that he expresses doubts is no surprise as he is one of the key figures in the research that has refuted this theory (cf. Roach, 1982). But, as with Wells on tones, more space is devoted to the pedagogic deception than to the realities of spontaneous speech. This is a real difficulty in addressing the realities of spontaneous speech. My own chapter on the rhythms of spontaneous speech (Cauldwell, 2013: 131ff) suffers from a similar problem: I have to explain (and therefore give considerable space to) the refuted theory, before I feel able to explain the reality. The route to reality is signposted, but our pedagogically convenient beliefs are significant obstacles.

Conclusion

The language teaching profession needs to recognise that (a) experts in phonetics and phonology have something important to tell us about the realities of everyday spontaneous speech, and that (b) the ELT field needs to get to grips with these realities and welcome the insights into the complexities, however pedagogically inconvenient they may be. Sometimes I get the impression that our field (ELT) acts like a voracious monster which demands the repetition of comfortable deceptions in textbooks and teacher-training materials, rather than seeking to learn something new.

We need this 'something new' because, although textbook rules have the advantage that they are readily understandable, teachable, and learnable, they have the disadvantage that they are not based on accurate statements about what happens in reality. They are thus inadequate for the teaching of listening to spontaneous speech in the real world, where the listeners have to cope with whatever comes at them – and most of what comes at them is absolutely not the proper speech of textbook English.

So let us embrace this question, and welcome the answers which come our way, however 'difficult' they may seem:

What happens in spontaneous speech?

References

Brazil, D. (1997). *The Communicative Value of Intonation in English.* [2nd Edition]. Cambridge: Cambridge University Press.

Cauldwell, R. (2013). *Phonology for Listening: Teaching the Stream of Speech.* Birmingham: Speech in Action.

Crystal, D. (1969). *Prosodic Systems and Intonation in English.* Cambridge: Cambridge University Press.

Halliday, M.A.K. (1970). *A Course in Spoken English: Intonation.* London: Oxford University Press.

Roach, P. (1982) On the distinction between 'stress-timed' and 'syllable-timed' languages. In D. Crystal (ed.) *Linguistic Controversies: Essays in Linguistic Theory and Practice*, 73-79. London: Edward Arnold.

Roach, P. (2009). *English Phonetics and Phonology: a Practical Course.* [4th Edition]. Cambridge: Cambridge University Press.

Swift, J. (1710/1975). *A Tale of a Tub and Other Satires.* London: J.M. Dent.

Wells, J.C. (2006). *English Intonation: an Introduction.* Cambridge: Cambridge University Press.

Richard Cauldwell has taught English in France, Hong Kong, Japan, and the UK, where he worked at the University of Birmingham. Since 2001 he has published electronic and print materials for listening and pronunciation. Two of his publications have won British Council Innovations in ELT awards (ELTons): *Streaming Speech* (online course) in 2004 and *Cool Speech* (iPad app) in 2013.

Email: richardcauldwell@me.com
Web site: www.speechinaction.com