

Prosodic resources for the management of turn-taking in Finnish

One of my interests as a phonologist concerns how, or perhaps whether, we put together CA's methodology for working on what most phoneticians still call "spontaneous speech" with theories of phonological structure. What would a linguistic theory based on talk in its most natural environment look like?

I'd like to explore what we mean by 'prosody'. My use of 'prosody' is drawn from the work of the Firthian linguists. The Firthians gave a privileged status to syntagmatic relations. Their phonological work shows their concern to find out for example how words are delimited from one another, and how they are held together. They were interested in how grammatical relations are expressed in the phonology. Prosodies for them are anything which relate to syntagmatic function. Of course, this includes, for example, intonation. But *a priori* there is no parameter which cannot be prosodic, because prosodies are determined not on a phonetic basis, but on a phonological basis, which for the Firthians included reference to how something functions.

Robins (1957) distinguishes two major functions of prosodies. I don't quite agree with him, but his distinction is helpful to newcomers to Firthian Prosodic Analysis.

According to Robins, prosodies can be 'extensional' or 'demarcative'. Extensional prosodies are a kind of glue which hold stretches of talk together. They produce internal coherence within units: pitch, nasality, and vowel quality and three common examples of extensional prosodies. Demarcative prosodies offset or delimit chunks of talk. They are like Trubetzkoy's *Grenzsignale*. They tell you when one piece of structure is over and another one beginning. Both extensional and demarcative prosodies have a clear syntagmatic function: they create and delimit chunks of local coherence.

The things that make it prosodic are its *function* and its *domain*. In the data in this paper, the function of the prosodies I examine is to mark relevant turn transition; the domain is over the end part of a turn.

So this is what I take from Firthian linguistics:

(Firstly) A concern for the establishment of phonological systems with formal categories, taking into account both what is paradigmatic and contrastive, and what holds together chunks of material like words, phrases and turns. Prosodies are a phonological resource for producing coherence in talk.

(Secondly) A concern for linguistic statements made at a number of levels, but all mutually compatible. At some point we have to reintegrate our analyses at various levels. What works at one level needs to be relatable to other levels.

Modern linguistics is rooted in approaches which make statements with respect to only one level at a time: for example phonology without syntax.

(Thirdly) Firth in 1957 and elsewhere insisted that the shape of language is determined by the occasions of its use and the social actions it promotes; and the environment of language, its context of situation is not separable from it, but embedded within it.

In the rest of this paper, I'll look at some of the prosodic resources for the management of turn-taking in Finnish. Producing locally coherent chunks and offsetting one thing from another are both needed in turn-taking activities.

There's very little literature on the linguistic work voice quality does, and with regard to turn-taking, most work has been done on intonation. So I'll concentrate on voice quality, but there will be other kinds of prosodies too. What you won't be hearing, though, is anything much about intonation. I want to show that creak (along with other non-modal voice qualities) is a phonetic resource that is used to signal transition relevance; and I want to explore some of the implications of this finding for phonological theory.

The data come from recordings of phone-in programmes made in Finland. There's more information on the handout. In the transcripts, P stands for the presenter, C for the caller.

creak and turn-taking

When I talk about 'creak', I mean a system of different kinds of voice quality which include voicelessness, whisper, breathiness, exhalation and creak. I use the word 'creak' because that is the most common form in my data. There's a bit more information about what I mean at 2.1 on the handout.

1. straightforward examples where creak (etc) leads to turn transition.

Turn-finality in Finnish is signalled using several resources: canonically, there is syntactic completion, a fall in pitch, and the current action is complete. There is also most commonly a change in voice quality, usually to creak. Creak in most cases then is a part of a more holistic structure of finality.

There are plenty of examples in section 3 on the handout where creak—marked with a capital C—occurs turn-finally.

In cases where there is turn-final creak followed by a change of speaker, it is possible to demonstrate that both current speaker and next speaker orient to the relevance of turn-transition. Current speakers orient to it by stopping their talk soon after creak is initiated; next speakers orient to it by coming in either just

after a creaky stretch, or in overlap with it.

So in example 3.1 line 69 (*onks sulle tuttu tämmönen*) notice that at the end of the 1PP there is creak followed by whisper and then exhalation. **SHOW OVERHEAD OF SPECTROGRAM. PLAY TÄMMÖNEN.** This is followed by the 2PP. The same pattern is shown in the other examples.

This is a normal pattern for questions + answers, first and second assessments, and other adjacency pairs; but not for pairs of greetings, which are typically done with stylised pitch. ((Anything that deviates from this pattern involves some extra work.))

2. examples which show the independence of creak:

(a) rising intonation + creak

A common claim in the phonetics literature, for example Laver 1994, is that creak is a natural accompaniment to low pitch; so it's important to show cases where creak can be produced with high pitch. Other data like this has been collected by Sara Routarinne at Helsinki, looking at so-called High Rising Terminals in teenagers' speech.

Fragment 4.1 on page 4 shows that creak and rising intonation can also occur simultaneously. P, in lines 16-18, is establishing C's connection with the group whose music he has requested. She does this using two yes/no questions, conjoined by *tai*, 'or'. These two questions are produced with rising intonation, where the accented syllable has low pitch, which rises to the end of the TCU. **PLAY THE EXAMPLE.** ((PITCH TRACE NOT POSSIBLE BECAUSE OF DIPLOPHONIA.))

((The f_0 of *soittaneet* in line 16 is at 165Hz and 184 Hz, while in *tunnetteko henkilökohtaisestikaan* it is at 155, 169 and 175Hz respectively for each of the underlined accented syllables.))

As well as rising intonation, each TCU is marked with creak finally. **THE OVERHEAD** shows this for *Leikkareis*, 'in Leikkarit', the final word of the first question. Notice that for the last vocalic portion, there is a change in the mode of vocal fold vibration and it's visible on the spectrogram. Instead of the glottal pulses being equally spaced, they are alternately louder and quieter, suggesting a change to a diplophonic mode of vibration. A similar spectrogram for the second question cannot be produced because the talk is in overlap.

Iivonen and others have shown that rising pitch is not associated with questions in Finnish. Here the questions are produced as part of a list. The yes-no format of the questions means that the questions themselves offer candidate answers.

Each of the two questions contain things which show P offering these as *candidate, tentative* answers. In the first one, the word *mahdollisesti*, ‘possibly’, and in the second one the clitic *-kaan* ‘at all’ mark out the provisional nature of the answers embedded within the questions.

The rising intonation here is not ‘question intonation’, but ‘list intonation’, and the rising intonation at the end of the second question implies that there may be a third list item; and if P does not supply this list item herself, then it is open also for C to do so, because there has been a TRP at the end of each of the two questions. Thus the list is constructed so as to facilitate co-construction by C.

Although both of the questions in this list is produced with rising pitch, the voice quality at the end of each list item marks the ending of that list item out as a TRP.

3. *creaky stretch + increment*

Now I’ll turn to a case where there’s a problem of how to join one thing to another, and which combines interactional work with phonological work and syntactic work: increments.

Increments—additions to talk that was hearable as complete—pose speakers with an interesting problem. An increment needs to be hearable as an attempt at finishing something that was already designed as finished. Increments are syntactically continuations of prior talk. Another way of seeing this is that they are syntactically coherent with prior talk; so we might expect *prosodic* coherence with prior talk as well. As one of my students, Gareth Walker has shown, the relation of the increment to the host is not straightforward.

The fragments under section 5 on the handout are increments to questions. The increments in these cases delimit the possible set of answers that are relevant: they narrow down the search space for a relevant answer to the question. This kind of increment is named “relevance delimiting increment” by Gareth Walker. In these cases, the phonetic properties of the increment match those of the host very closely. The end part of the host has the phonetics of turn-finality, and the increment is produced with the phonetics turn-finality throughout. The advantage of being designed like this is that the increment is produced as being a continued ending, and not as being a re-start, or a re-try. In other words, by being a continued ending, they are also a continued TRP.

I’ll talk more about Fragment 5.1. In the talk just before this extract, P has said a lot about the group that is playing C’s request. Her co-presenter, P2,

evidently knows that she has insider knowledge because she played in the group, and so she knows members of the group personally. P2's question in line 62 is designed then to allow P to explain the source of her knowledge to the audience. P comes in at line 64 and starts to answer, but P2 comes back at line 65. This increment delimits the answer space further: the question is now shifted away from being about the group to being about one of the members of the group. So this increment is a continuation of the question, and narrows down the answer space. It is this narrowed answer space that P takes up in her answer at line 66, even though she recycles the *no* at the start of this turn.

In line 63, the end of the turn is marked with creak; the question is syntactically and pragmatically complete at this point. The increment in line 65 continues with creaky voice. It is syntactically dependent on the prior TCU because the relative case form only makes sense in relation to the verb *tietää*, 'to know'. So the increment is formatted both syntactically and prosodically as a 'continued ending'.

What I've just shown is that when increments delimit the answer space of questions, their relation to their host is pretty straightforward with regard to voice quality: the voice quality of the increment matches the voice quality of the end of the host. At the end of the host, it signals transition relevance; it seems that the increments are designed as 'continued endings'.

However, increments are not *universally* marked with the same properties as the material they are attached to. Walker (2001) shows that at least in English, the phonetic relation between the host and the increment is dependent upon the interactional work which the increment performs. The Finnish data I have confirms this finding for Finnish.

Creak and other things.

Now I'd like to show an example where creak is just one of several prosodic resources used in the management of turn-taking. This is fragment 6.1 on p.7 of the handout. The presenter suggests to the caller that she puts the record on. The caller agrees. **PLAY FRAGMENT.**

This is an interesting fragment on several levels.

First, the 2PP of this pair re-uses the syntactic form of the 1PP: an impersonal form of the verb + an infinitive form in the illative case.

Second, it's a pun. The caller's request is entitled "The farm machines' day off". So "getting the machines turning" is both literal and figurative.

Thirdly, there are several interesting phonetic features to this stretch. Both turns are marked with creak at the end. But the turns are both also produced as very rhythmical. In coming in, the caller, places her incoming talk in overlap which starts on beat, as well as at a point where there is prosodic and syntactic completion are both clearly projectable.

SHOW OVERHEAD.

There's a consequence of rhythmicality here. In Finnish, plosives are not normally aspirated. But at the end of *koneet*, there is a clear portion of aspiration. If we look at what's in this foot as opposed to other feet, it's a shorter one. So one function the aspiration might be performing here is to fill in the space to the next beat. So here is a 'prosodic' thing which has an impact on 'segmental' detail.

It is relatively easy to find examples like this, where so-called 'prosodic' and 'segmental' detail aren't clearly separable. As modern phonological theory shows, many 'segmental' processes are restricted to a particular metrical domain; so that the prosodic hierarchy is a way of organising not just rhythm and intonation, but also organises some so-called 'segmental' processes.

conclusions

Let me try and draw this together. I've taken an old-fashioned view, which is at the same time a very radical one: you can't define prosodies phonetically; only phonologically in terms of syntagmatic relations. Defining 'prosody' this way means that it has to be defined with reference to function. Prosodies are things that produce chunks of coherence in talk. Prosodies join things together, and produce chunks of coherence which off-set some chunks against others.

Increments provide supporting evidence for this claim. The problem for a speaker in producing an increment is that of matching the increment to the host, so that it is hearable as a continuation of the host. Of course, intonation, tempo, voice quality and so on are also available to do this. But the point for phonology is that doing a "continued ending" is a question of generating local coherence, so it's natural that this should be done using prosodies.

WHAT ARE THE REPRESENTATIONAL ISSUES FOR PHONOLOGY

We've been asked to highlight methodological and theoretical issues. I'm a phonetician and a phonologist, so it's natural for me to ask how work described in this conference relates to other linguistic work. For phonology, the findings of interactional linguistics pose a set of representation issues:

- what are the units which best account for patterns we can observe in the data? Phonology talks in terms of things like ‘utterance’, and ‘intonational phrase’. Interactional Linguistics uses categories established interactionally: most obviously, ‘TCU’ and ‘TRP’. These need some account in phonology if they are phonetically delimited. This might be obvious to conversation analysts, but it’s a point that conventional phonology hasn’t got anything to say about at the moment. **MAKE AN OVERHEAD HERE SHOWING THE KIND OF THING (TREE + SPREADING)**
- One of the problems that phonology encounters is the time-bound nature of speech: phonological representations are ‘timeless’; but in extending a turn or a TRP the time-bound nature of talk is obvious.
- at some level, phonology, syntax and lexis all need to be represented so that the inter-dependencies between them are clear. This is needed so that things like increments can be modelled properly; so that Margret Selting’s view of the turn as a holistic construct among others, and which is supported by the data I have presented here from Finnish, can be formalised. This isn’t really possible in many linguistic formalisms, because they are based on assumptions which keep different linguistic components separate from one another. In one paper, Ogden 1999, I have tried to redress this somewhat, by using the signs of head-driven phrase structure grammar.

Like Firth, I think we should opt for prosody as a countable noun, not as an abstract noun. I hope I’ve shown you that we do indeed speak prosodies.

The challenging aspect to this view is that prosodies only make sense if they are warranted as rigorously as other categories of interactional linguistics: but this is exactly what makes the combined linguistic and conversation analytic approach such a potentially exciting one for developing our understanding of linguistic and conversational competence.