Prosodies in conversation

Richard Ogden
richard.ogden@york.ac.uk
University of York
Centre for Advanced Studies in Language and Communication
Department of Language & Linguistic Science
No study of meaning apart from a complete context can be taken seriously’

(Firth 1935a: 7).

The ‘prosodies’ of the title of this paper harks back to the title of Firth’s seminal paper from 1948, Sounds and Prosodies, where he pursues the idea that syntagmatic relations in linguistic (and especially phonological) analysis should take a central role. ‘Prosody’ is the name Firth gives to phonological units which handle syntagmatic relations.

In the Firthian conception, prosodies are phonological units with a structural domain. They are defined syntagmatically, and do not straightforwardly correspond with aspects of the speech signal that are normally thought of as prosodic in modern phonology, such as pitch, voice quality or tempo. It is implicit in Firth’s (and his colleagues’) work that prosodies are limited to the utterances of a single speaker. Today I will argue that prosodies can be extended to embrace consistent and meaningful relations between turns at talk in conversation: at least in some places in conversation, the phonetic design of one speaker’s talk stands in a meaningful relation to the talk of another. This conception of prosodies is not an original Firthian one, but is consistent with Firth’s concern that meaning is primarily social, because it is shared by people in society and contextualised by social settings.

Firth’s concept of prosody was, and remains, a radical one. First we will consider some of the main features of Firthian prosodies, and explore the concept of prosodies as phonological terms, established on the basis of syntagmatic relations as opposed to paradigmatic and merely locally contrastive ones, and which bind linguistic elements together. This lays the ground for extending prosodies to being units that hold between turns of talk by different speakers.
This is the fricative system of English, as established by conventional phonemic analysis. We will come back to this later: for now, the point to note is that it looks like a symmetrical system with four places of articulation and two degrees of voicing. However, real life is much more complicated than this set-up implies; and it also indicates that phonetic detail, grammar and phonology are more closely embedded than we might imagine.
in the pen, all the people; in thin pens, all thin people

\[
\begin{array}{cccc}
\text{n} & \text{v} & \text{l} & \text{i} \\
\text{n} & \text{v} & \text{l} & \text{i} \\
\ast \text{n} & \text{v} & \text{l} & \text{i} \\
\ast \text{l} & \text{v} & \text{i} \\
\end{array}
\]

This slide shows some possible and impossible sounds of English across word joins. Notice that for the ‘voiced fricative’, assimilation is possible; but the same is not true for the voiceless fricative. To put this another way: the definite article in English requires only voicing and dentality, so we get minimal pairs such as “in a park” vs. “in the park”, where the indefinite article has alveolarity, shortness and clear secondary resonances; but the definite article has dentality, length and dark secondary resonances. This is a pattern researched by, among others, Sharon Manuel.

Why should we get patterns like this?
The answer lies in the distribution of the two phonemes in English words. /th/ only occurs initially in content words: nouns, adjectives and verbs. /dh/ on the other hand only occurs initially in content words: determiners, pronouns, adverbs and the like. Function words form a closed system: languages rarely develop new determiners or pronouns, whereas new content words are added all the time. Content words form a relatively small set, and in English tend to be unstressed. So this gives rise to the possibility of more variation. /th/ and /dh/ are not therefore in complementary distribution in real speech.
This example illustrates some of the important properties of Firthian phonology. Firstly, there is the notion that language consists of many interacting systems, which is to say that phonology is polysystemic. Systems can be established on many grounds: here we have set up at least two: word-initial consonants, and two lexical classes.

Secondly, the phonetic details that relate to phonological items are sensitive to systems. That is to say: to know how to do the phonetics of /dh/, we need to know about its location in a word, and the word’s place in the lexicon. The implication of this is that in Firthian phonology — which is the kind of phonology I think works well for conversation too — phonetic detail can reflect facts about linguistic systems beyond the merely phonological.
So to return to the phonemic fricatives of English: a closer look at just two items shows us that while the inventory as a whole is symmetrical, in terms of real-life distributions and phonetic patterns, /th/ and /dh/ are related in ways that are more complex than ±voice.

Why does this matter? Well, when a listener hears something long, voiced and dental, they know that they are hearing a piece of grammatical structure; and that in turn can help them to construct a lexical and syntactic representation of what they are hearing, and in turn help them to project where the current speaker is going with their talk.

I could unpick similar patterns for other pairs too, especially s:z and sh:zh; but that would take us away from my aims today. For now just bear in mind what I’ve done: deconstructed an apparently regular and obvious fact about English in order to reconstruct it. More on this later.
Syntagmatic relations

Firthian prosodies are intimately bound up with syntagmatic relations, that is, relations between elements of structure. Henderson's (1949) *Prosodies in Siamese* illustrates this well. This slide shows a part of her analysis dealing with differences between different speaking styles. Each style is associated with different 'linking prosodies', at the top of the slide, each with its own set of phonetic exponents.

The difference between styles is the choice of linking prosody to join syllables together. The linking prosody is superordinate to the syllable level. As can be seen, when the syllables [jjaːŋ] and [rai] are combined, they can be pronounced as either [jjanajrai] (Combinative style) or as [jjanai] (Rapid Combinative style). Henderson treats this as (i) differences in realisation of a sequence of phonematic units, (ii) differences in Length and Tone (phonological units). Thus the analysis avoids derivation of the kind common in e.g. generative analyses of 'casual' or 'fast' speech, and treats the differences as differences of phonetic exponency and distribution of prosodies, which are structural entities, and can have multiple exponents.

Later on, we will extend this form of analysis so that it handles the phonetic format of the speech of one speaker relative to the speech of another as it is manipulated in linguistically meaningful ways.
What I’d like to do now is to argue that we can extend the idea of prosody in Firthian linguistics to conversation. The connection is that many phenomena in the phonetics of conversation — and by that I mean all phonetics, because I reject the division between segments and supersegmentals as a false one — often depends on system and structure and syntagmatic relations between units. In the case of conversation, the units are TCUs or turns; and by implication they often involve more than speaker. But one of the essential points is that we need a phonology that can deal with people in interaction; a dynamic phonology that focuses on the process of speaking in time, and not just the product of speaking; and a crucially, a phonology that is able to handle many levels of linguistic knowledge.

So I’m going to focus on illustrating with some examples that will be familiar to anyone who’s read my work: agreement and disagreement; and marking contiguity and disjunction.
VT19.12.03 sea kale

Louder; very wide pitch span, high in speaker’s pitch register

C it’s a much more humble thing than truffles actually
it’s sea kale
P2 1→ no [it’s not humble | ‘it’s de`LICious Gail;]
P3 [ m:::?!
] P2 sorry [sorry it’s a fab[ulous thing sea kale
C 2→ <<<f,h> it ↑\!IS! delicious. > he he
VT19.12.03 nuts

Wider pitch span; louder

01  P3   and you wouldn’t believe how much (.) extra flavour
02    1→ it gives to them because just `RAW;,
03    [they’re `really]
04  P2   [     I would    ]
05  P3  1→ Tpre- Tpretty `BORing <<p> aren’t they.>
06  P2  2→ <<h> ↑'I don’t think they’re ‘BORing?"
CallHome 4544.65

Quieter, lower in register; narrower pitch span, faster articulation

01 B 1→ so `that’s `good for `YOUR department.
02
03 A 2→ .h <<all, l> Yeah it’s `GOOD>.
04 it was like back n forth is it
05 good isn’t it good but I think in the end they’re
goona keep it
CallHome en_4365.163-195
01 B 1→ I mean <<all> ↑`talk about a little bit> th-
  `RU:DE; I'm sorry
02 (0.9)
03 A [well ]
04 B 2→ [that's <<f> `!R:]EAL!lly `rude if you `ask `me;
CallHome en_4485.624
01 A 1→ it `sounds like it’s a `really good
02 A 1→ `programme and you’re `learning a `LOT.
03 (0.4)
04 B 2→ .h ↑`well (0.6)
05 2→ .h ↑its ↑<<len> it’s an <<f> `O`K> `programme.>
06 I think they’re horribly underfunded
NB II.2;12 sun coming out
01 E      <<ff> the `SUN’s comin’ out.>
02 N      <<pp> ↑I `know it>
03 E      1→ `BEAU:tiful.
04 N      2→ [<<pp> just `BEAUTiful.>
05        (0.3)
06 N      `SO:: `Anyway, `let me: uh .hhh `call Raul’s `MOther.
Phonetics of second assessments

Sensitive to at least:

• whether the second assessment projects continued on-topic talk

• whether the second assessment projects an upcoming disagreement

• phonetic features of the first assessment: pitch span, tempo, loudness, articulatory setting
holds between two turns in an adjacency pair. The phonetic design of a second assessment is sensitive to (at least) the following:

• Whether the second assessment projects continued on-topic talk, or not
• Whether the second assessment projects an upcoming disagreement, or not
• The phonetic features of the prior turn at least in terms of pitch span, tempo, loudness, and articulatory setting or 'tightness'.

Since the phonetic design of a turn affects how it will be interpreted, an adequate theory of speech communication must necessarily treat relations between turns at talk as having phonological implications.

4. Aligned and non-aligned actions in adjacent turns

The foregoing analysis showed that the design of a turn that responds to an initiating action can reveal important things about the relation of the second turn to its prior. In this section, we consider more general relations between turns, and argue that the phonetic design of turns at talk is in fact a syntagmatic relation which has consequences for how sequences of talk run off.

Example 10 is part of a complaint by Mum to her daughter Lesley. In this extract, we can see alignment and disalignment of various kinds and degrees. Louisa's memorial service has just been held and Mum is complaining about someone who had 'expected to be there'.

(10) Complaints #11 Holt (Christmas) l: l: l

01 Mum: how dare she expect to be there
02 Les: I know yes
03 Mum: she was `so `WICked to Louisa;
04 (0.6)
05 Les: mm hhmm
06 Mum: `all those `years a`GO.>
07 Les: <<all, p, breathy, l> `yes.>
08 Les: <<f> [KAY then love]> 
09 Mum: [as usual
10 if Louisa had known she wouldn’t’ve uh (0.5)
11 carted Mrs Field about like she did all the time

The next piece of data comes from a collection of complaints which I wrote about in 2010. This is just a very typical piece of talk, and the things I'm going to say about it are true of the vast majority pieces of conversation that you can look at. I want to argue that the phonetic design of turns at talk is in fact a syntagmatic relation which has consequences for how sequences of talk run off.

Louisa's memorial service has just been held and Mum is complaining about someone who had ‘expected to be there’. Our focus is on three pairs of utterances:
1. she was so wicked to Louisa... all those years ago
2. Mum’s all those years ago... and Lesley’s yes
3. Lesley’s yes and okay love
At line 03, Mum produces a strong assessment, she was so wicked to Louisa. As an assessment it provides Lesley with an opportunity to display agreement or disagreement in her next turn. By agreeing, Lesley would be joining in the complaining; by disagreeing, she would be engaging with the activity of complaining, but contesting the complaint; and by withholding an assessment, Lesley would not be engaging with the act of complaining. Mum produces the first part of this turn high in her pitch range, a common device in these sequences for seeking an affiliative response (Ogden 2010).

Lesley withholds a response, with a silence at 04; and then she produces what might be described as ‘fake’ or ‘stylised’ laughter (Haakana, ms): that is, she produces ‘mirthless’ laughter particles (Sacks 1974).
Mum’s next talk is not a new Turn Constructional Unit but an increment (Walker 2004) to her prior turn, all those years ago. It is a syntactic extension of her prior turn. As such, it must also be designed so as to sound like a continuation of Mum’s prior talk. If it were not designed so that it were hearable this way, then it would risk drawing attention of the failure of the first assessment to attract an aligning response from Lesley. Phonetically, this turn is designed as a continuation: its phonetic features match those of the post-accented syllables –cked to Louisa: they are in the same pitch range, and have matching voice quality and tempo.

Now let us consider the relation between all those years ago at 06 and Lesley’s yes at 07. In providing an increment to her assessment, Mum also provides Lesley with another chance to provide a display of alignment with her complaint. Lesley does this minimally at 07 with a yes; a token agreement with the action proposed in the just-completed turn. This yes is quiet and low in Lesley’s pitch range. It disengages from Mum’s complaint because it offers no more than an agreement particle. Lesley is not contesting Mum’s complaint, but she is also not joining in with the complaint or escalating it, two other possible courses of action at this place in the sequence of talk. Her turn at 07 presents only token agreement, and no engagement with Mum’s complaint. So here, the low and quiet features of the production stand in opposition to the possibility of a design that is high and loud, which would propose continuation of the activity of complaining.
Finally, let us consider how Lesley’s two turn constructional units, yes, and okay love, at 07–08 are related. With okay love, Lesley initiates an exit from the sequence with reset pitch, and a recognisable ‘opening of a closing’ (Schegloff & Sacks 1973). This turn is disjoint both in terms of the action it promotes and its phonetic design. Here, Lesley’s speech increases markedly in volume and in pitch as compared to her prior yes. Sequentially, yes belongs with the prior turn and proposes closing down the current sequence. Okay love proposes the start of a new sequence which is disjunct in activity from the prior complaint sequence. As it turns out, Mum resists this closure of the sequence by coming in in overlap; but the fact remains that for Lesley, this is a place at which the sequence can be appropriately closed.
holds between two turns in an adjacency pair. The phonetic design of a second assessment is sensitive to (at least) the following:

• Whether the second assessment projects continued on-topic talk, or not
• Whether the second assessment projects an upcoming disagreement, or not
• The phonetic features of the prior turn at least in terms of pitch span, tempo, loudness, and articulatory setting or 'tightness'.

Since the phonetic design of a turn affects how it will be interpreted, an adequate theory of speech communication must necessarily treat relations between turns at talk as having phonological implications.

4. Aligned and non-aligned actions in adjacent turns

The foregoing analysis showed that the design of a turn that responds to an initiating action can reveal important things about the relation of the second turn to its prior. In this section, we consider more general relations between turns, and argue that the phonetic design of turns at talk is in fact a syntagmatic relation which has consequences for how sequences of talk run off.

Example 10 is part of a complaint by Mum to her daughter Lesley. In this extract, we can see alignment and disalignment of various kinds and degrees. Louisa's memorial service has just been held and Mum is complaining about someone who had 'expected to be there'.

(10) Complaints #11 Holt (Christmas) 1: 1: 1

01 Mum: how dare she expect to be there
02 Les: I know yes
03 Mum: she was `so `WICked to Louisa;
04 (0.6)
05 Les: mm hhmm hm
06 Mum: `all those `years a`GO.
07 Les: `yes.
08 Les: <<f> `OK then love;
09 Mum: [as usual
10 if Louisa had known she wouldn`ve uh (0.5)
11 carted Mrs Field about like she did all the time

This brief example shows that the relationship between one turn constructional unit and another is managed simultaneously on a number of levels. Among other things, speakers can mark continuity and disjunction with their own talk or the talk of another; they can mark whether their talk follows the line of action projected in a prior turn, or rejects it somehow. Much work on the phonetics of talk in interaction shows this: see, for example, papers in Couper-Kuhlen & Ford (2004).
Consequences of studying interaction

• The phonetic design of a turn at talk is in part a product of the sequence it is in.

• For responsive actions, (not) matching aspects of a prior turn is likely to be meaningful.

• Turns at talk stand in syntagmatic relations to one another.

Linguists often conceive of ‘communication’ as the transfer of meaning from one person’s mind to another, and language is the most important vehicle for this transfer. However, much work on the phonetics of talk-in-interaction sees meaning as a jointly negotiated accomplishment of interaction.

Agreement and disagreement are socially negotiated practices. They are necessarily responsive actions, which means they are linked to another earlier action. In making a second assessment, (dis-)agreement is a relevant action in a way that it is not in making a first assessment. Such linked social actions are, according to Heritage (1984: 225), ‘basic building blocks of intersubjectivity’, “the maintenance of a world... mutually understood by the participants as some same world” (Schegloff 1991: 151). The phenomena considered in this paper are the phonetic and phonological expression of how intersubjectivity is maintained. Since the resources for marking alignment or disalignment with another’s talk are so precisely interwoven with the other’s talk, a more satisfying and realistic model for interaction is one which stresses the dynamics of interaction.
Some traditional prosodies

Prosodies are the ‘syntagmatic glue’ of speech. Their domains in the Firthian tradition include:

• syllable initial and final
• word initial and final
• word and morpheme boundaries
• phrases, including intonation phrases and utterances
Prosodies in conversation

- Prosodies are the ‘syntagmatic glue’ of speech.
- So we can also see them as a method for relating turns at talk.
As we saw, various types of agreement can be indicated by manipulating aspects of the phonetic production of second assessments. This means that a sequence of two turns at talk creates a linguistic structure with syntagmatic properties holding between them.

Let’s say speaker A produces a first assessment. Here it’s represented as a sketchy, outline structure. This has general linguistic properties, which in the tradition of some varieties of syntax I’ve called SYN, SEM, and PHON. The details don’t matter here; it’s the principle. The action type, First Assessment, brings with it implications about the linguistic design of the turn.

Now speaker B does a turn which is a second assessment. It’s hearable as one because its syntax and semantics mirror those of A’s turn. Let’s note that, and pass over it, as a related, but different problem. We’re interested in the phonological structure of B’s turn: it relates to A’s turn by virtue of a syntagmatic relation, which is that it is upgraded or downgraded relative to A’s turn.

Why is this a phonological relationship, and not a phonetic one? The answer is that ‘upgrading’ is part of a system of meaningful contrasts. If, instead of upgrading a second assessment the speaker downgrades it, then the turn is not treated as marking agreement, but as either marking disagreement or at best prefacing a disagreement. Alternatively, ‘downgrading’ can handle ‘doing agreement in order to close the sequence down’. So ‘upgrading’ stands in opposition to ‘downgrading’, and there is a meaningful distinction between the two.
Syntagmatic relations in conversation

A: 

<table>
<thead>
<tr>
<th>1st assessment: [SYN, syn]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[SEM, sem]</td>
</tr>
<tr>
<td>[PHON, π]</td>
</tr>
</tbody>
</table>

B: 

<table>
<thead>
<tr>
<th>2nd assessment: [SYN, syn']</th>
</tr>
</thead>
<tbody>
<tr>
<td>[SEM, sem']</td>
</tr>
<tr>
<td>[PHON, (ug/dg)π]</td>
</tr>
</tbody>
</table>

Upgrading and downgrading are prosodies in the Firthian sense: they express a relationship between A’s and B’s turns.

This statement is context-bound: that is, it is a statement about how agreement is achieved in the circumstances described, and in the social setting described: the statement is socially situated. This is important, because ‘agreement’ is an order of social meaning. Agreement only makes sense if one can agree with someone else.

Our structure represents the fact that language is multi-layered. The lexical, syntactic and sequential design are all inherently bound up with the prosodic design of the turns at talk. Our analysis explores these connections between disparate levels of statement.

The analysis is partial, and leaves open the possibility that there are other statements to make about the way these turns at talk work, and the same, or related practices, may be implicated in conveying meaning that is generalisable in other ways.
Prosodies, not prosody

• Prosodies are linguistic features that bind structures together within the speech of one or more individuals.

• Prosodies are constructions: regular, socially conventional mappings between form and meaning.

• Prosodies can be formalised as constraints over larger pieces of structure, relating information at various levels: lexical, syntactic, sequential, and the level of action.

• This understanding of prosody gives linguists a way to understand phonetic detail, and gives conversation analysts a way to talk to linguists.

Let’s draw this together.
I’d like to point again to the use of prosody as a countable item, with a plural. I’d rather think of prosody as something with phonological status, not just a set of phenomena.
Prosodies capture syntagmatic relations: they bind other items together. I’m suggesting we think of prosodies as not just in the speech of one person, but as in the speech of interactants, binding together or keeping apart turns in sequences of talk. As phonological units, prosodies are constructions: they meet the requirements of linguistic signs in having form : meaning mappings which are conventional and socially negotiated and understood.
How can prosodies be formalised? Well, unlike prosody in the singular, prosodies must bind together information at several levels. We’ve seen evidence for lexical, syntactic and sequential connections; and the level of action is also implicated because a responsive action needs to be designed to be heard as responsive, and not for example initiating something new.
As a linguist I find this conception of prosodies useful because it enables me to understand how phonetic details work in conversation: something which I was brought up to believe was merely wishful thinking. But this conception of prosody is interdisciplinary: it can only work when we can all talk one another’s languages.
Implications:

- Take a holistic view of data
- Don’t be afraid of complexity: it is the richness of daily life
- Be sceptical about ‘automatic’ associations between form and meaning
- Be ready to deconstruct and reconstruct analytic categories

So let’s try to make this practical: what are the implications of the things I have been saying?
Firstly, take a holistic view of data. This means that a focus on one thing, like intonation, won’t tell you the answers you need. We need to understand all sorts of issues: sequential, interactional, phonological, including prosodic/metrical organisation.
Secondly: linguists like to simplify data. This is why many of us were brought up with citation forms in phonetics. But everyday life is much richer and more complex. We need to embrace complexity, and accept that we can normally only develop partial accounts of our data. But the data is still there for more analysis, just as Firth said it would be.
Thirdly: our hunch is often that things mean something. There are plenty of models of linguistic theory that encourage this. But the more we look at conversational data, the more we see that many things we took for granted are, surprisingly, not like how we imagined. And even if they are, we see that they are usually richer and more nuanced than we thought. So be sceptical if anyone tells you that such-and-such a prosodic phenomenon ‘means’ something.
Fourthly: we inherit categories from our own disciplines. Sometimes these are useful and helpful; sometimes they aren’t. We have to be ready to deconstruct them and then reconstruct them. Here’s a couple of examples. We don’t talk in sentences, but in turn constructional units. Some sentences, such as pivots, are typologically weird. But having deconstructed sentences, when we see them for real, we have a better understanding of them. The same is true for intonational phrases. The view from conversation is that intonational phrases of the kind we find described in the literature are just a subset of all the kinds of intonational units we need for conversation. And the view from conversation is that meaning comes about through connections and contrasts, through syntagmatic and paradigmatic relations. And Prosodies are about connections, about syntagmatic relations.