Phonetic detail in talk-in-interaction: on the deployment and interplay of sequential context and phonetic resources

John Local

Department of Language & Linguistic Science, University of York, UK
<lang4@york.ac.uk>

Abstract

This paper explores some methodological and empirical issues concerning phonetic detail and phonetic variability and the work they accomplish in talk-in-interaction. I argue that if we wish to explicate how phonetic design contributes to the meaning of utterances we need to develop a more refined understanding of phonetic detail and its relationships with «context» and «communicative function». I show that by paying close attention to «sequential context» it is possible to document the variable relevancies and communicative functions of phonetic parameters.

1.   Introduction

Surprisingly large numbers of linguists have been willing to operate with the convenient (and misleading) fiction that it is possible to study the phonetic and phonological organisation of speech independently of the interactional contexts in which it occurs. One outcome of this has been a willingness to treat prosodic features, especially intonation, as if they could be «stripped off» the utterances of which they are design features and assigned structure and meaning of their own. Viewed from the perspective of developing an understanding of the communicative functioning of phonetic parameters in speech this is neither an especially interesting or useful thing to do.

Work in phonetics and phonology has also routinely drawn a typological distinction between phonetic parameters such as pitch, loudness, tempo, rhythm and voice quality on one hand and vocalic and consonantal quality on the other. These two groups of parameters are typically allocated (usually without serious defence) to two different, independent phonological systems: «prosodic», «supra-segmental» versus «non-prosodic», «segmental». In part, this distinction has come about because of the ways linguists have treated (and given unwarranted importance to) lexical as opposed to non-lexical meaning («segments» are seen as key elements in encoding lexical contrast). In part it represents a particular kind of innocence

about how supra-segmental and segmental material could/should be treated phonetically and phonologically (cf. the non-segmental, long-domain interpretation of phonetic «segmental» material by Firthian Prosodic analysts (Firth 1948; Kelly and Local 1986).

I argue here that if we wish to explicate how phonetic detail and phonetic variability contribute to meaning in natural everyday talk relationships, we need to develop a more refined understanding of phonetic detail and its relationships with «interactional-sequential context» and «communicative function». We should take seriously the possibility that phonetic aspects of language should *in the first instance* be analysed and understood as shaped by interactional considerations. We should also be open-minded about the relevance of, and relationships between, phonetic parameters and not give analytic privilege to one kind of parameters over another.

2.   Phonetic and sequential organisation of everyday talk

In this section I present brief analyses of three different kinds of interactional practice which show the ways participants systematically draw phonetic parameters of all kinds in the design and interpretation of talk. Because these bundles of phonetic features cut across the traditional distinction between «prosodic» and «non-prosodic» features the analyses indicate that phonetic parameters are best treated as falling into «functional clusters», irrespective of their «prosodic» or «non-prosodic» status.

The analyses are the outcome of combining the rigorously empirical analytic techniques of Conversation Analysis (CA) (Drew 2005; Schegloff 2007) with parametric phonetic analysis (Abercrombie 1965a; Kelly and Local 1989a). (For more details see Auer et al. 1999; French and Local 1983; Couper-Kuhlen 2001; Curl 2005; Ford & Couper-Kuhlen 2004; Local, Kelly and Wells 1986; Local and Walker 2005; Odgen 2001, 2006; Selting 1996.) CA research has shown that participants systematically display, in the placement and design of their own talk, an understanding of each others' talk and of the actions which that talk implements. This means that we can use these displays to ground our analyses of phonetic organisation in the observable behaviours and reactions of the participants themselves. One important benefit of this is that it enables us to establish structural «sameness», and to compare «like with like», in a rigorous fashion.

The practices I explore here arise from the examination of many hundreds of cases in some 50 hours of recorded conversation which includes face-to-face talk, telephone calls and radio phone-ins. The recordings involve a range of speakers in terms of age, sex and social

class, a range of activities and a range of varieties of British and American English, including a number of non-standard varieties. The transcriptions of interaction given here are based on the conventions used in CA (Atkinson & Heritage 1984, ix–xvi). Turns at talk are shown sequentially down the page. Speakers are identified, by name or initial, at the beginning of a line. Audible in-breaths are indicated by sequences of «h»; increased duration is indicated by «:». Intervals of no talk are timed in seconds and durations are shown within parentheses; «(.)» indicates a brief no-talk interval of around 0.1-s. Vertically-aligned left square brackets indicate the start of talk from one speaker which overlaps that of another. Vertically-aligned right square brackets indicate the point of offset of overlapping talk. Turns of analytic focus are asterisked.

2.1   «Self-repetition and sequence closing»

The first practice I examine is a particular kind of clausal self-repetition which participants employ to close sequences of talk. I will refer to these as «doubles» (Curl et al. 2006). The analysis reveals that from a phonetic point of view speakers draw on a bundle of phonetic features and relationships between features which include tempo, loudness, pitch and constraints on articulatory variability. The data are drawn from a collection of 35 cases.

Examples 1 and 2 exemplify the canonical shape for the sequences in which doubles occur. That shape consists of six parts: some topic or sequence is made ripe for closing, followed by the mutual passing up on the opportunity to take an extended turn at talk (step 1). There is then some form of «appraisal» (step 2); these turns take a variety of different shapes, but all are recognisably produced as moves toward topic closure (see e.g. Drew and Holt 1998). These appraisals are followed by further passing up on the opportunity to take a turn (step 3). The appraisal is then redone (step 4), and followed again by the mutual passing up of the opportunity to take an extended turn at talk (step 5), which amounts, functionally, to the confirmation or acceptance of the move to topic closure. Finally, there is a next move to some new sequence (step 6), regularly, though not exclusively, performed by the «double»-producer. All of these steps are exemplified in Example (1). Prior to this example, both speakers have been assessing, and complaining about, work colleagues.

```
(1)   Holt.5.88.1.5.nevermind (telephone)
1    Rob:   you know she's very .hh sometimes she's quite
2           helpful and other times I feel you know I don't
3           know where I stand with her
4    Les:   no
5           (0.2)
6    Les:   no no
```

```
7*   Rob:    never mind
8            (.)
9*   Rob:    never m[ind
10   Les:           [no
11           (0.3)
12   Rob:    anyway (.) I will let you (0.2) [go
13   Les:                                    [oh yes
```

Robbie puts forward an assessment of a colleague at lines 1–3. The end of this turn marks a transition relevance place—a point of syntactic, prosodic, and pragmatic completion (see e.g. Sacks et al., 1974); thus it is relevant for a next speaker, in this case Leslie, to begin talking. Robbie's turn duly receives a response, «no», from Leslie (line 4) which acknowledges and agrees with Robbie's negatively framed turn at lines 1–3. She thus passes up on the opportunity to take a more extended turn and provides for the current speaker, Robbie, to continue. In the silence which follows, both participants are observably withholding any talk on this or indeed any other topic (line 5). Leslie then self-selects to talk but again passes on the opportunity to take an extended turn at talk. Her turn at line 6 shows characteristics of other designed-to-be and treated-as-complete utterances; that is, it slows down, shows none of the phonetic features of turn-holding (Local et al. 1986), and constitutes a complete, falling intonation phrase. Robbie's immediate starting up provides evidence that this talk from Leslie is hearably complete, and thus transition relevant. Robbie's «never mind» (line 7) marks a disengagement from the previous sequence of complaints in that it offers no further reporting of «complainables». Furthermore, this turn proposes that there is nothing more to be said about the problem and has no phonetic features which adumbrate more talk to come. In other words, talk from Leslie would appropriately follow it; however, none is forthcoming. Robbie then produces «never mind» again. This creates the «double». Leslie produces another aligning turn (line 10) in terminal overlap. This turn is sufficiently delayed relative to Robbie's talk at line 9 that it can be seen as responsive to it; furthermore, the turn shows no signs of competitive turn incomings (French and Local 1983), thus providing evidence of her own orientation to it as a legitimately placed turn. She makes, however, no attempt to take an extended turn at talk; nor does Robbie begin a turn, resulting in a 0.3s silence (line 11). This silence is brought to an end by Robbie's lexically explicit formulation that they begin a new sequence; in fact, she displays her orientation to the possibility of the call moving into closing at this point (Button 1987). Thus, the sequence consists of a pattern in which both participants disengage from a topic, one produces a repetition of his/her own talk (with no intervening move

to resurrect the failing topic), followed by a move to close the sequence.

Example (2) provides another example of this kind of sequence. It is taken from some way into a call made to the «Nightowls» late-night radio phone-in show, broadcast in the North East of England. Michelle (Mich) is an American, enrolled in a course at a local university. Talk has turned to some of the differences between British and American culture.

```
(2)   no.1.10.american.probably-are (radio phone-in)
1    Mich: you know that was my other biggest shock when I
2          got here .hhh was that there is cursing and
3          nudity on television
4    AR:   no
5    Mich: .hhhhhhh I was amazed
6    AR:   becaus[e i- in Ameri [ca     you     just] have
7    Mich:       [I::            [couldn't believe it]
8    AR:   preachers asking for money
9          (0.4)
10   Mich: well (.) true (.) but (.) you know
11         (0.3)
12   AR:   huh heh hih
13         (.)
14   Mich: 'cause
15   AR:   huh [huh]  [hih]
16   Mich:     [I  ] g[ue ]ss they think we're all sinners I
17         don't kn [ow
18   AR:   [huhhh huh hah   [hah hah hah
19   Mich: [hih hah huh huh [hih [huh .hh
20*  AR:                    [and [we
21         probably are
22   Mich: ah huh ((laugh))
23*  AR:   [we pr [o  ]bab [ly] are
24   Mich: [.hhh  [huh]    [hh
25   AR:   .hh that's lovely Michelle thank you for calling
26         (0.2)
27   Mich: thank yo[u
28   AR:   [what is your musical taste
```

((arrangements are made for AR to send Michelle some free CDs in return for her call; the call then runs to closing))

AR (the presenter) and Michelle have been engaged here in a teasing sequence regarding the appearance of televangelists on American (as opposed to British) television. At line 20 AR produces a summative appraisal of the prior talk formatted in such a way (i.e. beginning with «and») that it is linked to Michelle's talk at line 16. Although this turn starts in overlap with Michelle's laughter, Michelle stops and produces no more laughter or talk during the remainder of AR's turn, allowing him to complete his turn in the clear. This turn looks back to and agrees with Michelle's prior talk, but offers little more to take up.

Immediately on completion of AR's «and we probably are» Michelle produces more laughter. This laugh marks a passing up on an opportunity to take a turn at talk. Additionally, by not speaking here, Michelle offers no resistance to any move that AR might launch to close the sequence, or perhaps even the call. Following the laughter, AR repeats «we probably are» (line 23) simultaneous with more laughter from Michelle (line 24). In this way, both speakers are engaging in coordinated, though differently realised, displays of there being no more to say; AR in his production of «we probably are» as a double, and Michelle in her continued laughter. Following the double, AR moves into a new sequence in which he offers an appreciation of Michelle's call: «that's lovely Michelle thank you for calling» (line 25). This talk is also produced with a disjunctive step-up in pitch and loudness, characteristics which have been shown to have associations with the launching of new topics and sequences (Couper-Kuhlen 2003; Local and Walker 2004). Michelle collaborates with AR in this move to a new sequence with her reciprocal thanking (line 27).

Phonetic design

Doubles exhibit a number of recurrent phonetic characteristics, particularly in respect of their syllabic make-up, their pitch and their duration. For each instance in the collection

- the second part consists of the same lexical items with the same number of syllables as the first part; consonantal and vocalic shape display little variability;
- the second part has the same accentual pattern and same (falling) main pitch prominence as the first part;
- the main pitch prominence bearing syllable in the second part is shorter in duration than the main prominence bearing syllable in the first part;
- the second part is not noticeably louder or quieter than the first part.

One characteristic of talk-in-interaction is that the same lexical item (or string of lexical items) uttered by the same individual may have rather different phonetic shapes in different contexts. However, we find that when speakers redo lexical items in the second parts of doubles, they produce them with the same number of syllables and the same accentual patterns that those items had in the first part. In Example (2), for instance, AR produces "probably" with three syllables on both occasions (rather than with a two-syllable contracted form such as «prob'ly» in the second part). Additionally, one of the discriminably regular features of doubles is that speakers retain the rhythmic and accentual patterning of the first part when they produce

the second part. Specifically (i) words are metrically footed in the same way in the two parts of the double and (ii) the location of the main pitch prominence is the same in the two parts. (In the following transcriptions «/» is used to indicate the beginning of a metrical foot, underlining to indicate the location of main pitch prominence and «/ /» to indicate the end of an intonation phrase.) Thus, in Example (1) the first part of the double is footed such that both words are stressed and the main pitch prominence falls on the second word (/*never* /*mind*/ / /*never* /*mind*/ /). In Example (2) the repeated phrase «they probably are» is also produced with two metrical feet thus: *they /probably /are/ / they /probably /are/ /*

All of the doubles in the collection are characterised by being constructed as two distinct intonation phrases, each having a falling main pitch prominence. In addition, the second part of the double is typically lower in overall pitch than the first part, and the pitch range of the second part of the double is typically compressed relative to the first part. Each of these features is represented in the labelled F_0 traces of the double in Example (2) («we probably are») which are shown in Figure 1.
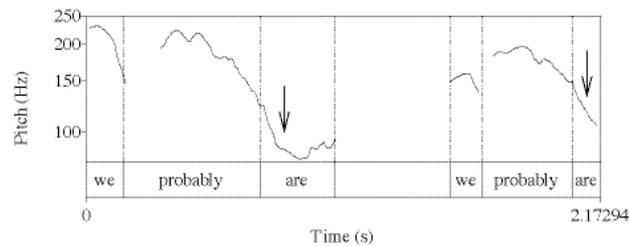
Figure 1(a): Falling main pitch prominence
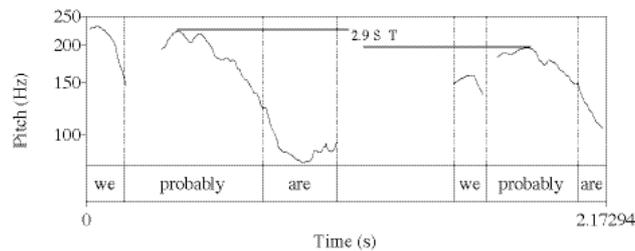
Figure 1(b): Lower overall pitch height in the second part
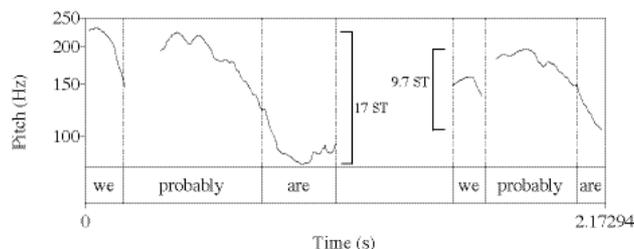
Figure 1(c): Compressed pitch range in the second part.

The arrows in Figure 1(a) show the falling main pitch prominence in each part; Figure 1(b) shows that the first metrically strong syllable of each part (the first syllable of «probably») is 2.9 semitones (ST) lower in the second part than in the first; Figure 1(c) shows that the first part has a pitch range of 17 ST whereas the second part has pitch range of 9.7 ST—a compression of more than 7 ST. Over the whole of the collection, the average difference between the pitch range of the first and second parts of doubles is 4 ST (with a maximum difference of 10.6 ST and a minimum difference of 2 ST). It can be seen, therefore, that doubles are associated with particular pitch relationships between the first and second parts, encompassing features of intonational phrasing, pitch height, and pitch range.

Over the collection of doubles, the second parts are shorter in duration than their first parts (the mean duration of the second part is 75% of the duration of the first part). The relatively faster production of the second parts of doubles is not achieved by uniform temporal compression of the repeated words and syllables. One systematic locus of temporal compression is the syllable bearing the main pitch prominence in the second, repeated part of the double. For all cases in the collection the pitch prominent syllable in the second part is noticeably shorter than its congener in the first part (mean: main pitch prominence in second part is 62% of the duration of the main accented syllable in the first part).

Regularities of two further phonetic characteristics are apparent across the collection. The first concerns constraints on variation in articulatory details between the two parts of the double. The second concerns the patterning of loudness between the two parts. Although it is well known that rather different productions may be recognised as the same lexical item(s), the articulatory details of doubles are organised in such a way as to *maximise* the similarity between the second parts and first parts. Other work on the articulatory variability observed in repetitions of individual lexical items has often attributed the variation to phonologically or physiologically triggered reduction

processes. Such variation, however, is not observed between the items in the first and second parts of doubles. There are not, for instance, systematic differences in the patterns of articulatory variation of the kind documented by Curl (2005). She found that some repetitions were produced to be maximally *distinct* from the first sayings, and that this was dependent on the sequential organisation of the talk *up to* that point. Such different phonetic realisations of re-done talk emphasise the importance of taking explicit and precise account of differences in sequential organization.

Finally, the first and second parts of doubles are loudness integrated, with no noticeable increases or decreases in overall loudness from the first part to the second. Among other functions, disjunctive step-ups in loudness have been identified as marking new sequences in talk-in-interaction (see for instance Local and Walker 2004). This lack of difference is one of the distinguishing characteristics of a «double», and contributes to the status of the second part *as* a second part. These systematic phonetic relationships between the first and second parts show that the practice of «doubles» is more than two consecutive utterances of the same lexical items; rather, it is the combination of the sequential location and phonetic design of the second part of a double, relative to the first part, which marks them as a distinct practice.

2.2   Standalone «so»

The «doubles» discussed in 2.1 manage aspects of sequence organisation. The analytic domain of the practice discussed in this section is primarily turn-taking organisation.

«So» can occur in a variety of sequential locations in talk —it frequently occurs in turn-initial and turn-final positions— with a range of functions. It may also occupy a «standalone» position in sequences of talk where it is separated from both the preceding and following talk by silence and may exhibit noticeable variability in its phonetic design (Local & Walker 2005). The tokens discussed here are drawn from some 130 instances in the transcribed parts of the CALLHOME American English Speech collection which consists of dual-channel recordings of 120 unscripted telephone conversations. The variability encompasses

- features of consonantal quality (e.g. audibly more or less tight articulatory occlusions associated with the initial consonant);
- vocalic quality (e.g. a range of monophthongal and diphthongal tokens which may be more or less central and/or more or less rounded at their beginnings and ends or throughout the token as a whole);

- duration (initial friction may be longer or shorter, vocalic portions may differ in their durations —some being noticeably long);
- laryngeal behaviour (phonatory types may differ);
- presence/absence of final supraglottal and glottal closures;
- differences in pitch contour, pitch relationships with surrounding talk and
- differences in loudness relationships with surrounding talk.

If the *sequential* context of the talk is taken into account, it becomes clear that variation in the phonetic design of «so» is implicated in the management of turn taking and the closure of topics. Examples (3) and (4) provide instances of the phenomenon.

```
(3)   6033-freewayCHAm
1     A:   I said what's the best to do take the freeway or
2          (w- w-)and take the city streets (0.4) the city
3          streets (0.3) don't take the freeway (0.3) you
4          know and so I thought (.) okay you know
5          know because that's (.) you know (0.2) high
6          high rise [you know] and I thought .hhhhhhhhh
7     B:             [mm   hm ]
8          (0.4)
9*    A:   so
10         (0.4)
11    A:   I started out (.) oh my gosh (0.2) pt<
12         .hhh I [  got   I ]
13    B:          [mmmmmmmmmmm]
14    A:   was just (.) I'm in (0.4) you know where I work
15         it's right down town
```

In example (3) as in all instances of stand-alone «so» silence sets off «so» from what preceded it, and what follows it. Here, «so» (after a silence of 0.4s) is followed by an on-topic continuation by the «so»-producer. The talk preceding the «so» is concerned with reporting a discussion which A had with work colleagues concerning the best route to take on a trip. Following the «so», A delivers what is clearly a next instalment of that reporting, continuing to recount how her journey was taken. There is no interactional evidence that when A talks at line 11, she has done something out of place or unwarranted or untoward: A's talk at line 11 is not designed with any features which might mark it out as a new action or new sequence, as sequentially misplaced or unoccasioned; speaker B does not attempt to come in immediately after the «so», nor does she produce any talk which could indicate inappropriateness of A's continuation; for instance, B could have talked in overlap with A at line 11 and designed her talk in such a way as to attempt to curtail A's turn (French & Local 1983). Indeed, when B does talk she produces an

extended, aligned in-overlap receipt of the launch of the next part of the telling (line 13). Example (4) shows an instance of «so» with a different interactional function.

```
(4)    4074-bizarreCHAm
 1    A:   it has an offboard power supply which they didn't
 2         steal
 3         (0.5)
 4    B:   hhhhh[h
 5    A:        [which makes the thing that they stoled
 6         absolutely worthless
 7         (1.1)
 8    A:   [huh huh
 9    B:   [hah
10    B:   hh (0.5) .hhhhhh
11         (.)
12    A:   .hhhhhhh
13         (0.2)
14*   A:   so
15         (0.2)
16    B:   bizarre
17         (0.2)
18    B:   bizarre
```

In (4), unlike (3), there is speaker change following the «so». Speaker A has been recounting at some length an incident in which various pieces of equipment had been stolen from his home. He brings the story to a possible conclusion with an assessment (lines 5-6). This does not get immediate uptake or appreciation from his co-participant, and neither speaker makes a move to take a turn. Instead, there are long silences, quiet laughter, and long inbreaths (lines 7-12). What is observable is a disengagement by both speakers from further on-topic talk. In producing a turn consisting entirely of the item «so» (line 14), A demonstrably does not produce talk which is topically linked with or topically develops prior talk nor does he take the opportunity to initiate talk on, for instance, a new topic. Subsequent to this turn neither speaker produces any further talk on the prior or indeed any other topic. Speaker A proffers no talk whatsoever while, following a silence, B produces at lines 16 and 18 a canonical «sequence closing double» (see section 2.1). The double provides an assessment of A's story («bizarre…bizarre»). In this interactional context, one of the functions of the «so»-turn is to indicate that, at this point, the speaker is not going to offer any further talk. In doing this the «so»-producer provides his co-participant with the opportunity to take a turn and initiate talk possibly with a new topic.

Phonetic design

The phonetic design of these «so» tokens displays a regular and consistent relationship with contrasting communicative functions: one

in which the same speaker continues with more on-topic talk («holding-so») and one where there is change in speakership «trailoff-so».

As a group the holding-so tokens :

- are noticeably louder than the same speaker's preceding talk;
- are noticeably higher in pitch than the speaker's preceding talk;
- have final glottal closure which may be accompanied by oral closure (final glottal closure is regularly held over the subsequent silence and released at the beginning of the speaker's next talk);
- may have a short period of final creaky voice before the final glottal closure but never elsewhere in the token (for those tokens with pre-glottal creaky voice Mean: 66.7ms (= 17% of voiced portion);
- may be accompanied by level, falling or falling-rising pitch.

By comparison as a group the trailoff-so tokens :

- are noticeably quieter than the speaker's preceding talk;
- are noticeably lower in pitch than the speaker's preceding talk;
- *never* have final glottal closure (though some have accompanying labial oral closure with voiceless egressive nasal airflow);
- may have creaky voice initially, medially, finally or throughout the whole of the voiced part of the token (for those tokens with creaky voice Mean: 203.6ms (= 80% of voiced portion);
- may be accompanied by level, falling or falling-rising pitch.

Importantly, there are no significant differences between the two groups in terms of the kinds of pitch contours which may occur (level, falling, falling-rising) or of the overall range of F_0 excursions or of the overall duration of the voiced portion of the «so» tokens. Nor are there any correlations between F_0 contours and phonatory features or between the alignment of pitch peaks and articulatory material. Kohler (1987) argued that for German an early F0 fall in the accented vowel, as against a rise-fall, strengthens low pitch and functions as an indicator of conclusion (as opposed to a new start) of an argument. One might hypothesise, therefore, that the two types of «so» token would differ in terms of pitch synchronisation with the vowel. However, this does not hold for the data-set: early F0 falls distribute across both types of «so» token.

Thus, unlike the data discussed in Section 2.1 where a designed-to-be «turn-final» pitch contour and a particular redoing of that pitch contour was required to accomplish a sequence-closing «double», the precise contour associated with «so» is of no sequential relevance. In

terms of pitch phenomena, it is relative pitch-height that matters (in concert with particular loudness and laryngeal features). This emphasises that in exploring the functional relevance of phonetic details it is important not to limit our focus to particular phonetic parameters. We should not assume that some phonetic parameter (e.g. pitch) is a priori of relevance whereas some other parameter (e.g. loudness, articulatory configuration, phonatory type, audible inbreathing) is not.

These findings are not simply related to productions of «so». For instance, they can be directly related to those of Local & Kelly (1986) who discuss the phonetic detail of transitions into and out of silence. They show *inter alia* that a highly recurrent sequential location for «holding» features of the kind discussed here is where a current speaking turn is subject to incursive talk. In such circumstances the current speaker regularly curtails their talk with glottal closure which is held until the incoming speaker has ceased to talk.

2.3   *News receipts*

It is well known that intonation patterns show different patterns of co-occurrence with different grammatical patterns. Here, however, I want to consider a different kind of distributional constraint on intonation, lexico-syntax and meaning which suggests that we should take care in treating intonational features as an «independent» strand which can simply be overlaid on segments and words. The final short piece of analysis considers the phonetic and sequential design of a small, but frequent speech particle «oh». The data is drawn from a collection of over 200 instances in some 5 hours of speech. This particle can occur in a number of contexts but a frequent one is in news-telling sequences (Heritage 1984). Someone in conversation imparts some news and the recipient produces «oh» proposing «that its producer has undergone some kind of change in his or her locally current state of knowledge, information or awareness» (Heritage 1984, 299).

«Oh» often appears in freestanding form in a turn responding to a-telling news. Freestanding tokens routinely terminate such telling or informing sequences. They are placed at points in talk where the informing in progress is possibly complete or may be strategically deployed to display that as far as the oh-producer is concerned the news-informing is for practical purposes complete. Examples (5) and (6) illustrate:

```
(5)   NB II.1: 2

1    Emma:   Bud's going play golf now up Riverside
2            he's just leaving
3                  (0.2)
```

```
4*   Lottie: Oh
5                  (0.5)
6    Emma:   so Katherine and Harry were supposed to come
7            down
8            last night but there was a death in the family
9            so
10           couldn't come so Bud's asked Bill to play with
11           the company deal so I guess he can play with
12           him so
13   Lottie: oh good
14   Emma:   what a miserable weekend
```

```
(6)   Rah II: 1
1    Jenny: hello there I rang you earlier but you were out
2    Ida:       must have been at Dez's mum's
3*   Jenny: oh
4    Ida:   but mind you we've been in a good hour and a
5           half to two
```

Evidence for the sequence terminating potential of these oh-tokens can be found in both sequential and phonetic aspects of the talk. As in example (5) and (6) we routinely find new topics (or reversions to previously curtailed topics) being started after such oh-productions. These topic changes are frequently lexically marked with disjunctions such as «but» (e.g. (6)) and with marked upgrading in pitch height and loudness features of the utterances (as at line 14 in (5)). Freestanding oh-tokens such as those in (5) and (6) are characterised a number of common features:

- they may or may not have an initial glottal stop but they never occur with a final glottal stop;

- they have utterance prominence (they are stressed); they are all done with terminal falling pitch movement (which ends low in the speaker's range); the range and starting pitch height varies from token to token: (5) has a mid-high start while (6) starts low;

- they are variably extended in time and done with tense articulatory setting.

- they may be accompanied by creaky voice quality

- they are typically diphthongal and close back, either throughout or in the closing part of the diphthong.

Freestanding oh-responses to question elicited informings have rather different properties and potentials with respect to the subsequent development of the interaction. They also typically have a different phonetic shape from other oh-tokens considered so far. Unlike other tokens, they may *end* with complete glottal closure and

they can be noticeably nasalised. Moreover, their vocalic quality is quite distinct from any of the oh-tokens considered so far. Most frequently, «oh» responses to question-elicited informings are realised as monophthongs. Typically, these monophthongs are back vocoids, usually open or half open. Qualities vary around cardinal vowels 5 and 6; if in the region of cardinal 6 the vocoid is routinely slightly unrounded.

The most striking intonational feature of freestanding oh-tokens is that they are routinely produced with *falling* pitch movement. One account for the design of these tokens with falls might be that such a contour here strongly projects finality/completeness (a common assertion in the intonation literature on English but see Local (1986) and Local, Kelly and Wells (1986) for a detailed refutation of this claim), and that coparticipants orient to this in not continuing with their talk or in proposing topic changes. What happens though if «oh» is produced with rising rather than falling pitch? Does this get the informer to progress the informing? Although such tokens with rising pitch are relatively rare they do occur. One such is shown in example (7).

```
(7)  NB II.2.2
1   Emma:    [Bud just left ]to play golf he's got to go to
3   Nancy:   [y   e  h    ah]
4   Emma:    Riverside
5   Nancy:   [o   h      ]
6   Emma:    [on a compan]y deal so, t.h[hhhh
7*  Nancy:                              [oh::
8   Emma:    go [it's bee-]
9   Nancy:      [to  River]side today
10  Emma:    .hhh yeah they they're going to tee off at twelve
11           it's a company deal so
```

The oh-token of interest is produced at line 7 at the end of Emma's turn, which begins with an out-of-the-blue announcement (line 1-2). Although Nancy produces a rising pitched oh-token, Emma does not orient to it as being a news-receipt which suggests that the telling is not yet complete and more information is being sought. That is, rising pitch in itself does not cue incompleteness. Although Emma is the person to produce the sequentially next talk, she does not produce more talk on Bud's golfing trip. Instead she begins an exclamation which prospectively opens up a new topic: «god it's bee-» (this gets redone a few turns later as «gosh this has really been a week hasn't it»). This utterance has the phonetic characteristics of new topic starts: specifically, it is louder and higher in pitch than preceding talk. It is not until Nancy produces the question-framed solicit «to Riverside today» that Emma provides an extended version of her news announcement. In this and other comparable cases, there is no

evidence that the pitch contour is central to determining the terminating potential of freestanding «oh».

As well as occurring in freestanding form (whether in response to out-of-the-blue or question-elicited informings) «oh» may also be accompanied in the same turn by other material (e.g. an «assessment»: «oh wow», «oh good», «oh dear» as in (8) or «really» as in (9)).

```
(8)    NB II. 2: 1
1    Emma:    .hh how you doing
2    Nancy:   .t hhh pretty good I got a raise h .hh[hh
3    Emma:                                          [goo[d
4    Nancy:                                              [yeah
5             two dollars a week
6                          (.)
7*   Emma:    oh[wow]
8    Nancy:    [uh ] uh hu[h          hu:h      ]
9    Emma:               [what are you going to]do with
10            it all
11   Nancy:   golly I really I just don't know how I'm going
12            to spend all that money


(9)    NB IV. 7: 6
1    Emma:    I've quit smoking you know and everything hh
2             (0.7)
3    Barbara:     well when did you stop that
4    Emma:        the day you left .h
5             (0.6)
6    Barbara:     left where
7    Emma:        from here in September
8    Barbara:     e-how many cigarettes you had
9             (0.5)
10   Emma:        none
11*  Barbara:     oh really
12   Emma:        no
13           (.)
14   Barbara:     very good
```

Like the freestanding oh-tokens discussed above, these «oh»+more turns are routinely placed at the termination of a topic/news-informing. In (8) Emma's «oh wow», which is produced in response to Nancy's news about being given a pay-rise, simply gets a laughter response from Nancy. It is not until the in-overlap question from Emma: «what are you going to do with it all» that Nancy produces further on-topic talk at lines 11-12. In (9) Barbara's «oh really» which responds to the prior informing (concerning Emma having given up smoking) is followed by a reconfirmation of the prior information: «no» from Emma, which in turn is followed by assessments from both participants. After this there is a reversion to a prior topic concerning Barbara's visit. (As Jefferson, (1981) notes these turns have the

sequential structure: (1) news announcement, (2) «oh really» (3) reconfirmation and (4) assessment (which is generally terminal or topic curtailing.) From a phonetic point of view these «oh+assessments» are rather different from freestanding «oh». While all the freestanding oh-tokens are produced with dynamic pitch movement, the pitch associated with «oh» in these oh+more turns may or may not be dynamic. Tokens of the kind illustrated in (8) and (9) also differ from freestanding oh-tokens in that they *all* begin with glottal closure and/or creaky voice. Figures 2 and 3 exemplify labelled F_0 traces for some tokens of «oh really» and «oh+assessments».
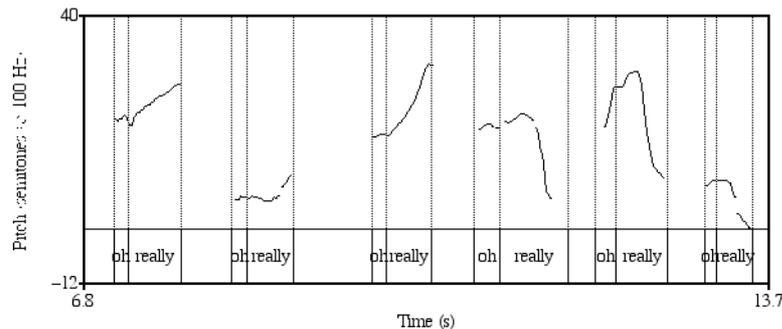
Figure 2: F_0 in semitones re 100Hz for six speakers' productions of «oh really»
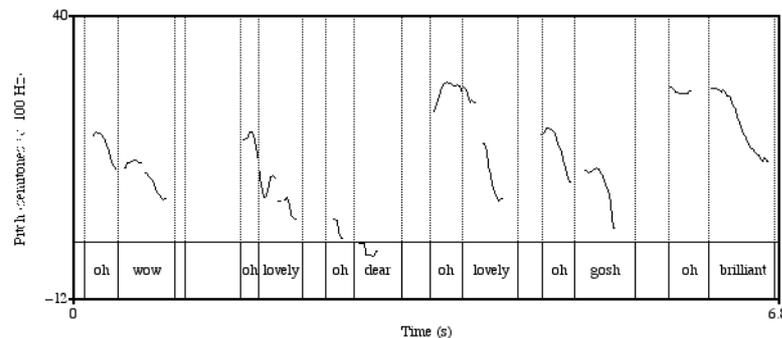
Figure 3: F_0 in semitones re 100Hz for six speakers' productions of «oh» + «assessment»

These figures show that while «oh really» may be produced with rising or falling pitch, news receipts such as «oh wow» are *never* produced with terminal dynamic rising pitch. At this place in sequential structure falling pitch is not in system (and therefore not in contrast) with other pitch choices. In retrospect, the association of

terminal falling pitch with oh+assessments may seem obvious given lexical content and the kinds of pragmatic work which these turn can be seen to be doing. But its obviousness trades on a naive and unexplicated sense of the «meaning» of rising pitch. As I have already suggested such an issue may not be nearly so transparent, or well understood as the literature on prosody and meaning might lead us to believe. It is important to register too, that the differences in pitch associated with «oh really» do not appear to be doing distinctive interactional work. Productions of «oh really» with rising pitch contours function in a similar fashion to all the other oh+assessment tokens and are similarly placed and treated in the course of the interaction—they occur at telling-termination points and they do not appear to engender more on-telling talk from the other participant despite the occurrence of rising pitch.

3.   Conclusion

The analyses I have presented illustrate a CA-informed approach to the interpretation of phonetic organization in everyday talk. My primary purpose has been to show that phonetic detail of all kinds is shaped by its particular place in sequence and the interactional work the talk is engaged in. Throughout, I have emphasised the need to extend our understanding of «context» to include, and take explicit account of, the sequential organisation of turns-at-talk. If this is done it becomes possible to document the differential and variable relevance of prosodic (and non-prosodic) parameters in different sequential environments. Just because some feature (e.g. pitch, tempo) is relevant to the shaping and interpretation talk in one environment does not mean that it is relevant (or relevant in the same way) in others. From a phonological point of view this should make us wary of setting up whole-language statements for such systems and making monosystemic claims about their phonological contrastive status.

As well as trying to understand how talk-in-interaction works and how particular phonetic features operate in its constitution, the work I have discussed here is also directed at changing the way we understand the conventional categories of description employed linguistic-phonetics and phonology. Results indicate that a priori assumptions about the putative importance or otherwise of particular phonetic parameters and what their function(s) might be are extremely dangerous. It takes serious interactional and phonetic analysis to show not only that something is there and systematic, but that it is also relevant to the participants. When (or indeed if) we wish to say things about the work that fine phonetic detail does in talk, interactional-phonetic work indicates that it is crucial to start with a sequentially grounded account of action. By doing this we can begin

to deconstruct and reconfigure our understanding of the constitutive elements of phonetics and phonology and begin to explicate in a serious fashion the different systems of phonological contrast which operate at different places in sequential organisation. Such an approach should significantly enhance our ability to give a robust account of both the cognitive processes involved in language production and understanding and the linguistic constitution of talk-in-interaction.

Acknowledgements

References

ABERCROMBIE, D. (1965a), «Parameters and phonemes», in Abercrombie, D. *Studies in Phonetics and Linguistics*, Oxford, Oxford, 120-124.

ABERCROMBIE, D. (1965b), «Conversation and spoken prose», in Abercrombie, D. *Studies in Phonetics and Linguistics*, Oxford, Oxford, 1–9.

ATKINSON, M. & J. HERITAGE (1984), *Structures of Social Action: Studies in Conversation Analysis*, Cambridge, Cambridge University Press.

AUER, P., E. COUPER-KUHLEN & F. MÜLLER (1999), *Language in Time: The Rhythm and Tempo of Spoken Interaction*, Oxford, Oxford University Press.

BUTTON, G. (1987), «Moving out of closings», in Button, G. & Lee, J.R. (eds), *Talk and Social organization*, Clevedon, Multilingual Matters, 101-151.

COUPER-KUHLEN, E. (2001), «Interactional prosody: High onsets in reason-for-the-call turns», *Language in Society* 30, 29–53.

COUPER-KUHLEN, E. (2003), «On initial boundary tones in conversation», in *Proceedings of the 15th International Congress of Phonetic Sciences*, 119-122.

CURL, T. S. (2005), «Practices in other-initiated repair resolution: the phonetic differentiation of 'repetitions'», *Discourse Processes* 39, 1–44.

CURL, T. S., J. LOCAL & G. WALKER (2006), «Repetition and the prosody-pragmatics interface», *Journal of Pragmatics* 38, 1721-1751.

DREW, P. & E. HOLT (1998), «Figures of speech: figurative expressions and the management of topic-transition», *Language in Society* 27, 495-522.

DREW, P. (2005), «Conversation analysis», in Fitch, K. & Sanders, R. (eds), *Handbook of Language and Social Interaction*, New Jersey, Lawrence Erlbaum, 71–102.

FIRTH, J. R. (1948), «Sounds and prosodies», *Transactions of the Philological Society* 1–26.

FORD, C. E. & E. COUPER-KUHLEN (2004). «Conversation and phonetics: Essential Connections», in Couper-Kuhlen, E. & Ford, C. E. (eds), *Sound Patterns in Interaction*, Amsterdam, John Benjamins, 3–25.

FRENCH, P. & J. LOCAL, (1983), «Turn competitive incomings», *Journal of Pragmatics*, 7, 701–715.

HERITAGE, J. (1984), «A change of state token and aspects of its sequential placement», in Atkinson, M. & Heritage, J. (eds), *Structures of Social Action: Studies in Conversation Analysis*, Cambridge, Cambridge University Press, 299-345.

KELLY,J. & J. LOCAL (1989a), *Doing Phonology*. Manchester, Manchester University Press.

KELLY, J. & LOCAL J. (1989b), «On the use of general phonetic techniques in handling conversational material», in Roger, D. & Bull, P. (eds), *Conversation: An Interdisciplinary Perspective*, Clevedon, Multilingual Matters, 197–212.

KOHLER, K. (1987), «Categorical pitch perception», *Proceedings of the XIth International Congress of Phonetic Sciences,* Tallinn, 331-333.

JEFFERSON, G. (1981), «The abominable «Ne?»», a working paper exploring the phenomenon of post-response pursuit of response. University of Manchester, Department of Sociology, Occasional Paper No.6.

LOCAL, J. (1986). «Patterns and problems in a study of Tyneside intonation», in Johns-Lewis, C. (ed), *Intonation and Discourse*, London, Croom Helm, 157-180.

LOCAL, J. & J. KELLY (1986), «Projection and silences: notes on phonetic and conversational structure», *Human Studies* 9, 185-204.

LOCAL, J., J. KELLY & B. WELLS (1986), «Towards a phonology of conversation: Turn-taking in Tyneside English» *Journal of Linguistics* 22, 411–437.

LOCAL, J. & G. WALKER (2005), «Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech», *Phonetica* 62, 120–130.

OGDEN, R. (2001), «Turn transition, creak and glottal stop in Finnish talk-in-interaction», *Journal of the International Phonetic Association* 31, 139–152.

OGDEN, R. (2006), «Phonetics and social action in agreements and disagreements», *Journal of Pragmatics* 38, 1752–1775.

SACKS, H., E.A. SCHEGLOFF & G. JEFFERSON (1974), «A simplest systematics for the organization of turn-taking for conversation», *Language* 50(4), 696-735.

SCHEGLOFF, E. A. (2007), *Sequence Organization*, Cambridge, cambridge University Press.

SELTING, M. (1996), «Prosody as an activity-type distinctive signalling cue in conversation. The case of so-called «astonished questions» in repair initiation» in Couper-Kuhlen, E. & Selting, M. (eds), *Prosody in Conversation*, Cambridge, Cambridge University Press, 231–270.