PHONETIC DETAIL AND THE ORGANISATION OF TALK-IN-INTERACTION

John Local

University of York, UK
lang4@york.ac.uk

ABSTRACT

This paper examines some methodological and empirical issues concerning phonetic detail and phonetic variability and the work they accomplish in everyday talk-in-interaction. By considering the phonetic and sequential design of a variety of conversational practices I show that phonetic aspects of language should in the first instance be understood as shaped by interactional considerations. I argue that in order to provide a robust account for the organisation and functioning of phonetic detail in everyday conversation we need to:

- enrich our understanding of 'context' and 'communicative function';
- develop a theory of phonetic exponency which derives from a sequential, action-based analysis of talk-in-interaction, and
- treat all phonetic resources equally and not give analytic privilege to one kind of phonetic parameter over another.

If we adopt this approach, it becomes possible to document systematically the ways in which speakers and listeners use fine phonetic detail and phonetic variability in producing and interpreting the moment-to-moment flow of everyday talk.

Keywords: Phonetic detail, Conversation Analysis, talk-in-interaction.

1. INTRODUCTION

One of the 'grand challenges' we face as phoneticians is how to make sense of the phonetic detail and phonetic variation we observe in everyday talk. Almost everything we do that concerns other people involves talk-in-interaction, yet the interactional organisation and functioning of phonetic events in everyday talk has received remarkably little attention in, and has had little analytic impact on, mainstream approaches to linguistics, phonetics, psychology and cognitive science. In this paper I suggest that we should urgently redress this balance and take talk-in-interaction as the core domain of phonetics and the primary data of linguistic interest.

I argue that the phonetic organisation of talk is most appropriately analysed as a *resource* which conversationalists use both to achieve things in the world and to manage the flow of talk. I propose that 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 — specifically by the organisation of utterances into sequences of turns. In what follows I want to engage in a little consciousness raising about what might constitute 'phonetic detail' and 'meaning' in speech by looking at talk-in-interaction. I suggest that it is timely to rethink what is meant by phonetic detail and 'meaning'/'communicative function' The approach I explore here builds on work that colleagues and I have been engaged in over the past 20 years [2], [3], [5], [9], [10], [15], [20], [22], [23], [25], [26], [31], [33].

2. ANALYTIC STANCE

If we are to investigate the ways in which speakers and listeners use phonetic detail and phonetic variability to shape and interpret their contributions to naturally occurring talk-in-interaction, we need to have ways of exploring how such talk is organised. The rigorously empirical analytic techniques of Conversation Analysis (CA) [7], [29], [12] combined with parametric phonetic analysis [14] provide a robust methodological framework to pursue this research. 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 is enables us to establish structural 'sameness', and to compare 'like with like' both phonetically and interactionally.

3. SYSTEMATIC PHONETIC DETAIL AND TALK-IN-INTERACTION

The four phenomena considered in this section are intended to provide evidence that phonetic detail is shaped by its place in sequence and the interactional

work the talk is engaged in. I use the term 'inter-actional work', or simply 'action' to indicate both the social actions that participants are engaged in — e.g. asking someone for something, answering, agreeing or disagreeing with a co-participant — and matters such as handling turn-transition and entry to and exit from talk, configuring talk as a continuation of some prior, abandoned talk or as a new departure, showing that current talk is correcting some trouble in prior talk. The practices described 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 nonstandard varieties.

3.1.   Turn co-construction

Consider the following stretches of talk produced in the course of everyday conversation. The first occurs in face-to-face interaction, the second during a telephone call: 'once those cameras start flashing particularly with the infants it puts them off'; 'but when we walk out of the class nobody knows what went on'. There is nothing particularly remarkable about these stretches of coherent speech — except that they are jointly accomplished by two speakers rather than one as shown in data fragments (1), and (2). The fragments are drawn from a data-set of some 180 cases which represent all the occurrences of such structures in approximately 18 hours of recordings of talk-in-interaction. (For ease of readability simplified orthographic transcripts are given. Transcription conventions are adapted from those of CA [29].)

(1)   **CMXmas-photo**

```
1  K:  once those cameras start flashing particularly
2      with the infants
3  C:  .hh it puts them off
4  K:  it puts them of[f and i[t it's such a [s h a m e ]
5  C:            [yeh  [yeh uh     [people were]
6      doing that last Wednesday
```

(2)   **Two Girls 5**

```
1  Bee:  .hhh and we nod when he wants us to say yes
2        (h)e[n ] .hhh
3  Ava:      [ye]ah
4  Bee:  we raise our hands when he wants to take a poll
5        [.hh you know
6  Ava:  [mm
7  Bee:  but when we walk out of the class
8  Ava:  nobody knows what went on
```

In data fragments (1) and (2) in the course of a not-yet-completed turn-in-progress, a second speaker produces talk which is syntactically, pragmatically and phonetically fitted to the prior talk. This incoming talk brings the turn-in-progress and the action it implements, begun by the first speaker, to a possible (syntactic, prosodic and pragmatic) completion point. For instance, in fragment (1) K's turn-in-progress is part of a complaint about the behaviour of parents at a school play. Her talk at line 1-2 is syntactically and pragmatically incomplete and projects that there is more talk to come (syntactically only the first part of a two-parted structure has been produced). At line 3 C provides a possible syntactic and pragmatic completion to K's talk which K receipts by redoing the lexis at line 4 and going on to do more on-topic talk. Similarly, in fragment (2) at line 8, Ava provides a possible completion of Bee's not yet complete talk at line 7 timing her start-up very precisely to latch onto the end of Bee's talk.

In producing a possible completion of another's talk, a speaker is confronted with a number of interactional problems. For instance, they must show that their talk is a continuation of the prior talk and not some new development. At the same time they must design it in ways which renders it a possible completion. Examination of large numbers of such co-constructed turns reveals that they are rather delicately designed for the interactional demands of their sequential placement and the work they do. Co-constructions such as those in fragments (1) and (2) display a number of systematically deployed phonetic characteristics:

1. they are timed to follow (rather than occur during) a coparticipant's talk;
2. they are not slower than the talk they complete (or than any redoing of the 'completion' by a coparticipant);
3. their ends do not show the kinds of 'lengthening' or slowing down associated with turn-delimitation and found around turn transition;
4. they are not louder than the talk they complete;
5. they are closely pitch-integrated with prior talk (this is particularly noticeable where a female speaker completes a male's turn or vice versa — speakers locate their talk rather precisely in the appropriate part of their own range to make it fit with the prior);
6. they are designed with overall pitch contours which are the same as those which can constitute complete turns and placed where transition to another speaker is relevant;
7. they end higher in the speaker's range than other designed-to-be and treated-as-complete

turns produced by the that speaker;

8. their pitch span is narrower than the talk they complete.

These 'collaborative completions' demonstrate that participants attend to the moment-by-moment evolution of complexes of parametric phonetic detail and what that detail encodes about other levels of linguistic organisation so that they can locate the precise temporal moment to begin their talk. Participants can entrain the phonetic characteristics of their speech to that which has just been produced by another speaker. Sequences of this kind are not rare or unusual; data from everyday talk makes it abundantly clear that participants systematically produce and attend to all sorts of non-lexical ('subphonemic') phonetic detail in the on-line production and understanding of what is being said, why it is being said and what sort of functions it has. That this is so raises the question of whether such details and variation play a central role in the way in which speech is parsed in real-time into chunks and how categories of various kinds are represented. The orientation of participants to a variety of parametric phonetic details suggests that current models of speech production, perception and understanding which concentrate on lexical distinctiveness under-determine the competencies of participants and the cognitive processes involved. Such phonetic detail is no less 'linguistic' and meaning-bearing than those details of the speech signal that express lexical items.

3.2. Stand-alone 'so': sequential context and phonetic design

One reason for examining the ways in which phonetic details and phonetic variability are linked with different kinds of interactional activities (and their linguistic components) in talk is to develop a richer, more grounded understanding of 'context' and 'function'. It has been long known that the phonetic detail of the same lexeme (e.g. *some, that, to*) can differ when it has different syntactic functions [32]. So too can interactionally different forms. 'So' can occur in a variety of sequential locations in talk — it frequently inhabits turn-initial and turn-final positions — and has a range of functions in the speaking turn [30]. It may also occupy a 'stand-alone' sequential position in which it is disconnected from both the preceding and following talk by silence and may exhibit noticeable variability in its phonetic design. For example:

$$[\text{s}̞ə̃ɤ̃ \quad \text{sɒ}ˑ^{\text{m}}_{̞} \quad \text{sɔ}ː_{̯} \quad \text{s}ː̃ɔ̞̆ʔ \quad \text{s}əu̯̟^{\text{p}}ʔ]$$

(The tokens discussed here are drawn from some 130 instances in the transcribed parts of the CALLHOME American English Speech corpus which consists of dual-channel recordings of 120 unscripted telephone conversations.) As the impressionistic records suggest, the variability encompasses features of vocalic quality, duration, laryngeal behaviour and presence/absence of final supraglottal and glottal closures (the variation also involves pitch contour, pitch relationships with surrounding talk and loudness relationships with surrounding talk). Is this variation simply happenstance, a bit of 'noise' in the signal? If all that we were interested in was modelling lexical contrast, the answer might be 'yes'. The tokens all occur in the same 'context': they are preceded and followed by silence. Thus the phonetic variation cannot be constrained by 'context' taken in any traditional sense. However, as Local and Walker [22] show, if the *sequential-interactional context* is taken into account, it is clear that these variations in the phonetic shape of the *so* represent a particular kind of systematic lawful and meaningful variation. These different phonetic designs of stand-alone 'so' are implicated in the management of turn taking and the sequential closure of topics.

3.2.1. Sequential design

Data fragments (3) and (4) provide examples 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:         [mmmmmmmmmmmm]
14  A:  was just (.) I'm in (0.4) you know where I work
15      it's right down town
```

In data fragment (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 installment 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 (for instance, she does not preface her turn with a lexical 'misplacement' marker' [16] or give it a particular prosodic shape [4]; 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 in such as way as to attempt to curtail A's turn [10]). 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).

Fragment (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:   .hhhhhh
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* (these doubles themselves have a particular phonetic design [6]). 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.

3.2.2. *Phonetic design*

The phonetic design of these 'so' tokens has a 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'* (see also [19]). As a group the holding-'so' tokens:
1. are louder than the accented part of the final foot of the speaker's preceding talk;
2. are higher in pitch than the accented part of the final foot of the speaker's preceding talk;
3. have final glottal closure which may, but need not, be accompanied by oral closure; (final glottal closure is regularly held over the subsequent pause and released at the beginning of the speaker's next talk);
4. if they have final glottal closure they 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).

As a group the trailoff-'so' tokens:
1. are quieter than the accented part of the final foot of the speaker's preceding talk.
2. are lower in pitch than the accented part of the final foot of the speaker's preceding talk.
3. never have final glottal closure, though some have accompanying labial oral closure with voiceless egressive nasal airflow;
4. 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).

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. Thus, unlike the data discussed in Section 3.1 where a 'turn-final' pitch contour was required to accomplish the task of co-completion, 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, phonatory type, audible inbreathing) is not.

3.2.3. A confounding case?

Although there is a systematic correspondence between particular clusters of phonetic events and interactional behaviour there are cases where, following a 'so'-token designed with trailoff phonetics, the 'so'-producer continues speaking, as in fragment (5).

(5) **4686-momCHAm**

```
 1   B:  I think I'll give him a call (0.5) next (0.4) next [coup]
 2   A:                                                    [yeah]
 3   B:  this weekend
 4   A:  yeah just to (0.4) surprise him
 5           (0.5)
 6   B:  yeah
 7   A:  act act better than him by giving him a call you know
 8       what I mean
 9           (0.2)
10   B:  yeah
11           (0.5)
12   B:  yeah
13            (1.0)
14   A:  so
15           (0.5)
16   A:  how's mom and dad
17   B:  doing pretty good
```

Do such cases refute the claim that trail-off phonetics does indeed mark completion and turn-yielding? Careful consideration of their sequential design indicates that they do not. Neither A nor B orient to the 'so' in fragment (5) as anything other than complete. It occurs in the environment of disengagement by both speakers from on-topic talk and is therefore like the trailoff-'so' in fragment (4). When speaker A does produce more talk (line 16) its design provides explicit evidence that his 'so' was designed to to be complete: his next move is to solicit talk on a new topic from his co-participant through a first pair part wh-interrogative. At line 17, B accedes to the initiation of the new topic, and in doing so orients to the topic change as a legitimate next action following the 'so'.

Fragments such as (5) demonstrate the importance of careful close analysis of the interactional structure. Just because a co-participant *could* speak after the production of a trailoff-'so' does not mean that they *have* to, or that it is problematic if they do not. There is no simple relationship between the

two different clusters of phonetic events and whether or not speaker change occurs. This lack of *simple* correspondence is a clear demonstration that detailed phonetic and interactional analysis must be conducted hand-in-hand, and on a case-by-case basis, so that each may inform the other.

The phonetic features of holding and yielding discussed here are not restricted to co-occurrence with the item 'so'. One sequential location where phonetic holding features recurrently occur is where a speaker, whose turn is subject to incursive talk, holds their turn until the incoming speaker has ceased to talk. Likewise, trailoff phonetics can be found where talk, designed to be non-turn-competitive, is produced in overlap with ongoing talk [33]. It also has a more general application in environments where on-topic talk is petering out [20].

The actions which are accomplished through turns at talk are done in sequences of turns. They occupy particular positions within sequences and their sequential position is a crucial determinant of how they are structured (phonetically, syntactically, lexically), understood and dealt with by co-participants in conversation. The sequential organisation of talk is therefore legitimately treated as a syntagmatic part of language structure. If we enrich our concept of linguistic 'context' to include sequential organisation and conduct careful and systematic analysis of that organisation and its phonetic design, we can develop a motivated analysis of the relationships between phonetic detail and communicative function.

3.3. Abrupt-joins

Continuing a turn past a point of possible completion is non-trivial and requires some kind of interactional 'work' [28]. The reason for this is that such recognisable completion points make it relevant for a coparticipant to begin talking. One resource which speakers use to build multi-unit turns involves a complex of recurrent phonetic events which Local and Walker refer to as an 'abrupt-join' [21]. One such case (drawn from a collection of some 150 instances) is exemplified in fragment (6) at lines 6-7 with the locus of the join indicated by ▶.

(6) **Holt.5.88.1.5.20.finger**

```
 1   Rob:  I just feel- (0.4) if they're going to go the way of
 2         the modern schools there's an awful- they're
 3         caught between the two that's their pro[blem
 4   Les:                                          [that's right
 5         (0.3)
 6   Rob:  and they've got to go (.) you know really get their
 7         finger out▶what do you think of Ann Percy
 8         (.)
 9   Les:  .hhhhh well do you know e-I wuh- I have a certain
10        sneaking respect for her
11   Rob:  mm
```

((talk continues with discussion of Ann Percy))

3.3.1. Sequential design

Robbie and Leslie are both supply teachers and the talk prior to this fragment has been about a local school which Robbie's children attend. While currently satisfied with the school, Robbie identifies that it is stuck between 'formal' and 'modern' methods (lines 1-3). The sequence is being brought to a close: Robbie has secured agreement from Leslie ('that's right') and then produces a figurative expression 'get their finger out'. (Figurative expressions are regularly found in sequence closing environments [8]). Robbie's talk at line 6 makes relevant a response from Leslie which will display her appreciation of Robbie's assertion (note also the 'turn-soliciting' 'you know') that the school needs to make improvements in the way it conducts its business. On the basis of prior talk Leslie might be expected to assent to Robbie's assertion. This would provide the basis for a sequence closing. However, Robbie does not stop talking after completing the figurative summary. Instead she proceeds to go straight into a next unit of talk. Note, however, that whatever Robbie is doing here with the production of talk beyond the first part of this turn it is not simply or only an attempt to preempt Lesley from coming in and taking a next turn. The talk that Robbie produces is shaped as an wh-interrogative which is designed to solicit talk from her co-participant ('what do you think of Ann Percy'). Lesley does indeed take the next turn, takes up the new direction initiated by Robbie and responds to her question. Robbie's talk past the first part of her turn in lines 6-7 is designed to manage Lesley taking a *particular* kind of turn — one which was not provided for by the sequential unfolding of the talk — at a *particular* time and place in sequence.

3.3.2. Phonetic design

Abrupt-joins are characterised by a distinct and regular co-occurrence of phonetic parameters encompassing both 'segmental' and 'non-segmental' features. These features are distributed over a highly localised domain, clustering around the edges of the two units concerned:

Tempo characteristics: it is well documented that ends of units of talk typically exhibit effects such as temporal extension of the final vowels and consonants, and general slowness of pace. By contrast, one striking phonetic characteristic of abrupt-joins is a localised speeding-up on the last syllable immediately prior to the possible completion point (e.g. in fragment (6) the word 'out'). The effect of this is to produce a noticeably short syllable and the impression of subsequent talk beginning earlier than might have been expected given the tempo of the talk leading up to the abrupt-join. Across the whole collection of abrupt-joins, vowels in the temporally compressed syllables are on average half of the duration of other final tokens of the same word or phonological structure and shorter even than comparable medial tokens (typically 30% shorter than turn-medial tokens with the same phonological structure).

Pitch characteristics: the talk leading up to the point of possible syntactic and pragmatic completion exhibits features typical of other designed-to-be complete talk. The terminal pitch of the temporally compressed syllable may be as low in the speaker's range as that on other, designed-to-be-complete, final items. The pitch excursions associated with these temporally compressed syllables typically exhibit 'tonal compression' [11], and may also be as great as those that occur with other, designed to be complete, final items. There is an audible step-up in pitch from the last syllable of the first unit to the first stressed syllable of the talk following the join. Over the whole collection, the upstep in pitch from the highest pitch of the last stressed syllable in the pre-join unit to the highest pitch in the first stressed syllable of the immediately following talk averages 7 semitones.

Loudness characteristics: as well as a pitch discontinuity across the join between the two units, there is also loudness discontinuity. There is relative quietness of the final, temporally compressed syllable followed by an increase in loudness from the onset of the interrogatively formatted post-join talk. Across the whole collection of abrupt-joins the percept of loudness disjunction between the first and second unit is all the more noticeable because of the sudden diminuendo on the final syllable of the unit that precedes the join. Impressionistically these final temporally compressed syllables are quieter than other syllables in the immediately preceding talk.

Articulatory characteristics: the phonetic characteristics described to this point can be grossly labelled 'disjunctive' and serve to mark the 'two-unitness' of the pre-join and post-join talk. However, there are a variety of other articulatory characteristics which occur around abrupt-joins, which contribute to their distinctive shape and serve in part to provide a phonetic cohesion between the temporally compressed syllable and subsequent, post-join talk. Speakers can employ a variety of phonetic resources to produce the auditory percept of compression or foreshortening of words, (or sylla-

bles) in ongoing talk. Some of these may sacrifice the lexical integrity and identity of the word or syllable implicated. For example, it is common for self-repair to be initiated by suddenly cutting-off the production of a word-in-progress by effecting some kind of abrupt oral or glottal closure. Such 'cut-offs' effectively prevent the production of subsequent parts of a word leaving it suddenly and audibly incomplete. The final syllables of the pre-join talk are never terminated with such glottal, or indeed any other, type of cut-off. This absence of cut-offs around abrupt-joins preserves the integrity of the talk around the join, preserves the lexical identity of the compressed syllable and serves to differentiate it from other types of premature curtailments of talk, such as initiation of self-repair.

A further characteristic of abrupt-joins is that the compressed syllable and the post-join talk occur in maximally close temporal proximity to each other. This temporal proximity is manifest, in part, in the encroachment of phonetic properties of post-join talk on pre-join talk. One characteristic way in which this happens is through what is conventionally referred to as 'assimilation'. Assimilation is *par excellence* a turn-internal (rather than a turn-edge) phenomenon and its principal interactional function is to project more talk to come. There are a number of instances in the current data set where some of the phonetic features of the beginning of the post-join talk encroach upon the end of the final syllable of the pre-join talk. For example, in fragment (6), the end of Robin's compressed syllable 'out' is produced with complete labial closure [p] which projects the labial place of articulation of the initial consonant in the following word 'what'.

Another kind of phonetic resource manifest at the join of the compressed syllable and post-join units, which also provides for the projection of more talk, is the maintenance of voicing. In all the cases where the compressed final syllable ends with a sonorant (vowel, nasal or lateral sound) and the post-join talk also begins with a sonorant voiced phonation is maintained across the join of the two units.

Taken as a cluster, the phonetic design features of abrupt-joins serve to bring off the turn as two distinct units implementing two distinct actions. By employing the kinds of disjunctive phonetics (pitch, loudness, tempo) participants can display that what they are now doing with their post-join talk is not sequentially connected with what they have previously been saying. The cluster of other phonetic characteristics provides for the projection of more talk to come from the same speaker.

3.4. Turn-beginning 'and-uh(m)'

There are, on occasion, places in talk-in-interaction where an ongoing activity may be interrupted by some other intervening action. Jefferson [13] in an analysis of this phenomenon describes the intervening actions as 'side-sequences'. One resource that participants use to display that their subsequent talk should be treated as returning to a prior course of action that was earlier halted or suspended by an intervening side-sequence, is to begin their turn with 'and' followed by 'uhm' or 'uh' (henceforth 'and-uh(m)') [34], [18]. Fragment 7 provides a representative instance.

3.4.1. Sequential design

(7) **Holt.SO88(II)2.2**

```
 1   Ron:   uhm Leslie the reason why I'm phoning is I'm just
 2          wondering whether you might be a hu- able to
 3          help a family in u-Upper Rayley .hhhh who've
 4          had rather a tragedy (0.2) t.hhh uhm (0.2) their
 5          youngest son was killed on the Upper Rayley
 6          bypas[s
 7   Les:        [some years ago
 8          (0.2)
 9   Ron:   couple of weeks ag[o
10   Les:                     [.t oh couple of weeks ago
11          there's another one yes.hh[hh
12   Ron:                             [.hhh and u-[uh-
13   Les:                                         [how old
14          is he- was he
15          (.)
16   Ron:   i- sorry
17   Les:   .hhhhhhhh how old was h[e
18   Ron:                          [he was just twenty
19   Les:   .m.t ah.h
20   Ron:   .hhh (.)  and um  .hhhh (.) I was talking to the
            [ ?aːnˑdʒːm]
21          mother hh uhm u-hu-i-u-her name is Mrs Sorrin
22          Mrs Noel Sorrin .h[hh
23   Les:                     [yes
24   Ron:   uhm uh- up until this point they'd had no
25          connection with the church at all
```

Ron, who is the local church minister, is in the process of attempting to solicit help from Leslie for a local family whose son has been killed in a road accident. As he is setting out the background to the request for help (lines 1-6), Leslie begins a side-sequence by interpolating questions first about the date of the accident (line 7) and then by inquiring about the age of the youth who was killed (lines 13-14, and then, following a repair initiation from Ron, again at line 17). Ron duly attends to Leslie's questions and in doing so temporarily interrupts the forward movement of his preliminaries-to-request-in-progress. After Leslie's questions and Ron's responses to them, Ron returns to his previous line of

action at line 20. With this 'and-uh(m)' beginning turn he goes on to provide further relevant background detail about the family concerned and the nature of the help that is being sought. In constructing his turn at line 20 as he does, Ron effectively treats the query/repair sequences (7-18) as incidental, background information to the on-going request. That Leslie orients to Ron's talk as a return to prior action can be seen in part from her subsequent contributions For example, at line 23, she produces a minimal (receipt) token and passes up an opportunity to produce a longer, responsive turn thus displaying an orientation that an not-yet-complete extended spate of talk (in this case preliminaries-to-request) is underway by her coparticipant [27].

When 'and-uh(m)' occurs at turn beginnings and implements a return to and continuation of prior action it comes off as a designedly unitary item with a distinctive cluster of phonetic characteristics. These phonetic features display a remarkable lack of variation when compared with, say, those of turn-initial 'and' which is not built with following 'uh(m)'. In fragments such as (8) and (9) we can observe participants using and-beginning turns to build talk incrementally on that of others or on their own immediately prior talk.

(8)   **Szcz Sausages**

```
1   Pa:  and we waited four hours for lunch
2            (0.3)
3   Ro:  ((cough))
4            (0.7)
5   Be:  ((cough))
6            (1.0)
7   Ro:  two hours for the beer
8            (3.0)
9   Pa:  and it was freezing cold to boot (0.3) remember that
10           (0.6)
11  Ro:  and there was a hell of a noise going on because they
12           were doing alterations
13  Be:  and were bang bang banging away
```

(9)   **Holt.5.88.1.5.2**

```
1   Les:  .hhh I know uhm I found they got very high towards
2            the end of last term .h[hhh
3   Rob:                          [ye[s
4   Les:                             [and I put it down
5            to uhm (0.6) .t.hhhhh (0.2) comic relief day
6   Rob:  yes
7   Les:  and Easter
8            (0.3)
9   Rob:  yes
```

3.4.2.   *Phonetic design turn-initial 'and'*

In sequences such as those illustrated in fragments (8) and (9), turn-initial 'and' takes a variety of phonetic shapes. Amongst other things, accentuation

can vary, vowel quality and vowel duration can vary, presence of alveolar plosion is variable and place of articulation of any nasal portion can vary (arising from well-known 'assimilatory' effects). For instance, in (8) we find [ɐm ʔɜ̆n ɜ̆n̩ ə̆m̆], in (9) [ă̰n̩ ʔa̰nd].

3.4.3.   *Phonetic design 'and-uh(m)'*

In contrast to the kinds of variability displayed by and in Fragments (8)-(9), 'and-uh(m)' tokens, such as that in fragment (7) display a stable cluster of phonetic characteristics involving pitch, tempo, laryngeal and articulatory features. Amongst the readily discernible features of 'and-uh(m)' are that:

1. it is typically preceded/followed by either breathiness and/or pause;
2. it is characterized overall by relatively level pitch; there is no striking dynamic on-syllable pitch movement on 'and' or 'uh(m)'; there may be a small step-up or step-down in pitch between the two syllables; (across the whole corpus this step between syllables has a mean value of 2.3 semitones);
3. it is typically placed around half-way up speaker's pitch range;
4. it exhibits perceptually 'equal-equal' rhythm [1] over the two syllables.

The first syllable of 'and-uh(m)' ('and') is characterized by the following features:

1. it exhibits creaky voice and/or complete glottal closure (glottal stop) at the beginning of the vowel;
2. it has a full (not reduced/centralized) vowel in the front, open region;
3. it has a noticeable period of sequenced alveolar nasality (i.e. [n]);
4. it has noticeable, orally released, alveolar plosion at the join between 'and' and 'uh(m)'; the release is effected straight into the following vowel of 'uh(m)' — there is no break between the two syllables; this contrasts with instances of turn-initial 'and' forms (not followed by 'uh(m)') which are produced with a final alveolar plosive before a following vowel: in such cases we often find a glottal stop following the plosive and initiating the vowel-onset of the following word; this does not occur in the 'and uh(m)' cases
5. it exhibits relatively level pitch (mean variation in pitch for 'and' across the core collection is 1.25 semitones).
6. it is typically produced more slowly than surrounding talk; this is manifest in the noticeable duration of the vowels of the first and/or

the second syllable and/or the consonantal portions in the tokens (the average duration of 'and-uh(m)' tokens across the whole corpus is 733ms).

'Uh(m)' may show variability in whether or not the syllable is closed by labiality and nasality ([m]). It is further characterized by the following features:

1. a central to back-of-central open vowel [ə ɜ ʌ]
2. no glottal closure at the vowel onset — (though it may begin with creaky voice as in fragment (7)).
3. relatively level pitch (mean variation in pitch for the 'uh(m)' syllable across the core collection is 1.9 semitones).
4. if there is no final closure with nasality and labiality the vowel not terminated with glottal closure (as may be the case where e.g. [əʔ] appears as an indicator of repair-in-progress).

As with the phonetic design of holding and trail-off 'so' and 'abrupt-joins', there is an intimate relationship between the particular phonetic design of 'and-uh(m)' and its interactional function (return to on-going action). The bundle of phonetic features systematically deployed with 'and-uh(m)' draw on parameters (e.g. pitch, tempo, vocalic and consonantal quality) which are conventionally allocated to distinct, different, independent phonological systems — prosodic, non-segmental and segmental. One of the aims of the analyses presented by here is to challenge such an artificial separation of phonetic parameters by providing a more thorough-going account of the phonetic detail associated with various interactional practices.

4. CONCLUSION

One central outcome of research in Conversation Analysis has been the finding that no order of detail can be dismissed, *a priori*, as disorderly, accidental or irrelevant [12]. Analyses of data from naturally occurring talk-in-interaction indicate that this is particularly true in the phonetic domain. The phonetic design of talk (particularly patterns of variation) is one of the *orderly* details of interaction. It provides a resource which speakers use to accomplish social action, shape meaning, and guide its interpretation.

The analyses that I have sketched here seek to develop a more refined concept of phonetic exponency by explicitly grounding accounts of phonetic form in their sequential 'context' in everyday talk. One outcome of this may be a new understanding of 'function' in phonetic and phonological analysis. This might involve the reconfiguration of phonetic and phonological systems and structures in order that they represent the kinds of actions and

practices which I have shown here to have communicative function. It is clear that such a phonology would require a broader theory of meaning than is available through lexical distinctiveness and propositional meaning [24]. It will certainly need to be sensitive to the polysystemic and multi-structural organisation of talk exemplified here by different interactional sequences [17].

By grounding analyses of 'context' in the behaviours of participants we can begin to elaborate accounts of phonetic organisation which are relevant for the participants themselves. The analyses presented are intended to show that we *can* make serious theoretical claims about function (and its phonetic exponents) in everyday talk-in-interaction and can elaborate a theory of meaning without relying on intuitions either as individual analysts or the pooled intuitions of experimental participants. Moreover it provides a systematic way of investigating the phonetic design of non-lexical meaning without analytically problematic appeals to 'speech style' or 'speaker attitude'. Much contemporary phonological work trades on assumptions about lexical meaning, syntactic structure and intuitively ascribed pragmatic functions with little or no attention to their locus of primary occurrence — in stretches of real-time talk-in-interaction. If phonology is to be truly concerned with function and linguistic contrast we need to induce those functions and domains of contrast from a thorough-going analysis of talk-in-interaction.

By extending our formal understanding of 'context' to include the precise sequential organisation of talk, we can also begin to separate out those combination of phonetic parameters which are recruited for particular interactional functions from generic phonetic resources which are deployed for different functions in different sequential contexts. By examining interactional tasks in comparable sequences of talk we can also gain a better understanding of phonetic regularities and phonetic resources across languages [9]. Certainly if we want to build robust models of speech production, speech understanding and phonological representation we need to entertain richer ideas about 'context' and the ways in which phonetic details may relate to the construction of meaning.

5. REFERENCES

[1] Abercrombie, D. 1965. Syllable quantity and enclitics in English. In: *Studies in Phonetics and Linguistics*. Oxford University Press 26–34.

[2] Auer, P., Couper-Kuhlen, E., Müller, F. 1999. *Language in Time: The Rhythm and Tempo of Spoken Interaction*. Oxford: Oxford University Press.

[3] Couper-Kuhlen, E. 2001. Interactional prosody: High onsets in reason-for-the-call turns. *Language in Society* 30, 29–53.

[4] Couper-Kuhlen, E. 2004. Prosody and sequence organization. In: Couper-Kuhlen, E., Ford, C. E., (eds), *Sound Patterns in Interaction*. John Benjamins 333–376.

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

[6] Curl, T. S., Local, J., Walker, G. 2006. Repetition and the prosody-pragmatics interface. *Journal of Pragamtics* 38(10), 1721–1751.

[7] Drew, P. 2005. Conversation analysis. In: Fitch, K., Sanders, R., (eds), *Handbook of Language and Social Interaction*. New Jersey: Lawrence Erlbaum 71–102.

[8] Drew, P., Holt, E. 1998. Figures of speech: Figurative expressions and the management of topic transition in conversation. *Language in Society* 27, 495–522.

[9] Ford, C. E., Couper-Kuhlen, E. 2004. Conversation and phonetics: Essential connections. In: Couper-Kuhlen, E., Ford, C. E., (eds), *Sound Patterns in Interaction*. Amsterdam: John Benjamins 3–25.

[10] French, P., Local, J. 1983. Turn competitive incomings. *Journal of Pragmatics* 7, 701–715.

[11] Grabe, E. 1998. Pitch accent realisation in English and German. *Journal of Phonetics* 26, 129–144.

[12] Heritage, J. 1989. Current developments in conversation analysis. In: Roger, D., Bull, P., (eds), *Conversation: An Interdisciplinary Perspective*. Clevedon: Multilingual Matters 21–47.

[13] Jefferson, G. 1972. Side sequences. In: Sudnow, D., (ed) *Studies in Social Interaction*. New York: Free Press 294–338.

[14] Kelly, J., Local, J. 1989. *Doing Phonology*. Manchester: Manchester University Press.

[15] Kelly, J., Local, J. 1989. 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.

[16] Levinson, S. C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

[17] Local, J. 2003. Variable domains and variable relevance: Interpreting phonetic exponents. 31, 321–339.

[18] Local, J. 2004. Getting back to prior talk: *and-uh(m)* as a back-connecting device. In: Couper-Kuhlen, E., Ford, C. E., (eds), *Sound Patterns in Interaction*. Amsterdam: John Benjamins 377–400.

[19] Local, J., Kelly, J. 1986. Projection and 'silences': Notes on phonetic and conversational structure. *Human Studies* 9, 185–204.

[20] Local, J., Kelly, J., Wells, B. 1986. Towards a phonology of conversation: Turn-taking in Tyneside English. *Journal of Linguistics* 22, 411–437.

[21] Local, J., Walker, G. 2004. Abrupt-joins as a resource for the production of multi-unit, multi-action turns. *Journal of Pragmatics* 36(8), 1375–1403.

[22] Local, J., Walker, G. 2005. Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. *Phonetica* 62, 120–130.

[23] Ogden, R. 2001. Turn transition, creak and glottal stop in Finnish talk-in-interaction. *Journal of the International Phonetic Association* 31(1), 139–152.

[24] Ogden, R. 2006. Phonetics and social action in agreements and disagreements. *Journal of Pragmatics* 38, 1752–1775.

[25] Ogden, R., Routarinne, S. 2005. The communicative functions of final rises in Finnish intonation. *Phonetica* 62, 160–175.

[26] Plug, L. 2005. From words to actions: the phonetics of *eigenlijk* in two communicative contexts. *Phonetica* 62, 131–145.

[27] Schegloff, E. A. 1982. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In: Tannen, D., (ed), *Georgetown University Round Table on Linguistics 1981. Analysing Discourse: Text and Talk*. Washington: Georgetown University Press 71–93.

[28] Schegloff, E. A. 1987. Analyzing single episodes of interaction: An exercise in Conversation Analysis. *Social Psychology Quarterly* 50(2), 101–114.

[29] Schegloff, E. A. 2007. *Sequence Organization*. Cambridge University Press.

[30] Schiffrin, D. 1987. *Discourse Markers*. Cambridge University Press.

[31] 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*. 231–270.

[32] Sweet, H. 1910. *The sounds of English*. Clarendon Press.

[33] Wells, B., Macfarlane, S. 1998. Prosody as an interactional resource: Turn-projection and overlap. *Language and Speech* 41(3–4), 265–294.

[34] Wright, M. 2001. Conversational phonetics: The case of 'and' and 'but'. Master's thesis Department of Language and Linguistic Science, University of York.

6. Acknowledgements