

DETAILS AND CONTEXTS: COMMENTS ON THE PAPERS

Richard Ogden

Department of Language & Linguistic Science, University of York, YORK YO10 5DD, England
raol@york.ac.uk

ABSTRACT

In this paper I make two main points: (1) we need a better understanding of context, (2) there may be naturally-occurring phenomena in conversational data which offer a good basis to see the interplay of segmental and prosodic factors in constructing meaning; it may be possible to use such data as the basis for further work. I aim to open a dialogue between quantitative and qualitative approaches to the study of language.

Keywords: Question, context, conversation, meaning, prosody.

1. INTRODUCTION

The papers in this section nicely illustrate the main themes of S2S; I aim to push the discussions on by referring to my own interest: conversational data, which is the best source of natural speech. A pessimistic view is that conversation provides intractable data because it is so complex. An optimistic view is that conversation is a rich source of phonetic detail, and the medium in which people really do have to turn sound into sense [1, 4, 5]. But it is also, of course, a highly problematic area from the point of view of traditional phonetics and its methodologies. One of the challenges that those of us working on conversational data have to address is: how can we use experimental methodologies to make potentially more robust claims? Or, to ask the question the other way round: how can those using quantitative methodologies use the linguistically relevant categories of conversation to reduce some of the ‘noise’ in the data?

Meaning and the interrelatedness of parameters are two themes which help us address these questions. S2S is interested in what role phonetic detail has in producing meaning, taken broadly: we assume that the listeners’ task is to find meaning in what they hear. As Post, D’Imperio & Gussenhoven, Volín & Studenovskiy, Garcia Lecumberri & Cooke in particular show, speech contains enormously complex information about context, and (as Volín and Studenovskiy put it)

“each context is multidimensional”. This paper takes up the themes of context and the interrelatedness of phonetic parameters.

2. CONTEXT AND MEANING: WHAT IS A QUESTION?

There are at least two problems in looking at phonetic detail in conversational data: (1) Are any two instances of a phenomenon comparable? (2) What kinds of mapping are there between what someone *does* with their talk, and *how they do it* in their talk.

A good example of these comes at the start of Post et al’s paper: “a bigger pitch excursion on a rise can sound simultaneously more questioning and more polite, where the rise is discretely different from a fall signalling assertion and the range varies gradiently with a less polite realisation”. This comment raises lots more questions: *to whom* does it sound more questioning? is “more polite” a socially/culturally determined category, an analytic one, or one that is real in the here and now of those speakers? What do we mean by *context*? We could consider the organization of utterances relative to one another, or social-indexical terms relating to aspects of the speakers’ identities, etc., as well as the more standard notions of context, such as prosodic or segmental context of the kind controlled in other papers in this section.

The apparently obvious category of “question” highlights the problems of working with conversational data. I aim to show that such categories might not actually be so obvious, and, for the purposes of linguistic analysis, not as useful as we thought, because they are too coarse.

Linguists often talk as if “statements” and “questions” were fundamentals of a theory of meaning. In a sense, they are: all languages, as far as we know, have some device for a speaker to produce a turn that makes relevant another turn by another speaker which addresses some issue in the first turn. Question-answer pairs are a very obvious case of this:

R: is Gina there?
M: no Gina's not here.

Here, R produces an interrogative sentence, stops speaking, and then M produces a reply: we know that this second turn is designed as a reply because it starts with a response token, *no*; refers to the same person as the first turn; and like the first turn, the second turn is concerned with Gina's location. But defining questions is not trivial: there is no straightforward mapping between form and function. Interrogative forms can serve to do other things than elicit information: e.g. they can perform a request (*can you pass me the gravy?*), or make an assessment (*isn't that good news?*). In each case, the response takes a different form: for the first one, the answer might not be just "yes", but more probably a physical action, passing the gravy. For the second one, a possible response is another assessment. While 'question' and 'answer' are related, the appropriate form of an answer depends what the speaker is doing by asking the

1	M:	[Ye:h- Hello?
2	R:	Hello uh: Gina?
3		(0.2)
4	R:	Is Gina there?
5		(.)
6	M:	No Gina:'s not he:re.
7		(0.3)
8	M:	Can I take a message?
9		(0.2)
10	R:	Ye:ah Uh:: (0.7) has she mo:ved.
11		(.)
12	M:	No:.
13	R:	Oh: (O[)
14	M:	[`hhh She's working today: Right now.
15	R:	Okay. `hh Ah:: tell her Ro::n ca:lled, ah:: this this is
16		Who is th[is.
17	M:	[whh This is: her mo:m.
18	?:	=[nye:
19	R:	=[Yes:. I-I for I forget your name.
20	M:	Hm:?
21	R:	I forget your name. This is Ron F***** d'you remember me?
22		(0.3)
23	M:	`hh ↑Oh fer heaven sa:ke. Ro:n. Yeh this is Marsha. `hh
24	R:	Marsha right. °([)°

question, and not just on the form of the question.

The example in the text box shows the opening of a phone call. I will use it to show that conversation is structured into sequences of turns, through which speakers perform social actions. One action may make another action relevant, and provide a context for talk to be understood in (e.g.

an invitation makes relevant an acceptance or refusal, and a next turn will be understood in that context). An understanding of the organisation of sequences of turns, and the actions performed through them, can provide us with a way to decide whether two or more utterances are comparable, and with a clearer understanding of how a turn in conversation functions, and how it is treated (or *oriented to*) by the participants themselves. This method lets us escape what Ladd [2: p.39] calls "the Linguist's theory of meaning": essentially, native-speaker intuitions often shared by linguists.

This fragment contains several possible "questions", which are built syntactically as interrogatives, and are responded to with things we can identify as answers ("||" indicates a change of speaker): 1.4: *is /Gina there? || no Gina's not here*; 1.8: *can I take a /message || yeah*; 1.10: *has she /moved || no*; 1.16: *who /is this || this is her mom*; 1.21: *do you remember /me || oh for heaven's sake, Ron*.

Although English yes-no questions are usually said to be done with final rises, and wh-questions with final falls, 1.10 contains a yes-no question with a final fall. Exceptions to apparent generalisations need an account: after all, neither R nor M treat this intonation contour as ill-formed (e.g. neither one of them initiates repair; on the contrary, M provides a timely reply).

There are a number of other turns here which have final rises — a contour often identified as "questioning". If we take the single-word turns with a rising contour in this data, we find: 1.1: *hel/lo*; 1.2: */Gina*; 1.20: */hm*. Are these "questions"? They are not syntactically interrogatives. Some other candidate questions (notably 1.4 and 1.8) have a final rise too. How are these turns treated? L.1 is M's answer to the summons of the phone ringing. But it is also the first of a pair of greetings. L.20 is treated as a repair initiator: R treats this turn as M making a claim that she did not hear him: he repeats his immediately prior turn, and so treats it as the trouble source. Only 1.2 might be a "question": its action is to attempt to identify the person on the other end of the phone. This turn is met with a 0.2s silence from M, which R handles with the interrogative *is Gina there?* This is not quite an expansion of 1.2: 1.2 identifies the person on the phone as possibly Gina, but 1.4 asks whether Gina *is there*, so R treats the silence as indicating that the person called is not Gina: in other words, the 'pause' is treated as *lack of an answer* which in

turn is understood as a *negative answer*. On these grounds we might argue that the turn /*Gina?* made relevant a confirmation that R is indeed speaking to Gina.

Linguists are interested in *contrast*. So one question to ask is: *what else could have come here?* Here is where “question” as a category is too loose. E.g., in 1.20, other possible repair initiators are *what? huh? sorry? pardon?* These all make a claim not to have heard the immediately prior turn, and a possible response to all of these is to repeat the trouble source. So *huh* is in contrast with a limited set of things, but probably not with interrogatives. Whether *huh* is a question or not is not the issue: what matters is where it stands in sequence, its relation to prior and subsequent turns, and the action it initiates. To understand *why this intonation contour here and now*, we need to think differently: what is the speaker displaying through such turns?

Volín & Studenovský say that contexts are multidimensional: this is so not just at the level of e.g. phonological shape, but also in terms of where they are in sequence, and the actions they initiate or respond to.

Speakers also have knowledge about appropriate use of their language. For instance, notice the following exchange:

- M: can I take a message?
R: yeah. uh (0.7) has she moved.

R’s response to M’s question is linguistically well formed. In response to a yes-no question, *yeah* is a “grammatical” answer, in the sense that a yes-no question makes possible *yes* or *no* as answer. But in response to an offer to take a message, agreement to let the offerer perform the offer is not enough: R also needs to say what the message is that M should take. At this point, R has some ‘unfinished business’, which he might leave unfinished (and risk looking socially incompetent), or finish later.

In fact, R finishes this business in line 15, after some intervening talk. But more importantly, M can recognize the turn in 1.15 as the message to be taken because the question-answer sequence which intervenes as a result of R’s *has she moved* is dealt with: he has closed that particular sequence with his *okay* at 1.15. Despite that sequence being closed, M does not come in: instead, R returns to dealing with the *message*. To understand the turn *tell her Ron called* as a response to *can I take a*

message involves using top-down information. The following structure is common in conversation [3: Chapter 6]:

- A: initiating turn 1
B: initiating turn 2
A: responsive turn 2
B: responsive turn 1

By making reference to this kind of sequential organisation, conversationalists have a frame through which they can understand each other’s talk. A turn of one kind often projects a turn of another kind, and this is frequently the “context” in which turns are understood.

In summary, the notion of “question” is complex because it draws on function and form. It is not easy to see how “question” as a formal category of analysis (and therefore also of experiment) can be readily validated. It is more useful to think of: (1) locations in sequence (e.g. at the start of a new sequence, immediately after a turn that makes the current turn relevant); (2) actions (e.g. making an offer, initiating repair); (3) syntactic design (interrogative vs. declarative; elliptical vs. complete syntax); (4) the design of a turn e.g. so that it is hearable as e.g. continuing a sequence or starting a new one.

Finally, “question” (whatever we mean by it) is *by its very nature* an interactional category. Questions are addressed to co-conversationalists. They are a basic form of initiating action. This very obvious fact affects the shape and design of turns at talk, and imposes a set of social requirements for conversationalists’ behaviour. If we want to talk about ‘meaning’ in conversation, then we must remember that conversation involves more than one party. ‘Meaning’ is not inherent in single turns, but is constructed and negotiated through sequences of talk. This is so because talk is used to do things, and these things generally take some negotiating or require certain social behaviours for their satisfactory completion. Our conception of ‘grammar’, therefore, needs to allow for multi-dimensional meaning which belongs to two or more parties, and to take into account not only form but also sequence and action.

3. INTERRELATEDNESS OF PARAMETERS.

Another recurrent theme in several of the papers in this section is the interrelatedness of phonetic parameters. Interactions between things like

loudness, duration, pitch excursions, voice quality and vowel quality provide a complex area for study. As Post et al. comment, these relations can mark non-lexical aspects of meaning. As Garcia Lecumberri & Cooke show, some knowledge about language is subtle, and not necessarily accessible to second language learners. As Volin and Studenovsky point out, variability is multidimensional.

In designing experiments, we need to work out what things to control. When we look at FPD, this issue is even more critical, because FPD is often subtle and tied to particular contexts. Let us now look at a phenomenon from conversation which illustrates the complex of phonetic material, context and meaning.

In outline, the phenomenon (which is a kind of phonetic construction) has the following form: a word contains a pitch peak (H*), though not necessarily high in the speaker's range, followed by a fall with a rather wide pitch span; the tempo of the talk before the contour is fast; consonantal or vocalic articulations are elongated, and some of the articulations are unexpectedly 'strong': there are often instances of fortition. For instance, in *oh very sensible*, there is complete labiodental closure at the start of *very*; and in *you couldn't have a better marriage*, there is voicing throughout the bilabial closure at the start of *better*. These articulations fit iconically with having 'strong' intonation contours, e.g. with a wide pitch span.

The construction is implicated in the activity of assessing [5] or evaluating. It ought to be amenable to speakers' judgements; and a similar thing might be found in other languages: an open question.

The following examples are a representative sample taken from a single call to radio phone-in programme, with four speakers:

{all the all} \w:onderful Margue/rite
cape gooseberry || \wɔ:nderful
{all it's got the most all} \w:onderful aroma when
you open the- the-uhm \bottle up
cherries and redcurrants mixed together || oh
\v:ery sensible
redcurrants providing \m:asses of pectin, you
couldn't {all have a all} \b:etter marriage
thank you \s:ɔ much

The close relationship between the supralaryngeal and laryngeal articulations makes this construction of interest in studying e.g. the alignment of f0 peaks with supralaryngeal

articulations in a natural setting, and as part of an analysis of the construction as a whole. Any study of the construction needs to look closely at the phonetic details of its production and perception. For example, is it easier or harder for listeners to access words with 'unusual' 'hyperarticulated' sounds like [ɥ]? Is the accurate retrieval of such words facilitated or hindered by the accompanying temporal and intonational characteristics? How does the marking of non-lexical meaning affect the phonetic design of talk? Studying a phenomenon like this is one way to see interactions between parameters more clearly, but in a more natural setting. Once the construction has been properly delimited, we could collect a sample of comparable instances larger than this: with large databases, this might be more achievable.

One way to approach looking at conversational data is to identify other kinds of linguistic construction and see whether there are generalisations to make over them both within and across languages. For instance, at what point in time do listeners recognise this construction for what it is? Can we manipulate examples of naturally-occurring data and test how natural they sound or how well they are perceived by native and non-native speakers? A common and seemingly unremarkable construction in natural talk provides a rich ground on which researchers from different perspectives can start to work together.

4. SUMMARY

S2S promises to look at some complex questions for production, perception, machine recognition and synthesis. The questions are difficult because they are so finely nuanced, and because the data are complicated. In this commentary, I have tried to draw out questions that I hope will help to move the research forwards.

5. REFERENCES

- [1] Couper-Kuhlen, E., Ford C.E. 2004. *Sound patterns in interaction*. Amsterdam: John Benjamins.
- [2] Ladd, D.R. 1996. *Intonational phonology*. Cambridge: CUP.
- [3] Levinson, S.R. 1983. *Pragmatics*. Cambridge: CUP.
- [4] Local, J.K., Walker, G. 2005. Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. *Phonetica* 61, 120-130.
- [5] Ogden, R. 2006. Phonetics and social action in agreements and disagreements. *Journal of Pragmatics* 38, 1752-1775.