This is a repository copy of *The independence of phonetic form and interactional accomplishments*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/92000/

Version: Accepted Version

The independence of phonetic form and interactional accomplishments

Abstract
In this response to Peter Auer's commentary, I revisit the question of
phonetic form and interactional meaning as well as the question of what the
aim of transcription actually is (or should be). What I advocate is a careful
look at the ways in which our analyses link linguistic forms with actions.

Peter Auer's commentary shows that it is possible to take my paper forward in ways I did not foresee, and for that I am grateful. Whilst in the main I agree with most of his points, there are two areas upon which I would like to expand: the question of phonetics, phonology and conversation analysis; and the production and use of transcripts. At first these struck me as two rather unrelated topics, but upon further reflection it becomes clear that one's—my—attitude toward the latter is undoubtedly conditioned by my understanding of and philosophy about the first.

Peter Auer raises the point that I have equated form (or practice) with phonetic realisation, and not with the phonological structure. This is deliberate, and based on my belief that "[t]he need for a clear distinction between functional and phonetic criteria in phonological description is most pressing in so-called 'intonational' phonology." (Local, Wells & Sebba 1985:311). Here I would like to address Auer's points by summarising and extending the core ideas of this paper. In the conclusion to Local, Wells & Sebba's paper, as in the practice of the phonetics of talk-in-interaction as developed at the University of York, conversation analysis is used as a method for formulating the functional categories that a phonology must handle. The careful description of the phonetic design of a particular function then can be mapped onto a phonological unit, if useful. Few have done this, but see Jasperson (2002) for a consideration of closure cut-off in self-repair from both a phonetic and a phonological perspective. What is important to understand here, however, is that the function or action being accomplished by a particular bit of talk is the object of investigation. The phonetic design of that action may, or may not, be constitutive of it. For instance, Ogden & Walker (2013) shows how the phonetic design of turns used as offers is related to issues of sequence management, not to the activity of offering. That is, there isn't a systematic 'phonetics of offering', raising the question of whether we should expect other (inter)actions to have phonetic/phonological signatures.

Whether or not the phonetic design plays a constitutive role in other actions (and certainly I believe that it does) is an important analytical question. But once a phonetic or phonological structure has been identified, for a particular action or function, it cannot be imported to another action or function taking place at a different place in the conversational structure. This is why it is basically incorrect to say that rising pitch means anything. Rising pitch is only rising pitch. It becomes involved in meaning when it occurs at a particular place in structure, in contrast with another phonetic exponent.

Thus, when rising pitch is realised as part of the system of repair initiators involving wh-words produced in single-TCU turns, it forms one part of a contrast (with falling pitch) and the whole unit, taken together, can be described as having a meaning or a function (see Benjamin 2013 for a thorough explication of how the pitch movement on final syllables pinpoints a trouble source and simultaneously diagnoses the trouble-type).   There is no analytical grounding for abstracting from this particular finding to say that all final falling pitch means X and all rising pitch means Y.

Auer doesn't seem to take up the argument against equating meanings with phonetic realisations like rising or falling pitch; it seems almost petty to take issue with his claim that functional glosses such as "rushing" or "hesitating" can be used in the early stages of transcription to stand in for phonetic phenomena that require further analysis or more detailed description. But petty I will be, because I want to make the point that the way we allow ourselves to think about the data at any stage may colour our expectations and analyses.   The functional glosses he provides as examples do seem particularly innocuous; I submit, however, that many CA practitioners use what I consider more pernicious glosses such as 'distractedly' or 'surprised'. As I have already stated, it is likely the case that in our everyday lives we could so describe our actions and those of our conversational partners by either overtly or covertly using such labels; this of course doesn't make them appropriate analytical constructs unless they are shown to be so (see eg., Local & Walker 2008, Wilkinson & Kitzinger 2006). And it is especially important that the phonetic instantiation of an achievement — an action — like 'surprise' as done in a particular sequential environment, as a particular part of an adjacency pair is not exported wholesale to any other environment in which 'surprise' is a handy shorthand label for what is probably a complex interactional achievement.

This is why I have claimed elsewhere (Benjamin & Walker 2013:133), apparently controversially, that a given pitch pattern (namely, a high rise fall)   does not "do surprise". It is indeed the case that Local (1996) and Selting (1996) have demonstrated that 'surprise' may be differentiated from another function or action via a difference in the pitch and loudness in comparison to either the surrounding talk or to the other item(s) in the contrastive set. This simply does not mean, however, that every use of eg., high, wide-range, rising-falling pitch is designed to, or must display, surprise. As Local (1996:206) cautions, we should not assume "a simplistic assignment of meaning to pitch contours independent of the interactional, lexical and grammatical environments in which they occur."   This is all I am arguing for; to some, perhaps it is too much, but I remain to be convinced that there is something worthwhile in the drive to attach a meaning to intonational contours any more than there should be (but interestingly isn't) a drive to attach a meaning to a syllable like /da/ or /mi/.

One point of difference between me and Peter Auer is that he describes transcription as a "method to

produce data from audio (or video) recordings". This suggests that transcripts can be regarded as the data, but surely the first caveat about transcription that any CA practitioner learns is that transcripts are a record of the data, not the data itself. Indeed, it is even argued that the audio or recordings are not the data, but rather the closest facsimile we have (the real 'data' being only that precise moment of interaction itself, insufficiently captured by technology). What Auer must mean by this statement is the problem of producing data that we can work with from the recordings that we have, and the problem is a serious one. As our corpora of spoken data grow, how can an analyst handle (ie., search) the data without a transcript, since listening to or watching a recording in real time takes far longer than scanning a written record. Obviously, transcripts are an invaluable tool, and I don't disagree with that.

On the contrary, I agree fully with Auer's suggestion that a distinction should be made between the purposes transcription serves (and this is also made explicit in the GAT system). He suggests one purpose is the rendering of recordings as manageable/searchable, and another, use in analysis and presentation of a phenomenon. Instead, I would suggest three types of transcript: a reference transcript, a working transcript, and a presentation transcript (see also Kelly & Local 1989, and Curl, Local & Walker 2006). The first, the reference transcript, is motivated by the manageability question. This is the kind of transcript that I would imagine Auer has in mind when he describes the "data" produced from the recordings. It is a transcript that stands in for the data when an analyst is searching for instances of particular phenomena, a transcript that probably is treated as the data until a potential instance of 'the thing' is found, which (presumably) is then found in the audio file and listened to. As noted above, conducting conversation-analytic research using large corpora of spoken data is a growing problem.   Colleagues in fields such as linguistics and psychology are developing an interest in interaction, but may only be convinced by analyses built on collections of hundreds of instances, not dozens. Conversation analysts themselves are becoming more aware of the analytic worth of being able to see distributional patterns alongside the intricacies of single-case analyses (see eg., Benjamin 2013:87ff). In an ideal world a researcher would listen to and transcribe the data her/himself; but few if any of us live and work in such a world. We now have the technology and the ability to access large amounts of data, more than any single researcher could hope to transcribe entirely on his/her own; thus, obviously, reference transcripts are necessary and helpful.

What should be captured in this transcript is indeed a question of granularity, as Auer identifies. And a point that bears emphasising is that many of the ways in which data is transcribed confounds the form of the talk with its function, and obscures rather than enhances analysts' ability to spot potential instances of 'the thing' they are collecting. Both my paper and Auer's response detail some of the

ways in which this still goes on. To me, a fairly bare reference transcript is a constant reminder that there's much more detail back in the real data — the recording — and encourages the researcher to go and listen. The inclusion of notes like "hesitating" or emotive labels ("sounds sad/upset/depressed" — which is it?) is unnecessary, because these are all analytic claims that don't belong in a reference transcript.

An obvious question that this construal of a reference transcript raises, however, is what kind of transcription should we teach? Students in university CA courses and in fact all newcomers to CA are indoctrinated as to the importance of transcription, yet I am not at all sure that a transcript itself is important. I submit instead that it is what the process of transcribing reveals that is important, rather than the product (the transcript). By the process of transcribing, I mean the immersion in the emergent, moment-by-moment unfolding of talk. When we teach transcription, aren't we really trying to get students to attend to the placement of talk by each participant, the timing of the talk, the relationship of this turn to those before and after, the very sounds of a turn at talk as they are produced (as well as what is not produced)? And it's true that we probably can't get students to do this unless we insist on a highly detailed transcript, but then we run the risk of the students — and worse, others who see our work — seeing transcription as an end unto itself: as the data, or even as the analysis.

The working transcript, then, would be created from the fragments of talk collected together because they exemplify some phenomenon of interest. Here, the analyst is free to include as much description as possible using whatever technical language, jargon, or shorthand that is meaningful to him or her. If you are noticing that across the different fragments there is a turn which contains a noise that sounds to you like your first-born child gearing up to cry, make a note of it. If you can't shake the sense that there's some element of avoidance going on, note that too. A working transcript is a personal indulgence. It is where we work through the data and get to grips with just what is going on, how the turns are being designed, what they are displaying and how they are being treated. The working transcript is a tool that can sometimes lead to the eureka moment; they are not for general consumption, as the reference and presentation transcripts are.

The presentation transcript, then, is the summation of all the analyst's efforts. This is the level of transcription described so well by Peter Auer at the end of his article, and I cannot add much to it. The aim of my paper, however, was not to advocate new or different kinds of transcription but to alert us to the ways that the forms of language are so often conflated with the functions that language is helping to accomplish in certain sequential environments.

In summary, I do not think that I am advocating anything too shocking. I want to alert others that we ought to be striving for nothing more (and nothing less) than careful attention to the way we describe the activities in which participants in talk-in-interaction are involved—the labels we use for the forms that instantiate the functions we describe. I am by no means the first to do so; I have already cited Local, Wells & Sebba (1985), just one of the sources of inspiration for an approach which carefully separates functional and phonetic criteria (ie., functions and forms, or actions and practices). The other, of course, is Schegloff (1997:505-6), which provides support for my contention that a rise in pitch deployed for a particular kind of function, in a particular place in sequence, need not be considered as 'the same' rise in pitch that occurs at another place in sequence, when he points out that "... practices, deployed always in some position, can accomplish different actions; and actions can be accomplished through a variety of situated practices [emphasis in original]." I feel fairly justified in aligning myself with his call to "focus attention more broadly on the generic problematicity of the linkage between practices and the actions they effectively produce."

REFERENCES

Benjamin, Trevor M. 2013. Signalling trouble: on the linguistic design of other-initiation of repair in English Conversation. PhD thesis, University of Groningen.

Curl, Traci & John Local & Gareth Walker. (2006). Repetition and the prosody-pragmatics interface. Journal of Pragmatics 38, 1721-1751.

Jasperson, Robert. 2002. Some lingustic aspects of closure cut-off. In Cecilia Ford, Barbara Fox, Sandy Thompson (eds), *The Language of Turn and Sequence,* 257-287, Oxford: Oxford University Press.

Kelly, John & John Local. 1989. Doing Phonology. Manchester: Manchester University Press.

Local, John. 1996. Conversational phonetics: some aspects of news receipts in everyday talk. In Elizabeth Couper-Kuhlen & Margret Selting (eds.) Prosody in Conversation: Interactional Studies, 177–231, Cambridge: Cambridge University Press.

Local, John & Gareth Walker. 2008. Stance and affect in conversation: On the interplay of sequential and phonetic resources. Text and Talk 28(6), 723–747.

Local, Wells, & Sebba, M. 1985. Phonology for conversation: Phonetic aspects of turn-delimitation in London Jamaican. Journal of Pragmatics 9, 309-330.

Ogden, Richard & Traci Walker. 2013. Social and Linguistic Units in the Construction of Activity. In Beatrice Szczepek-Reed & Geoffrey Raymond (eds) Units of talk — Units of action, 277-312, Amsterdam: John Benjamins.

Schegloff, E.A. 1997. Practices and actions: Boundary cases of other-initiated repair. Discourse Processes 23, 499-545.

Selting, Margret. 1996. Prosody as an activity-type distinctive cue in conversation: the case of so-called 'astonished' questions in repair. In Elizabeth Couper-Kuhlen & Margret Selting (eds.) Prosody in Conversation: Interactional Studies, 231-271, Cambridge: Cambridge University

Press.

Wilkinson, Sue & Celia Kitzinger. 2006. Surprise as an interactional achievement: reaction tokens in conversation. Social Psychology Quarterly 69, 150–182.