



This is a repository copy of *Form ≠ Function: The independence of prosody and action*.

White Rose Research Online URL for this paper:

<http://eprints.whiterose.ac.uk/91998/>

Version: Accepted Version

---

**Article:**

Walker, T. (2014) Form ≠ Function: The independence of prosody and action. *Research on Language and Social Interaction*, 47 (1). 1 - 16. ISSN 0835-1813

<https://doi.org/10.1080/08351813.2014.871792>

---

**Reuse**

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Form ≠ Function: The independence of prosody and action

## Abstract

This paper argues for the importance of describing form independently of function, especially for prosodic and phonetic forms. Form and function are often conflated by language-in-interaction researchers when they give descriptive labels to the sound of talk (e.g. "upgraded" pitch, "continuing" intonation), and that tempts researchers to see a given form as having a given function or practice - often one that is influenced by the descriptive label. I argue that we should discipline ourselves to keeping to a purely technical description of any form (practice); that will then make it possible unambiguously to show how that form contributes to a particular function (action), without presuming the relationship to be exclusive.

## 1 Introduction

In this paper I argue for the importance of separating linguistic forms (practices) from interactional functions (ie., actions). The discovery and specification of form-function correlations is the basis of (at least some forms of) linguistic inquiry. What does a particular linguistic structure mean, how does it vary, how does its insertion in a larger structure affect that structure? Take for instance the English morpheme (ie., the form) -ED. When attached to a certain kind of verb, it changes the meaning (ie., the function) of that verb from denoting an action occurring in the present to one that has already been completed: I WALK, I WALK-ED. Note that here I am only using the spelling, not the actual phonetic realization of the morpheme.

Linguists also recognise that languages often recruit one form for more than one function – for example, the morpheme spelled with -s in English. This sound is used to mark agreement with the number of the subject when attached to verbs, and possession or plurality when attached to nouns; eg., THE DOG EAT-S THE CAT-S BISCUIT-S. Thus, having a clear technical description of a form (or structure, or practice) makes it possible to unambiguously show how that form contributes to a particular function (or meaning, or action), because they are logically independent and described with different terminology.

In this paper I equate the conversation analytic term ‘practices’ (for example, a practice like delivering a stretch of talk with rising pitch) with what linguists call form, and the term ‘action’ (for example, requesting) with what linguists call function. So throughout the paper, I will use the pairs form / practice, and function / action, interchangeably, though occasionally I will remind us that, in each pair, the terms come from different disciplines.

Linguists, by strictly separating forms from functions, strengthen the scientific value of their analyses by making clear statements about the many-to-one and one-to-many relationships between them. My argument is that conversation analysts – and all students of language-in-interaction – ought to do the same. Such attention to what may at first glance seem only a terminological detail is important because it allows us to more clearly conceptualise the multiplicity of linguistic forms that comprise any turn at talk alongside the multiplicity of functions that turn is performing.

Much of the analysis that describes the phonetic or prosodic design of talk-in-interaction, however, conflates form and function. This contributes to a difficulty in deciding what is practice and what action, as well as obscuring the fact that for many forms, there is not a 1:1 mapping to function. This is not to say that the use of a particular phonetic form may not be crucial to performing a given function, only that it confounds the analysis to say that the phonetic form ‘means’ or ‘is’ that function.

The paper is organised as follows: in Section 2, I show how form and function are readily separated in studies of grammar. Section 3 turns to how form and function are conflated and the problems this can cause. It’s noted here that this conflation is not

unique to work in CA, but may indeed be unique to work in prosody and phonetics. The following section continues this theme, explaining how and why it is important to maintain the division of form and function found in studies of grammar in studies focussing on prosody and phonetics. The conclusion of the paper recaps the relationships among turns, forms and functions; the value of strictly separating them, and raises some questions about what future studies of the relationships among them might reveal.

## 2 Separating form and function in grammar

When we discuss the grammatical forms (or practices) that are used to instantiate particular functions (or actions), we generally do not conflate the terms. Perhaps one reason for this is that the forms are so obviously disparate: they have different words in different sequential orders within the turns, and we have learned that the location of a turn in a sequence is necessary to understand what it ‘means’ on that occasion. Additionally, we recognise that different forms, even those employed in different sequential positions, can sometimes be used to accomplish the same action. Such is the case for the action of requesting, as discussed in Curl & Drew (2008). Here there is no single grammatical form employed to make a request. Rather, the argument is that various grammatical forms are employed to make requests for objects or actions dependent on the amount of entitlement and/or contingency that the participant chooses to acknowledge or display. The utterance types Curl & Drew describe are formally unrelated: modal verbs (could you bring up a letter); irrealis conditional clauses (I was wondering if uhm it was possible to see him one day next week); simple declaratives (I want you to telephone the doctor); and imperatives (pass me the Wishbone). Yet, importantly, they are all used for a similar function – making a request. The specific form that is used in a particular context is responsive to and constitutive of other

pressures (namely, displaying entitlement to the requested object and awareness of any contingencies surrounding its granting), but they all can be described under the functional umbrella of ‘requesting’.

Crucially, the argument in Curl & Drew (2008) does not preclude the use of any of these forms for other interactional functions. I was wondering if is not claimed to indicate requesting outside a particular sequential location; if it prefaced it would rain today, no one would think to search for a way to link the action of requesting and whatever action I was wondering if it would rain today might be being employed to do. The same for the conditional form – conversation analysts would be unlikely to analyze could you get off my foot as (only) a request.

Curl (2006) is another analysis respecting the separation between grammatical form and function. This work examines the different syntactic forms used to make offers in a corpus of telephone calls, and shows how the distribution of different syntactic constructions is systematically related to the interactional situation and the sequential placement of the offer. The syntactic form is shown to be matched to finer details co-occurring alongside the larger activity of making an offer. For example, offers are done with conditionals (If X, (then) Y) when callers present the offer as the purpose of getting in touch, whilst offers of remedy for problems that emerge (or are treated as emerging) from the previous talk within the call are produced with the DYW syntactic form (Do you want (me to X)). A third ‘form’ of offer occurs when overt problems have been raised during the call to which the offer is responsive; these offers have various syntactic patterns that generally match the syntax used in the talk about the problem, but not the DYW format.

This research shows, then, that speakers have a variety of forms available to perform the function of making an offer. The claim is not that any form can be used at any

point to perform that function; there is still an interplay between the sequential location and the syntactic form, but the place in sequence does not necessarily determine the form used. Curl's example 13 shows how an offer produced near the close of the call, the same location that most of the DYW offers were found, is not produced with that format but with another that is better fitted to the overt problem just stated by Mum.

(1) Holt: X(C):1:2:7:18

1 Mum: oh that's w'I was going to tell you I'm:- (.) I'm not bringing an (.)  
2 I'm not bringing any big j:um- (.) big ca:rdigans  
3 (0.2)  
4 Les: no: y[ou can  
5 Mum: [or a dressing gow:n I- (0.2) I [can ba-  
6 Les: [no-  
7 (.)  
8 Les: you can borrow mi:ne

Despite the fact that they are engaged in a move out of closing, Leslie's choice of syntax reflects the activity she is engaged in (rather than displaying a more mechanistic use of the DYW format). Thus we see that the syntax of offers is sensitive to more subtle influences than 'just' those of sequential placement.

What is of special importance is that there is no attempt to assert that the construction do you want intrinsically performs offering, any more than if X, (then) Y does. Rather, the emphasis is on explaining why the different forms are used in different sequential environments, arguing that the subtle differences in form provide additional nuances to the (meta?) function of offering.

It is not only research in which I have been involved that is of interest in this regard. Fox, Thompson, Ford et al. (2013) provides an engaging and thorough review of the interplay between conversation analysis and linguistics in discovering the distribution of and relationships among forms and functions. Additionally, this focus on the separation of form and function is not unique to linguists employing conversation analytic methods. The work

of Thompson & Mann (1987) in particular explicitly addresses the importance of the logical independence of form and function, and will repay a brief summary.

Whilst not employing a strictly conversation analytic approach, Thompson & Mann (1987) adopt an interactive methodology to the analysis of grammar. This particular work looks at one aspect of the organisation of informal written texts. On the one hand, a text is just a string of clauses<sup>1</sup>; what makes a written text coherent, however, is the way that the clauses are connected and/or combined. Many kinds of clause-combining relationships exist; the focus of the research summarised here is concession.

Concession is a discourse relationship previously described in semantic terms as indicating the ‘surprising’ or ‘incompatible’ nature of one clause given the proposition put forth in another clause. What Thompson & Mann note, however, is that this explanation of concession fails to specify who it is that is surprised, or is supposed to be surprised; additionally, assigning such an absolute value to the relation makes it inseparable from simple contrastive relations (that is, where nothing is being conceded but rather only compared).

Thompson & Mann argue, however, for an interactive view of language incorporating the insight that writers want readers to get a particular understanding from what they write; they are not just manipulating symbols when writing but rather are communicating with purpose. Although formally concession and contrast appear to be signalled in the same way, the context surrounding the clause combinations is the real clue to the relationship. In other words, the overall construction of a sequence can signal a concessive relationship between the clauses comprising it, rather than being subject to the use of a particular connective such as the so-called concessive markers “although” and “but”.

---

<sup>1</sup> A clause is, put simply, a subject and a predicate, where a predicate means a verb and everything else having to do with that verb, including a direct object, prepositional phrases, adverbs, and so on.

In fact, in some instances, a concessive discourse relationship exists even though no connective term is used, as shown in this example from Thompson & Mann 1987:443)<sup>2</sup>:

1. Some of you have occasionally given me receipts for Xeroxing done off-campus.
2. Until now I have never had any trouble getting these reimbursed for you.
3. Now the accounting department is clamping down and enforcing a regulation that they claim has been in effect since July 1976 that all Xeroxing on University accounts must be done through the copy centers on Campus.

Here, concession is ‘signalled’ (or displayed) through the (interactive) construction of the sequence, ~~not by the use of a concessive/contrastive conjunction~~. The author clearly displays that reimbursement will no longer be possible (3), after having acknowledged what recipients already know has been the case in the past (2), and pointing out the change in the situation (clamping down and enforcing a regulation that they claim has been in effect in (3)).

By focusing on the interactive construction of sequences in addition to the use of certain connectors, Thompson & Mann are able to demonstrate the “functional unity of the relation” (Thompson & Mann 1987:443). That is to say, they can demonstrate the mapping of several forms (including a ‘zero’ form) to the single function of concession, by maintaining an awareness of what a ‘function’ is – an action here being ‘done’ by writers so as to influence the behaviour or opinions of the readers. Such an analysis would not have been possible without a strict separation of form and function, allowing for the possibility that the same function could be accomplished by various forms.

The logical independence of form and function does not mean that there is no relationship between form and function, or that languages or communities of speakers of a language do not construct or recognise relations between forms and functions: this is where grammaticalization starts (see eg., Haiman 1983; Hopper & Traugott 1993). Certain

---

<sup>2</sup> I have retained Thompson & Mann (1987:437)’s division into (numerical) units, defined thus: “each unit consists of one clause, except that embedded complement and relative clauses are considered part of the same unit as the clauses with which they are associated.”

linguistic forms, whether morphemes, words, clauses, or even intonation patterns or tone groups, rhythmic patterns, ‘bundles’ of phonetic features, can and do become associated with a particular function, because they are repeatedly deployed in the appropriate sequence – not because the form is biologically best-fitted to do that job.

There are cases in which reasonable labels for forms are almost synonymous with descriptions of the functions they are used to perform, especially when those forms are bundles of phonetic features. Curl (2004, 2005) shows how upgraded/non-upgraded phonetic forms seem to equate with upgraded/non-upgraded functions. The analysis of a collection of other-initiated repairs resolved by self-repetition revealed two separate phonetic patterns. One consisted of repetitions that had an expanded pitch range, and were louder, longer in duration, and had altered vocal tract settings compared to the first saying (the turn treated by the co-participant as a trouble source). These were called the upgraded phonetic pattern. The non-upgraded phonetic pattern, conversely, consisted of repetitions that had a compressed pitch range, and were quieter, shorter, and had similar vocal tract settings to the first saying.

The deployment of these phonetic patterns was found to be differentiated by the type of sequence in which the trouble-source turn was embedded. The upgraded phonetic pattern occurred after trouble-source turns that were fitted; they were appropriately designed to follow the previous turn, continuing the sequence-in-progress, or beginning a new sequence if the prior one had been collaboratively closed. The non-upgraded phonetic pattern was found to occur on repetition repairs produced after trouble-source turns that were disjunct, i.e., not designed as relevant next actions, and that lacked a clear link to the just-prior turn, and failed to display shared understanding.

The descriptive terms for the sequential organization of the trouble-source turns – fitted and disjunct – sit comfortably alongside the phonetic descriptions of upgraded and non-upgraded. I would argue, however, that this is due to a metaphorical understanding of ‘positive’ and ‘negative’ functions or actions. It is easy to accept that a turn that was

fitted when produced, but nonetheless treated as a trouble-source, would be repaired with an upgraded phonetic pattern. It is easy to then say that perhaps this upgraded phonetic pattern is a display of a rejection of the proposed role of trouble-maker, or ‘offender’, whilst the non-upgraded phonetic pattern (being quieter, and shorter) accepts responsibility for the breakdown in intersubjectivity. And it is possible that this is all true – that these are the very functions that these forms accomplish in this position. This is indeed the contention put forward in Curl (2004, 2005).

What is important to remember, however, is that this linkage is the outcome of an analysis, not its starting point. The ‘danger’ of a finding such as this is that it can reinforce a non-analytic approach that equates, *a priori*, talk that is loud or slow or higher in pitch with a particular function, and talk that is quiet or fast or lower in pitch with the opposite function.

It should also be stressed that this research is an analysis only of the phonetic patterns employed on other-initiated repetition repairs. There is no claim in either of those publications that the bundle of features that comprise the upgraded phonetic pattern will always co-occur with, or mean that, a turn containing them is designed as fitted to its place in sequence. Until the research is done, we do not know whether quieter, shorter utterances with compressed pitch ranges are used to downgrade speaker claims in other sequential positions. Anecdotally, intuitively, it is easy to associate this pattern with submission and deference. But just as anecdotally and intuitively, what of the deadly quiet and flatly-intoned voice of a parent chastising an unruly adolescent? Just because a particular intonation contour (or difference in loudness, or quality of voice) is used in the performance of a particular function, we cannot assume that contour will always be used to perform said function. It must be shown to be doing so, just as research might show that a modal verb could be used in an interrogative utterance to do some function other than requesting – namely, complaining, as in the (invented) example could you get off my foot.

### 3 Conflating form and function in prosody and phonetics

The work of linguists doing CA differs remarkably from that of ‘mainstream’ linguists primarily in that the former’s work relies on demonstrating the observable orientation of coparticipants in talk-in-interaction. This contrasts with the use of native speaker intuitions to guide research and test hypotheses. Intuitions about the meanings of particular pitch contours form the basis of work in the field known as Intonational Phonology, as illustrated by this quote from Cruttenden (1997:90):

For instance, if we consider [...] what meanings that tone has when combined with each of the sentence-types [...] we may end up describing a number of local meanings like ‘weighty’, ‘impatient’, ‘dispassionate’, ‘serious’, and ‘powerful’, which are all meanings suggested by a low-fall ...

This quote (and there are many others like it) reflects the drive to attach a single meaning to prosodic (mainly intonation or pitch) patterns; here, Cruttenden is not suggesting that there is a particular intonation contour associated with each of the terms he places in scare quotes, but rather that these meanings are all fine distinctions of some more basic meaning that is attached to the low fall. The drive to associate a given intonation pattern with a (single, basic) meaning is not unusual; arguments for or even assertions of tone-meaning associations abound in both research-based theoretical articles and books as well as texts aimed at teaching English to speakers of other languages (eg., Ladd 1996; Wells 2006). In the case of English-teaching textbooks, the appeal is clear: to attain native fluency, you must know not only the sounds of the words, but the melodies that entire utterances should have. In such texts, however, a great deal of surrounding context (albeit invented) is supplied, and the words themselves also provide much of the ‘sarcastic’ or ‘polite’ meanings under discussion.

To what extent the claims are true about what intonation pattern ‘sounds polite’ or ‘sounds disinterested’ is another story. The works cited offer little more than their author’s native speaker authority. More importantly, this literature also postulates that there are a

limited set of intonation contours with associated meanings in a language. Thus, the meanings associated with intonation contours cannot be as specific as ‘weighty’, ‘impatient’, ‘dispassionate’ or ‘serious’; there simply cannot be a 1:1 pairing of form and function. This is why many of the ‘basic meanings’ proposed for intonation contours are so vague as to be vacuous, if not indistinguishable from one another. Notice that the basic meaning postulated in the quote above must be able to encompass both ‘dispassionate’ and ‘impatient’ on different occasions of use.

Cruttenden is clearly aware of the interaction between the words being used (which he calls the sentence-types) and their phonetic-prosodic realization. Given that, to have any intonation at all, these sentence types must be spoken, we can also presume that they will be spoken to be heard by another person – and if to another person, then they are also being produced in a particular context, in a sequence, requiring a response of some kind. Surely then all of this will contribute to how the sentence/utterance is interpreted. Indeed, the intonational phonology literature often provides little vignettes preceding the description of a contour claimed to express a particularly subtle meaning. The words, the fine details of their production (not only their pitch and the overall melody of a turn, but also their timing and aspects of their articulation), what they come before and what they come after – all of this contributes to how participants interpret what a turn at talk means (or, in CA parlance, ‘does’). To researchers accustomed to the importance that sequential location exerts on the function of a particular item of talk, this should not be surprising. What is apparently more contentious to researchers of all persuasions, however, is to suggest that prosodic phenomena like pitch and duration, patterns of sound that are not equatable with alphabetic symbols, perform Function A in one place in sequence, but Function B in another.

Now, to my knowledge no one has ever posited that the sound (form) [s], as used to mark various functions in English (as discussed in the introduction), has a basic or root meaning that encompasses the concepts of both possession and plurality, and that this is true for all languages, not just English. But this seems to be exactly what

intonationologists are proposing for pitch contours. Even if the “local meaning” of a contour is allowed to vary based on context (as per Cruttenden 1997:90), there is still the assumption of a basic, immutable meaning.

The claim that there are basic meanings for intonation contours (whether biologically-determined or not) seems to run directly counter to the arbitrariness of form-function mappings found in the rest of language. It is probably uncontroversial to say that every language makes use of different prosodic forms (ie., high and low pitch peaks) and employs prosody functionally. It is, however, more controversial to agree with Hirst (2005:33) and others that “although many prosodic functions and prosodic forms seem to be quasi-universal, the mapping between form and function is certainly not universal.” Despite what you may have heard, or read, there is not a universal consensus that rising intonation contours ‘mean’ questioning (see eg., Hirst & DiCristo (1998), cited in Hirst (2005); see also Local, Wells & Sebba 1985; Local & Kelly 1986; Wells & Peppé 1996). Intonation is made up of pitch rises and falls (or lack thereof). These rises and falls can be described as technically as any segmental articulation, and then can be associated with a particular function. This isn’t, however, what generally happens. There is a great deal of conflation of form and function in prosodic labelling (for a non-interactional argument against this conflation, see Hirst 2005) rather than an analysis of the mappings between forms and functions. In the CA and Interactional Linguistics literature, this surfaces most perniciously and pervasively in the labels final and continuing intonation.

It’s well known that researchers from the ‘York School’, of which I’m honoured to count myself one, are and have been opposed to the use of such labels (and their corresponding transcription symbols [.] and [,]) for some time. But what seems to be less well understood is the true or full reason for that opposition. The problem is that the use of such symbols and labels for intonation contours encourages a conflation of the form of a contour with its function. To transcribe, and thus label, a turn at talk as exhibiting continuing intonation is to define that contour by its function. It obscures the linguistic-

phonetic description of the sound – level or slightly rising – and replaces it with the function it [the pitch] is doing at that place in the sequence. What ought to be done instead is to say that slightly rising or level intonation can be/is used for the function of continuing, and it is encouraging to see that (Hepburn & Bolden 2013:62) do in fact say that the [,] transcription symbol “indicates slightly rising intonation, not necessarily a clause boundary and not necessarily marking that the speaker is continuing”. It’s interesting that the authors recognise the necessity of warning against conflating the function of continuing with this kind of pitch contour; it’s clearly relevant to them to mention it. And well it should be, because it’s not clear how readily such a definition will be adopted. It is not difficult to find papers using the terms continuing and final intonation – 50 emerge just from a nonsystematic search of the electronic papers in a personal bibliographic collection (in other words, not from a systematic search of the literature, but merely from a computerized search of a collection of electronic versions of papers).

I contend that the continuing conflation of form and function is anathema to scientific enquiry of the sort that conversation analysts ought to be engaged in. I am arguing not for the proliferation of jargon, but for its elimination. There is a growing (and welcome) concern in the CA literature with providing rigorous, verifiable definitions of the actions and practices under investigation (see the argument in, eg., Shaw & Kitzinger 2012). Maintaining a division between the form of a contour (rising, falling, level) and its function (which cannot be specified a priori, but requires analytic work to demonstrate), we should not – must not – combine the form and the function in a label.

This is not just wrangling over ownership of terminology. The following quote from Robinson & Kevoe-Feldman (2010:fn4) exemplifies how confusing two levels of description can lead to overall confusion – perhaps not on the part of the authors<sup>3</sup>, but on the part of the readers.

---

<sup>3</sup> I in no way mean to single out these particular authors as the source or root of the problem; theirs is but one out of many uses of the terminology that I wish to argue against!

turn-final-level or continuing intonation . . . can project the lack of turn completion, and thus turn continuation

Leaving aside the difficulty of understanding how turn-final-level pitch could be continuing (if the speaker continues, the place in question may be TCU-final but presumably not turn-final), the quote is deeply confusing. First, why is it necessary to gloss the description “turn-final-level” with “continuing”? Turn-final-level is a perfectly adequate description of what the intonation, or pitch, sounds like. Turn-final-level describes the form, including a specification of where the form is found. Next, however, the authors describe what this intonation pattern – this form – can project – its function; in this case, turn continuation. But what is gained by incorporating the function into the label for the form? This case seems to rather transparently exemplify the ease with which form and function can be conflated. The function (or action) being described is that of turn continuation past a point of possible syntactic and pragmatic completion. This action may be accomplished (at least in part; but see Local & Walker (2012) for a fuller account of the phonetics of turn projection and turn completion, as well as the discussion in Section 4) by employing the form (or practice) of level pitch. But if this level pitch is (mis)labelled as continuing intonation, it neatly but misleadingly packages the form and function as a single construct.

Making this point often leads to the question, “but then how would you transcribe that?” Excellent discussions of the various uses and pitfalls of transcription, including comparisons between systems, can be found in eg., Kelly & Local (1989); Selting (2010); Walker (2013), and I’d caution against what sometimes seems to be the fetishization of the transcript. It’s widely accepted that creating a transcript involves a certain level of analysis, and over-reliance on the standard Jeffersonian transcription symbols can encourage the conflation of form and function<sup>4</sup> In my experience a fairly ‘bare’ or minimal transcript is

---

<sup>4</sup> The transcription system GAT2 for English (Selting, Auer, Barth-Weingarten et al. 2011) addresses many of these concerns.

an excellent tool for focussing the mind on listening to the data once a phenomenon has been identified based on sequential-interactional function.

## 4 (how to) Separate form and function in prosody and phonetics

It is not my intent here to advocate a modular or hierarchical approach to language structure, in which a syntactic form is given phonetic implementation in a kind of assembly line production. The division between the two is simply an artefact of the way we (have to?) do analysis. Both the syntax and phonetics of an utterance have to be produced in an appropriate way to accomplish an intended action, and a good way to proceed with analysis is to provide accurate descriptions of the various forms that seem to be used for a particular function. These accurate descriptions should be built out of the appropriate terminology, which does not mean that everyone interested in how an utterance sounds needs to complete a doctorate in phonetics (see also Selting 2010, especially pp 24-25).

Rising pitch can simply be called rising pitch, rather than question intonation; level pitch can be called level pitch, rather than continuing intonation. Words that describe forms – rising, falling, quiet, loud, short, long – should be used for descriptions of forms, and words that describe functions – questioning, continuing, offering, requesting – for description of functions, not a mixture of the two.

Once the forms have been described, the next part of the analysis is to discover whether, and show that, the form is consequential for performing a certain function. Local & Walker (2012) does just this for the use of phonetic features in the projection of turns. Through a careful analysis of all the phonetic parameters attending points of possible syntactic and pragmatic completion, they find that a fall-to-low pitch contour at such a point is not a reliable indication of turn finality. They report that in 13 cases with fall-to-low pitch co-occurring with other markers of turn-continuation, only 3 resulted in turn transition. However, 67 cases with fall-to-low pitch co-occurring with other markers of turn-

completion, 60 resulted in turn transition. In other words, fall-to-low pitch is but one of a set of phonetic markers of possible turn finality. When it co-occurs with these other markers, turn-transition does indeed take place; but when it occurs without these other markers, the so-called ‘final intonation’ pattern does not regularly mark turn finality. How much more accurate, then, to dispense with the conflated label ‘final intonation’ and replace it with the more descriptively accurate fall-to-low pitch.

What Local & Walker show, then, is that fall-to-low pitch is systematically but not consequentially associated with turn transition. The aspects of phonetic design that do seem to be consequential for turn transition are aspiration of TCU-final voiceless plosives<sup>5</sup> and outbreaks.

Another function-form pairing that employs aspects of phonetic production not usually captured in CA transcripts (though explicitly provided for in the ‘fine’ level of GAT 2) is that of ‘doubles at closing’ as described in Curl, Local & Walker (2006). These are self-repetitions employed to close down a sequence of talk, as in the following example (fragment 5 of Curl, Local & Walker 2006:1726). The arrowed lines have been retranscribed according to the GAT 2 transcription system, to show (some of) the relevant phonetic detail.

## (2) 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  
3 don't 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

The repetition of “never mind” in line 9 is constructed as a ‘double’ by virtue of its relationship to the first saying (line 7) in terms of accentual pattern and rhythm, pitch,

---

<sup>5</sup> A phonetic feature not usually captured in any conventional CA transcription system.

duration, and loudness. In other words, not every phrasal repetition will display the same relationship between the first and second saying; what Curl, Local & Walker (2006) show is that the production of a double is an accomplishment that plays a functional role – that of closing down a sequence of talk.

The phonetic forms that are a part of this practice are not unique to closings or endings; in fact, at least one of them is used to indicate continuation of a turn, rather than closing of a sequence. Though it is not immediately apparent from the transcript shown above, Curl, Local & Walker (2006:1744) provide this as part of the description of the phonetic characteristics of a double: “second parts are shorter in duration than first parts”. However, longer durations are one of the indicators of turn finality, as shown by Local & Walker (2012). Therefore, it is clear that simply describing the duration, even if done as a relationship between one realization and another (eg., lengthened or shortened) is vastly preferable to using a (misleading) label that incorporates any mention of an interactional function, since the functions that shorter durations can be put to are in this case nearly opposite.

Separating out the various forms of phonetic production (eg., pitch, voice quality, duration) and analysing each independently also helps to detach the analyst from the data and show, in cold hard fact, what is regularly occurring in utterances that perform a particular function. This works to prevent the analysis being colored by the analyst’s folk knowledge of what it is to ‘sound unfinished’ or to ‘sound angry’. We are all (I presume!) human, and participate in a daily basis in the very same kind of interactions that we also study in a professional capacity. In our everyday interactions, we (also presumably) rely on socialised, shorthand labels for the action sequences we are co-constructing: ‘He sounds angry’; ‘she always does that when she wants me to hurry up’ is the kind of self-talk that we probably wouldn’t want to publish as an analysis. When we get to the office, however, and build a collection of turns that are treated as displaying anger, and begin to break down and investigate what design features they share (ie., the forms that produce the function), we

begin to build a scientific explanation of what it means to ‘sound angry’. If we cannot show that there are features in common, this must be taken to mean that our impression, or our intuition, that there is a ‘sound’ that means ‘angry’ cannot be scientifically sustained.

However much I would argue against them as scientific or analytic constructs, I would be loathe to abandon the folk notions of ‘sounding angry’ etc. Having these folk labels and being able to talk about the forms that talk takes (again in a non-technical sense) seems to me to be an invaluable social-interactional tool. Indeed, Local & Walker (2008) show just how such utterances are in fact deployed and treated in talk-in-interaction.

By conducting our research in this way – strictly separating our folk notions of what something sounds like from what we can scientifically describe – we may also make it easier to conceptualize the possibility, indeed the likelihood, that a particular phonetic form may be used for more than one function. For example, Benjamin & Walker (2013) describes the use of high rise-fall intonation contours in repair sequences. Specifically, the claim is that repairs initiated by other-repetition, with this distinctive intonation pattern, are designed to mark the unacceptability of the trouble source. The unacceptability may be a problem with the veracity of the prior talk given the co-participant’s knowledge; or the co-participant may be questioning the contextual appropriateness of what was said. The contour that is used on these repair initiators, a rise-fall, may also function in other sequential locations to display surprise, but this doesn’t mean that the rise-fall need be described as surprise – just as I wonder if needn’t be (and wouldn’t be) described as ‘requesting grammar’.

To show how to separate form and function in research involving prosody and phonetics, and what it can gain us, the research of Benjamin & Walker on high rise fall repetitions will be used to exemplify (and perhaps demystify) how research in this area proceeds.

The collection upon which Benjamin & Walker (2013) is based was built as part of a larger collection of other-initiated repairs accomplished by other-repetition (see Benjamin to appear). In the course of building the larger collection it became apparent that the

phonetic pattern employed on the other-repetition varied. The questions this raised for us were, how could that variation best be described, and was there a functional difference linked to the different phonetic patterns. Our answer to the second question led to the published paper (and to ongoing work); the procedure we followed to answer the former, and which can be extended to other data sets, follows below.

First, all those repetitions that ‘sound similar’ based on repeated listening are gathered together. There’s no great mystery here; just repeated listening to repeats. The next step, however, does require some analytic thought and decision-making. First, what makes them ‘sound similar’? Either they share certain characteristics in common, or they share certain characteristics relative to the talk they repeat. If it’s the latter, you have to listen carefully to, and perhaps measure, the difference between the loudness, pitch range, tempo, and rhythm of the trouble source turn and the other-repetition. Repeated listening to our collection didn’t suggest this as the commonality, so we had to try to pin down what phonetic characteristics they shared among themselves.

Although the high rise fall pitch contours were certainly salient, many other phonetic regularities could be occurring as well, and might have contributed to the sense of sounding similar. Therefore, we set about measuring duration, loudness, tempo, and rhythm. In other words, at this stage we looked for ways to define and delimit the form we trusted ourselves to have heard. They weren’t yet “high rise fall” repetitions, nor had we yet begun investigating the function they were accomplishing.

The absence of any extended discussion of the duration, loudness, tempo, and rhythm of the repetitions in Benjamin & Walker (2013) indicates our failure to find any regularities. Thus while it might look like we only analyzed the most perceptually salient parameter, we did investigate other aspects of phonetic production as well.

So what then of the form-function relationship? Our collection had already been guided by the function of engendering a repair sequence, accomplished by the production of the form of an other-repetition. Were the finer details of this form accomplishing an

additional function? Our answer is yes – that using the form of a high rise fall repetition claims a problem with the appropriateness of the prior turn. This appropriateness might be the talk's veracity, or moral implications, or contextual fittedness. This function, we claim, can be accomplished by using this form. And because we had a clear description of the form, on its own, it was not necessary to invent a label like 'claiming a problem with veracity/morality/contextual appropriateness contour'.

This division between form and function also allows us to acknowledge a potential relationship between, but also separation of, this contour and the function of displaying surprise. They do sound similar. Put in a forced-choice experimental condition, I could imagine labelling the high rise fall repetitions as 'surprised.' But the context they occur in, and the function these repetitions perform, is demonstrably different, as shown by the data extracts below. In each, the repetition is produced with a high rise fall contour.

Fragment 3 comes from a conversation between two male friends discussing John, a mutual friend neither of them has been in touch with for some time. In lines 3-4, Speaker A proffers some knowledge about where this person might now be.

### (3) CallFriend 4175: Utah

1 B: uh yeah (I) think he's just (0.3) you know a real standup guy  
2 or whatever [he's like- ] really (.) workaholic [and (every)  
3 A: [yeah ] [he got ]  
4 a job in uh Utah right  
5 (0.4)  
6→ B: .t ^Utah  
7 (0.2)  
8 A: I think so  
9 B: .t I can't re[member]  
10 A: [ Ida]ho  
11 A: Utah  
12 (0.3)  
13 A: yeah Utah  
14 (0.9)  
15 A: cause he's [(near) ] Idaho  
16 B: [really ]  
17 (0.6)

(For a more detailed analysis of this fragment, see Benjamin & Walker (2013:120ff).) It is of course possible that Speaker B may feel surprised by Speaker A's suggestion that John is in Utah. There's little in this fragment, however, to indicate that a display of surprise is relevant from Speaker B, or that his turn in line 6 is designed to be heard as a display of surprise. Speaker A has done nothing to work up to "he got a job in Utah right" as a potentially surprising bit of information, or indeed as news of any sort. Quite the contrary; he adds a tag question particle, "right", displaying a lack of certainty. Rather than surprise, Speaker B displays conviction of the falseness of this information in the turns that follow. He never accepts Speaker A's continued assertion of Utah (see lines 16 and 20).

The next fragment comes from a conversation between a father and daughter. Though the other-repetition is produced with a high rise fall contour, the context preceding the repetition is quite different from that of fragment 3, as is the treatment of the repetition itself.

**(4) CallHome 4184:Hendricks**

1 Bill: and guess who called here last night (.) looking  
2 for your address  
3 (0.7)  
4 Jen: who  
5 (0.6)  
6 Bill: Aiden Hendricks @:: @  
7→ Jen: Aiden Hend[ricks  
8 Bill: [@ @ @ @ @ @ @ @ @ @  
9 Jen: [why [hhhhh  
10 Bill: [.hhhhh well  
11 Mommy and I were sea- seated with his mother  
12 I'm sure this is why

In this fragment, Bill builds his turn at line 6 to be heard as news, prefacing it with "guess who" (line 1). Jen plays along, giving him the go-ahead to deliver the news, or in this case punchline, and produces an aligning display of surprise in her turn in line 7 (Wilkinson & Kitzinger 2006). Bill treats her surprise as relevant by continuing to laugh; Jen doesn't pursue any repair work, instead asking "why" in her next turn; it turns out she is looking for some explanation of why someone from her past would suddenly

come looking for her address, a situation especially delicate as she (Jen) is at the time of the call living abroad and engaged to be married.

What the juxtaposition of these fragments of talk clearly shows is that the form of a high rise fall intonation contour can't easily be equated with a single function like displaying surprise. Separating form and function allows us to describe many-to-one relationships that go both ways; one form involved in constituting many functions, and one function that can be accomplished by several forms. This also allows for the flexibility of talk-in-interaction, and participants' ability to co-construct the meaning of a turn by their treatment of it.

The work described in this section and the previous one show that careful attention to the form of an utterance allows us to discern functions with a function. Take for instance Curl & Drew's requests. Several different forms are all shown to be used to do the function of requesting, but each individual form displays the speaker's awareness or estimation of the attendant contingencies as well as his/her entitlement to the requested item. Thus displaying entitlement and contingency might be thought of as sub-functions of the function of requesting. A similar argument could be made for the other-repetitions. They function on one level as repair initiators, but through their phonetic design they indicate what trouble needs addressing in the repair: a sub-function of initiating repair.

## 5 Conclusion: A practice for every action?

In this paper I have argued that describing form independently of function will allow the broadening of our understanding of what the phonetic design might be accomplishing. When it comes to the sound of talk-in-interaction, we are far too quick to attach functional, often emotive labels to practices rather than to provide formal descriptions of them. Many prosodic transcription systems do this as well (see

Hirst 2005), and the most commonly used CA transcription system is unfortunately no exception.

This paper is not intended to be an argument over how to transcribe, but how to conceptualise. Conversation analytic principles support the idea of one turn at talk being able to accomplish many actions, which I have equated here with functions; I have shown that any function may be instantiated by several forms, and vice versa, that a particular form may be involved in signalling more than one function. I would suggest that this rich interrelationship – 1 turn → many functions, 1 function → many forms; 1 form → many functions – underlies another concept basic to CA, that of co-construction of meaning. The multiplicity of the relationships between forms and functions, the layering within any given function, allows participants in a conversation to reach different yet equally sensible interpretations of each individual turn at talk. Should any of these interpretations fail to match what the co-participant understands, the mechanism of repair is available to rescue the sequence.

Questions remain about the specific details of the forms and functions. For instance, will increasingly sophisticated analyses, as well as the sheer increase in the number of functions and actions that are studied, allow us to describe how a form or practice that accomplishes a particular action or function is (however subtly) different from a form/practice that accomplishes a different (however subtly) action/function? If we achieve the right level of granularity, will we find a 1:1 mapping between forms and functions, between practices and actions? I simply don't know, choose to remain agnostic about the issue at this time. But we'll never discover the answers, or even come to a better understanding of whether these are sensible questions to ask, unless we analyse form and function separately.

## References

- Benjamin, Trevor M. To appear. Signaling trouble: On the linguistic design of other-initiation of repair in English conversation. Ph.D. thesis, University of Groningen, Groningen, The Netherlands.
- Benjamin, Trevor M. & Traci Walker. 2013. Managing problems of acceptability through high-rise fall repetitions. *Discourse Processes* 50(2), 107–138.
- Cruttenden, Alan. 1997. Intonation (second edition). Cambridge: Cambridge University Press.
- Curl, Traci S. 2004. ‘Repetition repairs’: The relationship of phonetic structure and sequence organisation. In Elizabeth Couper-Kuhlen & Cecilia E. Ford (eds.) *Sound Patterns in Interaction*, 273–298, Amsterdam: John Benjamins.
- Curl, Traci S. 2005. Practices in other-initiated repair resolution: The phonetic differentiation of repetitions. *Discourse Processes* 39(1), 1–44.
- Curl, Traci S. 2006. Offers of assistance: Constraints on syntactic design. *Journal of Pragmatics* 38, 1257–1280.
- Curl, Traci S. & Paul Drew. 2008. Contingency and action: a comparison of two forms of requesting. *Research on Language and Social Interaction* 41(2), 1–25.
- Curl, Traci S., John Local & Gareth Walker. 2006. Repetition and the prosody-pragmatics interface. *Journal of Pragmatics* 38(10), 1721–1751.
- Fox, Barbara A., Sandra A. Thompson, Cecilia E. Ford & Elizabeth Couper-Kuhlen. 2013. Conversation analysis and linguistics. In Jack Sidnell & Tanya Stivers (eds.) *The Handbook of Conversation Analysis*, chap. 36, 726–741, Oxford: Wiley-Blackwell.
- Haiman, John (ed.) . 1983. Iconicity in syntax, vol. 6 of *Typological Studies in Language*. Amsterdam: John Benjamins.
- Hepburn, Alexa & Galina B. Bolden. 2013. The conversation analytic approach to transcription. In Jack Sidnell & Tanya Stivers (eds.) *The Handbook of Conversation Analysis*, chap. 4, 57–77, Oxford: Wiley-Blackwell.
- Hirst, Daniel J. 2005. Form and function in the representation of speech prosody. *Speech Communication* 46, 334–347.

- Hirst, Daniel J. & A. DiCristo (eds.). 1998. Intonation Systems: A Survey of Twenty Languages. Cambridge: Cambridge University Press.
- Hopper, Paul & Elizabeth Closs Traugott. 1993. Grammaticalization. Cambridge: Cambridge University Press.
- Kelly, John & John Local. 1989. Doing Phonology. Manchester: Manchester University Press.
- Ladd, D. Robert. 1996. Intonational Phonology. Cambridge: Cambridge University Press.
- Local, John & John Kelly. 1986. Projection and ‘silences’: Notes on phonetic and conversational structure. *Human Studies* 9, 185–204.
- 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, John & Gareth Walker. 2012. How phonetic features project more talk. *Journal of the International Phonetic Association* 2(3), 255–280.
- Local, John, Bill Wells & Mark Sebba. 1985. Phonology for conversation: Phonetic aspects of turn-delimitation in London Jamaican. *Journal of Pragmatics* 9, 309–330.
- Robinson, Jeffrey D. & Heidi Kevoe-Feldman. 2010. Using full repeats to initiate repair on others’ questions. *Research on Language and Social Interaction* 43(3), 232–259.
- Selting, Margret. 2010. Prosody in interaction: state of the art. In Dagmar Barth-Weingarten, Elisabeth Reber & Margret Selting (eds.) *Prosody in Interaction*, 3–40. Amsterdam: John Benjamins.
- Selting, Margret, Peter Auer, Dagmar Barth-Weingarten, Jörg Bergmann, Pia Bergmann, Karin Birkner, Elizabeth Couper-Kuhlen, Arnulf Deppermann, Peter Gilles, Susanne Gü nthner, Martin Hartung, Friederike Kern, Christine Mertzlufft, Christian Meyer, Miriam Morek, Frank Oberzaucher, Jörg Peters, Uta Quasthoff, Wilfried Schü tte, Anja Stukenbrock & Susanne Uhmann. 2011. A system for transcribing talk-in-interaction: Gat 2. *Gespraechsforschung – Online-Zeitschrift zur verbalen Interaktion* 12, 1–51.
- Shaw, Rebecca & Celia Kitzinger. 2012. Compliments on a home birth helpline. *Research on Language and Social Interaction* 45(3), 212–244.

- Thompson, Sandra A. & William C. Mann. 1987. A discourse view of concession in written english. In Scott DeLancey & Russell Tomlin (eds.) Proceedings of the Second Annual Pacific Linguistics Conference, University of Oregon.
- Walker, Gareth. 2013. Phonetics and prosody in conversation. In Jack Sidnell & Tanya Stivers (eds.) *The Handbook of Conversation Analysis*, chap. 22, 455–475, Oxford: Wiley-Blackwell.
- Wells, Bill & Sue Peppé. 1996. Ending up in Ulster: Prosody and turn-taking in English dialects. In Elizabeth Couper-Kuhlen & Margret Selting (eds.) *Prosody in Conversation: Interactional Studies*, 101–130, Cambridge: Cambridge University Press.
- Wells, J.C. 2006. *English Intonation: An Introduction*. Cambridge: Cambridge University Press.
- Wilkinson, Sue & Celia Kitzinger. 2006. Surprise as an interactional achievement: reaction tokens in conversation. *Social Psychology Quarterly* 69, 150–182.