InLiSt no. 51

The fuzziness of intonation units:
Some theoretical considerations and a practical solution

Dagmar Barth-Weingarten
(University of Freiburg – Hermann Paul School of Language Sciences)

Abstract

Talk-in-interaction is structured according to various dimensions. One of them is prosody and the concept of 'intonation unit' has been widely used to capture prosodic structuring. When looking at prosodic chunking in detail, however, transcribers and analysts alike face the problem that, in a considerable number of cases, intonation units do not seem to do justice to language reality, or our perception thereof. This paper adopts earlier suggestions that it may be more helpful to analyze the prosodic-phonetic structuring of talk in terms of what separates the 'units', namely the unit boundaries, or 'cesuras' in talk. In addition, it argues that also weak cesuras should be notated to allow us to adequately reflect the complexity of prosodic-phonetic structuring of talk for notating and analytic purposes. To enable this, notation conventions of various degrees of granularity are suggested. All of this can serve as a basis for investigating the interface of prosodic structuring with other language-organizational dimensions, such as syntax, and the accomplishment of actions.

1   Introduction[1]

When listening to spoken language, one gets the impression that longer turns are sub-divided into smaller spurts in many languages of the world. This is observable in terms of syntax, semantics, the accomplishment of actions, but perhaps most obviously so with regard to prosody (see, e.g., Chafe 1994, Ladd 2008: 288, Szczepek Reed 2010). This paper will focus on prosodic "chunking".

   Linguists have, for quite some time, attempted to capture prosodic chunking by notions such as 'tone unit' and 'intonation group' (cf. the British School), 'intonation unit' (cf. the autosegmental approach, for instance) and, more recently, 'turn-constructional phrase' (see Szczepek Reed 2010a). Criteria for identifying such prosodic chunks recurring across approaches are a coherent intonational contour with at least one prosodically prominent syllable, a significant final pitch movement, final lengthening together with final loudness diminuendo and a shorter stretch of non-modal voice quality typically followed by a (micro)pause and anacrustic syllables as well as a pitch step-up/down in the next intonation unit (IU).[2]

   Different approaches to the IU put different amounts of emphasis on the various definitional criteria (see section 2). Yet, they agree in assuming that IUs are the basic way to capture the prosodic chunking of spoken language (the unit approach). Analyzing talk in terms of IUs then requires the analyst to apply these criteria, and there are indeed many instances of clear prosodic chunks. Consider ex. (1) from the "Callhome corpus"[3] as an instance of a succession of clearly delimited IUs. For the auditory impression, the reader is referred to the sample sound files available at http://www.inlist.uni-bayreuth.de/issues/51/index.htm. Each line contains one IU; line numbers refer roughly to seconds in the recording.

(1) No way (Callhome, 6479_318-322)[4]
(AE telephone conversation between two women friends. Barbara is telling Anita about how she eventually decided to take a sabbatical year.)

```
        318     Ani:    [rea[lly   ]
        319     Bar:  °hh[h  [and th]en
        320             i was like
->      321             no way
->      322             this would be really fun
```

In this excerpt, Barbara's turn (l. 319-322) is divided into four clearly recognizable IUs. If we focus on l. 321-322, for instance, *no way* is clearly separated from *this would be really fun* by means of a clear boundary tone (a (rising-)falling-to-mid pitch movement), lengthening and loudness diminuendo on the last syllable *way*, a pitch up-step, anacrusis and a loudness upsurge at the beginning of l. 322, together with the distribution of focus accents (*way* and *fun*) as well as the high pitch register and stylization embracing *no way*. In cases such as these, the notion of IU is completely unproblematic and seemingly offers support for analysis in terms

---

[1] I gratefully acknowledge the discussions with Sandy Thompson as well as data sessions with Elizabeth Couper-Kuhlen, Richard Ogden, Jack du Bois and numerous other colleagues during the past five years, amended by ideas that came up during the workshop "Grammar and interaction revisited" in Helsinki, 10-12 March 2011. All of these contributed to developing the ideas presented in this paper. Thank is also due to Elizabeth Couper-Kuhlen, Margret Selting and Peter Auer for commenting on earlier versions of this paper. All remaining errors are, of course, my own.

[2] Thus, prosodic chunks are not delimited by intonation, i.e. pitch, alone, but by a variety of prosodic parameters. They should therefore rather be referred to as *prosodic* units (cf. Szczepek Reed 2010). However, the term 'IU' is used in this paper to refer to prosodic phrases in general because it is an established term with an established abbreviation. Most of all, it is to be noted that this choice of terminology does not favor one theoretical approach over others.

[3] Vgl. http://www.ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC97S42, last access: 23 May 2011.

[4] For Transcription conventions, see the Appendix.

of syntax and the accomplishment of actions: With the (reported) interjection, Barbara contextualizes a negative stance against her own earlier objections, which is supported by an accounting clause.

At the same time, when trying to employ IUs exhaustively across larger amounts of material, e.g. for transcriptional purposes, one soon runs into difficulties. In particular in natural talk-in-interaction which was consequential for the participants at the time it was recorded, there are a considerable number of cases in which two phenomena pose problems: For one, it is sometimes not so clear *where* an IU ends. Consider seconds 566-570 of ex. (2), for instance.

(2) Went there (Callhome, 5046_563-570)
(AE telephone conversation between two women friends. Donna is telling Sue how mutual friends found some spot worth buying as land property.)

```
      563   DON:    anyway
      564           so that's how that spot
      565a  Sue:    [°hhh]
      565b  Don:    [°hhh] came into the picture
->    566-570       an' they finally went there an' stayed there an' had a
                    wonderful time
```

The last bit of this excerpt (l. 566-570), upon listening, could be notated in various ways. Among them are those shown in (2a) and (2b), each of which assigns the connective *and* to different IUs.

(2a)
```
      566-570    a  an' they finally went there
                 b  an' stayed there
                 c  an' had a wonderful time
```

(2b)
```
      566-570    a  an' they finally went there_an'
                 b  stayed there_an'
                 c  had a wonderful time
```

Notation variant (2a) is supported by the position of the boundary tones on *there* and the anacrusis on *and*. For variant (2b), we can find arguments in terms of the position of a pitch upstep after the *and*s as well as latching/cliticization and non-modal voice quality, which integrate each of the *ands* with the preceding IUs (cf. Auer 2010 for a similar case in German).

Second, in some cases it is not even clear whether there is an IU ending at all. Consider l. 897 of ex. (3) as a case in point:

(3) Sunday evening liturgy (CallHome, 4705_896-905)
(AE telephone conversation between two acquainted nuns. Ann is in the process of suggesting what a travelling missionary priest could do at their convent school.)

```
      896   Ann: °h an' maybe working out
->    897        some kind of a deal whereby he comes an' says (0.29) a
                 sunday
      901        offers a sunday evening
                 liturgy
      902        °hh
      903   Bon: oh
      904   Ann: an' v[isits with the] students before or after.
      905   Bon:      [that's          ]
```

It could either be notated as one long IU, or – on the basis of the distribution of accents, lengthening and rhythm – with IU endings after *comes* and *says* and perhaps even after *deal*.

An apparently easy way out would be to fall back on syntax, as suggested by Cruttenden (1997) and Crystal (1975), for instance: The prosodic boundary is placed where the syntactic boundary is located. However, this would circularize research on the syntax-prosody interface. While some correlation between the two language-organizational levels has been found, it is not there all the time (cf. Cruttenden 1997: 68-69). Similarly problematic are the correlations between IUs and units on other language levels, such as semantics and information structure (cf. Chafe 1994) and action (Szczepek Reed in prep.). Yet, if we restrict ourselves to prosodic, and phonetic, parameters, we run into the difficulties mentioned above. The relevant features are just not there to a sufficient extent or with sufficient overlap.

Several researchers have pointed out these problems (Brown et al. 1980, Cruttenden 1997, du Bois et al. 1992, etc). Yet, so far, few have suggested, let alone implemented, an alternative (see Brown et al. 1980 for a little accepted attempt at pause-defined units). Hence, analysts are still forced to take a yes/no decision on cases such as those provided in (2) and (3). This, however, leaves them with a considerable dilemma: Either way, the decision feels untrue to language reality. No matter whether they decide to treat the instances as one or two IUs, the impression of a "bump" in the prosodic flow is, in comparison to other, clear cases of IU separation, either blown up – in the case of treating them as two IUs – or wiped out – in the case of treating them as one IU. In any case, the decision the unit approach forces them to take makes them "throw away" a considerable amount of information on the realization of IUs and their boundaries (see also Barnwell 2011). This is highly undesirable, in particular in view of the role that is assigned to IUs for quite a number of interactional phenomena including multi-unit turn formation, clause combining, the accomplishment of actions et cetera. Not to speak of the novices to transcription and Interactional Linguistics, or student assistants to interactional-research projects, who are hesitant to move beyond the initial project step of transcribing because they assume that they "just don't hear it right yet". On the other hand, pushing the identification of prosodic units into "the experts' corner" and not trying to tackle it at all (Szczepek Reed in prep.: 20) is equally undesirable. There is research (e.g. Barth-Weingarten/Couper-Kuhlen 2011) that has shown that the way in which, not to say the degree to which, IUs are connected to each other may be as relevant as identifying IUs per se.

What this paper intends to do is propose a way to deal with these problems. It will adopt earlier suggestions to shift the focus from the intonational units to their boundaries (cesura approach) and develop them further to provide for dealing with the clear and the problematic instantiations of IUs alike. For this, it will (1) explicitly acknowledge the phenomenon of fuzzy, or weak, prosodic boundaries, and (2) understand this fuzziness, or weakness, as a gradient phenomenon. To capture the latter, a notation system will be proposed which can be adapted to the level of granularity needed for the specific research question pursued by notating prosodic structuring. Therefore, the approach suggested may be of interest to novices and more advanced students of prosody alike in that it allows us to notate the impression of fuzzy IU boundaries in as fine a detail as is needed or desired.

The remainder of this paper is structured as follows: Section 2 will summarize the unit approach to prosodic structuring and its problems. Section 3 will lay out the alternative cesura approach. Section 4 will suggest how this can be implemented in the transcription process at various levels of granularity.

2   The unit approach to prosodic structuring and its problems

Most previous work on prosodic phrasing, regardless of its exact theoretical provenance (for extensive surveys of previous work see, e.g., Couper-Kuhlen 1986, Tench 1990, Selting 1993, 1995, Cruttenden 1997, Szczepek Reed 2010 and Ladd 2008, among others, also Barth-

Weingarten in prep.), models prosodic structuring in terms of the stream of talk being sub-divided into smaller chunks, units ("unit approach"). For these, they employ various labels, among them 'tone group' (Palmer 1922, Halliday 1967, 1985), 'intonation-group' (Cruttenden 1997), 'tone-unit' (Crystal 1969, 1975, Brazil/Coulthard/Johns 1980, Brazil 1985/1997), 'rhythm unit' (Pike 1945), 'intonation phrase' (Pierrehumbert 1980, Shattuck-Hufnagel/Turk 1996, Selting et al. 2009, 2011) and 'intermediate phrase' (Beckman/Pierrehumbert 1986), 'informational phrases' (Gumperz/Berenz 1993), 'intonation unit' (Chafe 1988, du Bois et al. 1992, Selting 2005) and 'intonation contour' (Selting 1995).

   The identifying features for these are differently weighed, but they include the internal criteria of the occurrence of a prominent syllable (nucleus), possibly accompanied by other, less prominently stressed or unstressed syllables, and external criteria, i.e. cues to unit boundaries, such as pause, tempo changes (anacrusis, latching), final lengthening, change in pitch level and/or pitch direction on unaccented syllables, changes in loudness and voice quality and rhythmic breaks (cf., e.g. Pike 1945, Crystal 1969, Cruttenden 1997, Beckman/Ayers Elam 1997, Schegloff 1987, 1988, 1996a, 1998, du Bois et al. 1992, Gumperz/Berenz 1993, Chafe 1994, Selting 1995, Lerner 1996, Hermes 2006, du Bois 2008).

   Auer (2010) points out the problems involved with this list of criteria employed to segment spoken language:

> "Dabei bleibt unklar, ob eines oder mehrere dieser Kriterien vorliegen müssen. Statt einer Operationalisierung haben wir es also eher mit einer komplexen Heuristik für das Erkennen von IP-Grenzen zu tun, deren Anwendung auf das Datenmaterial erhebliches Geschick erfordert – und manchmal dennoch nicht zu klaren Entscheidungen führt.
> [In doing so, it remains unclear whether one or several of these criteria need to be met. Instead of an operationalization, we are rather dealing with a complex heuristic for identifying intonation-phrase boundaries, whose application to the data requires considerable skill – and even then sometimes does not lead to clear decisions – my transl., DBW]" (2010: 8)

This may explain the considerable number of problems scholars acknowledge in prosodic-unit identification, especially with natural, conversational data (cf. Brown/Currie/Kenworthy 1980: 41, Cruttenden 1997: 29, Brazil 1985/1997: 16, 17, 149, du Bois et al. 1992: 100, 112-113, Wennerstrom 2001: 31, Selting 2005: 36, Birkner 2008, Szczepek Reed 2010: 197, also Breen et al. 2010).

   Various solutions to these problems are suggested, most of which are, however, not acceptable for various reasons:

- retreat to syntactic and/or semantic criteria (Cruttenden 1997: 30, Crystal 1969, 1975: 207, Gumperz/Berenz 1993: 95) may produce potential circularity in studies on the prosody-syntax interface (see, e.g., Ladd 2008 on this issue in the autosegmental-metrical (AM) approach, also du Bois et al. 1992: 101, 106, Couper-Kuhlen/Selting 1996: 15 for conversational data),
- retreat to action identification (e.g., Szczepek Reed in prep.) may become circular in studies on the prosody-action interface (cf. Ford/Fox/Thompson 1996),
- retreat to apparently more objectively identifiable prosodic features, such as pausing (Brown/Currie/Kenworthy 1980: 47, Szczepek Reed 2010) is too restrictive, as pausing may occur also within an IU (cf. Couper-Kuhlen 1986: 75-76).

Similar problems arise with the more holistic concept of turn-constructional units (TCUs) when they are defined on the basis of (a combination of) syntactic, prosodic and action completion (Sacks/Schegloff/Jefferson 1974, Schegloff 1996a, Selting 2000).

   An alternative solution was seen in going beyond the categorial approach. Some scholars suggested a prototype approach to prosodic units (Brown/Currie/Kenworthy 1980, Gum-

perz/Berenz 1993, du Bois et al. 1992, du Bois 2008, Schuetze-Coburn 1992, 1994), which, however, has been rarely adopted in practice. The same applies for attempts to amend the two-category approach. Du Bois et al. (1992), for instance, refer to 'intonation subunits', "something intermediate between a full intonation boundary, and no intonation unit boundary at all" (1992: 112). The intonation subunit boundary "is sometimes seized upon as a compromise, in cases when one transcriber hears an intonation unit boundary, while another hears none" (1992: 68) and is supposed to "represent a juncture which displays some of the features of a prototypical intonation unit boundary but not all" (1992: 68). It has, however, not been applied systematically (cf. Barnwell 2011 for a similar criticism), and even if it were, three, instead of two, categories do not overcome the general problem: the units model as such.

Auer (2010: 10) has argued on the basis of structuralist assumptions on segmenting into units – exhaustivity, atomicity, discreteness and consistency of descriptive level – that the general idea of segmenting talk is in fundamental opposition with the online character of spoken language. A model that assumes clear beginnings and ends is often too restrictive to be able to take the "usual" local contingencies of everyday talk into account (cf. Thompson 2011). While being produced, practices employed in, and projecting, one kind of structure (e.g. Auer 2000, 2005) can be cut-off, repeated, expanded and re-analyzed, or rather re-employed, in(to) (fitting) other structures, because current speakers feel the need to locally manage aspects of turn-holding, recipient-design, preference and stance-taking. All of this happens in real time, so that previous utterance parts cannot be adjusted, let alone be erased. Similarly, participants, unlike analysts (Selting 2000: 491), usually do not (have the time to) categorize them retrospectively. Moreover, management of interaction can be accomplished on a range of dimensions, including syntax, prosody, semantics-pragmatics and visual behavior. Cues from these dimensions may, or may not, co-occur (cf. Ford 2004: 31). Problems such as these may be even more common with prosodic structuring because of the range of cues that lead to the impression of a boundary, such as pausing, pitch step-ups and anacrusis, which may, or may not, co-occur (e.g., Du Bois et al. 1992: 100, see also section 3).

Moreover, the unit approach to prosodic structuring also brings with it a number of unwanted associations that are, at least in part, responsible for the problems encountered: A units conceptualization involves the general problem of hypostatizing the process of talk.

1. Treating prosodic units as static entities suggests that a categorial treatment is possible. And indeed, this is how all previous approaches treat the IUs: either a chunk is, or is not, an 'IU' on the basis of the criteria stated. The chunks themselves are considered an unquestionable fact, despite difficulties encountered in determining them in spontaneous conversational data exhaustively, where some chunks are "not so apparent" (Cruttenden 1997: 29, cf. also Auer 2010). Note that Cruttenden's choice of wording already suggests fuzziness of the concept. Only few authors have attempted to go beyond categoriality.

2. Defining the units of talk includes a definition of the number of types of chunks (see, for instance the autosegmental-metrical approach). Yet, the number of differentiable chunk types may depend on the complexity of the utterance produced. An assessment can be responded to with just a short second assessment produced in one prosodic unit, while a story may require more intricate, not to say nested, chunking. There is evidence that this is reflected in prosodic marking (cf. Barth-Weingarten 2009). An *a priori* fixation of the number of unit types may underestimate just how much language can adapt to its local employment, and thus just how much variation is going on.

3. The chunking of talk into (neatly separable) units (in a row), in effect, also encourages the perspective of parallelism between the prosodic-phonetic language organizational level and other (unit-based) dimensions, such as syntax (Schegloff 1996a, Selting 2000), information structure (see, e.g., Chafe 1994, also Halford 1994, 1996, Mukherjee 2001) and action (e.g., Selting 2005, Szczepek Reed 2010). This, in effect,

discourages observing and describing the prosody and phonetics of talk-in-interaction in its own right (see also Knowles 1991: 160 for a similar criticism).

Hence, the unit approach to capturing the prosodic structuring of the flow of talk in interaction involves a number of difficulties (see also Ford 2004, Szczepek Reed 2010: 205).[5] As a consequence, Auer (2010: 10) even questions the reasonableness of *segment*ing talk altogether and suggests instead that we focus on what comes between the chunks, the boundaries, because that is what participants need to monitor in talk. This approach is adopted here.

3   The cesura approach

3.1   Cesuring in talk

Auer (2010: 11-12) argues that participants, in natural interaction, are forced to constantly monitor talk for "Gestaltschlüsse" (*Gestalt* endings), which may be relevant for turn-taking, and they have to do so on-line. These *Gestalt* endings occur where the projections on the syntactic, prosodic and semantic-pragmatic dimensions are satisfied (cf. Auer 2000, 2005, cf. also completion points (CPs) in Ford/Thompson 1996). In this regard Auer introduces the term "zäsurieren" to refer to what it is the *Gestalt* endings are doing (2010: 12, see also Peters/Kohler/Wesener 2005). They divide up talk, without, however, focusing on the units between the *Gestalt* endings. They draw attention to boundaries, rather than units.

For want of a similar term in English – compare the terms 'chunking', 'phrasing', 'segmenting', which all by their very stem associate the existence of 'chunks', 'phrases', 'segment', i.e. units – I will employ the words *cesuring* and *cesura* for this concept, adopting their meaning of 'cutting'. Note the similarity to the uses of this word as a term in poetics and musical notation (alternative spellings are *cæsura* and *caesura*), where they refer to a stop in a line of poetry and a cessation of musical time respectively.[6] Similarly, prosodic cesuras "cut into" a participant's utterance flow.

With his shift of focus to the boundaries, rather than the units, Auer also explains the fuzziness of segments: Most salient *Gestalt* endings, i.e. CPs, are those, Auer (2010: 12) claims, at which the projections on all three dimensions (syntax, prosody and semantics-pragmatics) are satisfied (cf. complex CPs in Ford/Thompson 1996), while less salient endings result from incompletion on one or more dimension(s), or from the cues to completion being ambiguous on one or more dimension(s). Note that this modeling of fuzziness includes two ideas: First, the overlay of CPs on the various dimensions. For this, Auer draws upon the concept of projection. This concept crucially depends on the existence of (more or less well-entrenched) patterns (*Gestalts*) on the relevant dimensions. Uttering the beginning of such a pattern allows the participants to infer what, roughly, it will take for it to be completed.[7] In his sample transcript, Auer seems to employ this idea mainly for identifying CPs on the syntactic and semantic-pragmatic dimensions. With these he distinguishes between syntactic *Zäsuren* ('cesuras') and syntactic segment boundaries, which seems to refer to the distinction between possible CPs and actual syntactic CPs/turn endings respectively. Prosodic completion, apparently, collaborates with (possible) syntactic completion to differentiate the two.

It is in discussing prosodic completion that Auer employs the second idea of explaining fuzzy boundaries: ambiguous completion cues, i.e. the cases this current paper is concerned

---

[5] Interestingly, similar problems have been encountered when employing the unit approach at other levels of language organization (see, for instance, Schegloff 1996b and Ford/Fox/Thompson 1996 on TCUs, also Thompson 2011 on syntactic units).

[6] I am grateful to Mack Thompson for drawing my attention to this.

[7] Sacks/Schegloff/Jefferson (1974: 702) have, in fact, connected such an idea closely to the unit concept in terms of the TCU.

with. Auer observes that with fuzzy prosodic boundaries some cues for completion are there, while others are missing or cue non-completion (2010: 8-9, 13-15). Nevertheless, at these points he still takes a binary decision for, or against, a prosodic boundary, indicated by placing, or not placing, the symbol for a prosodic cesura in the transcript. This suggests that we are still obliged to take a categorial decision in these cases. I would like to argue that we can gain something if we put even more emphasis on fuzzy prosodic cesuras.

3.2 Fuzzy prosodic cesuras

3.2.1 Cues to prosodic cesuring and candidate prosodic cesuras

Judging from the way Auer placed the prosodic cesura symbols, he assigns major importance to the presence and kind of boundary tones (see also Gilles 2005). Thus, prosodic cesura symbols in his sample transcript regularly co-occur with a marked boundary tone. No such symbol occurs when a boundary tone is missing. Auer does notate, and discuss, pausing and prominent syllables (focus accent) (2010: 13) as well as latching and level intonation (2010: 14-15). Yet, these are referred to only in terms of "unklare prosodische Zäsurierung" ('fuzzy prosodic cesura', :13) and "mangelnde prosodische Grenzmarkierung" ('insufficient prosodic boundary marking', :15) and are left without a cesura symbol. This notation convention, in effect, levels out the potential cesura points and places them in the same category as non-cesuras.

I would like to argue that we can gain something by notating our perceptual impression. For one, we could capture language reality, and/or our perception of it, more satisfactorily. If we can perceive something as between a cesura and a non-cesura (see also 'boundary blur' in Schegloff 1996a: 72, also Du Bois et al. 1992, du Bois 2008, for a similar perception in terms of TCUs see Ford/Fox/Thompson 1996), why not notate it as exactly that. Notating a *candidate prosodic cesura* does not necessarily require us to make any claims as to what kind of cesura exactly we are dealing with. But it allows us to notate our perceptual impression of a "may-be cesura".

Second, as has been pointed out in the literature, cues to cesuring are manifold. They can occur as a bundle of features, of which not all need to be present at every (candidate) cesura (see, e.g. du Bois 2008). For all we know, pausing, prominent syllables and tempo changes *are* candidate cesural features (cf. also section 2). There may be more, such as final lengthening, for instance, and there may be others that are possibly language-specific, such as glottalization, for instance. In any case, I venture to say that we do not yet know enough about the relevance of these features for the participants in talk-in-interaction to rank one feature higher than the others *a priori*. Also, it is an empirical question whether, and if so, which features are decisive ones. If we conduct our cesura notation on all features alike, we provide ourselves with the opportunity to find out how participants actually rank them.

Thirdly, notating also candidate prosodic cesuras (see candidate cases, Schegloff 1996b) gives us the opportunity to come back to such points in talk and study them further, when a more exact description of what is going on is necessary because it seems to be relevant for the participants (for potential interactional relevance of "fuzzy" prosodic cesuras see, for instance, Lerner 2003, Ford 2004). It is a methodological prerogative of CA and interactional-linguistic work to assume that talk-in-interaction, including its actual linguistic practices, are inherently orderly (Heritage 1984: 241) Therefore, the observation of candidate cesuras as such must not be put aside.

Hence, there are a number of advantages to be gained by simply notating the very fact that some cesural cues are there and others are missing, for which reason we perceive a "fuzzy" cesura. Note that notating candidate prosodic cesuras is not the same as introducing a third

type of chunk (see section 2); all we are doing up to now is notating the fuzziness of a point in talk (on the practical issues involved see section 4).[8]

   For most researchers on spoken language this may be as far as they need/want to go because their research interests are not centered around prosodic structuring.[9] There may, however, be even more to be gained if we pay more attention to what exactly happens at the ("fuzzy") cesuras.

3.2.2   Degrees of prosodic cesuring

There are a number of hints that suggest that perhaps even treating the phenomenon of prosodic cesuring in terms of 'cesura – candidate cesura – non-cesura' needs to be refined and that this can be done so profitably.

   To begin with an observation by Auer (2010), a candidate cesura based on the occurrence of a focus accent followed by a micro-pause is certainly different in *kind* from one based on the co-occurrence of a level boundary tone and latching. Perhaps for this reason, these are referred to by Auer as 'fuzzy prosodic cesura' (2010: 13) and 'insufficient prosodic boundary marking' (2010: 15) respectively.

   Going beyond this, there are even hints that the nature of the 'cesura – candidate cesura – non-cesura' distinction is *gradual*. Interestingly, almost all proponents of the unit approach have observed that some prosodic units are more closely connected to each other than others (see also Fox 1984). This cohesion has, for instance, been seen to be produced by larger pitch phenomena (declination, Cruttenden 1986/1997, e.g.), but also by the gradient variation of prosodic boundary features. Very high onset, a long pause and a very low final fall are assumed to be characteristic of a major paratone, for instance, while minor paratones have lower onsets, shorter pauses and (only) low final falls (cf. Yule 1980, Brown/Currie/Kenworthy 1980: 71, for similar ideas see, Trim 1959, Wichmann 2000, Brazil 1997: 124-125, for instance, also 'weak rhythm group' in Pike 1945, 'internal vs. terminal juncture', in Trager/Smith 1951). Cruttenden's (1986: 43) notion of 'intonational sandhi' even refers to "the merging of two basically independent intonation-groups" (ibid.), i.e. the reduction of the boundary markers to "zero".

   The notion of a degree of boundary, or separation, has been introduced by Bolinger (1964/1972: 25). Bolinger (1984, 1989) suggests that there are different degrees of intonational subordination and finiteness depending on pitch range. He claims that "the deeper the fall, the slower the rate, and the longer the pause, the higher up on the scale of separations the particular separation is" (1989: 81, also 83-84), although he relates this to separating intonational contours and syntactic constituents at the same time.

   Interesting results along the lines of gradual cesural strength have been obtained by acoustic studies of prosodic boundary features. 't Hart/Collier/Cohen (1990) observed that pitch resetting between sentences is greater than that between clauses (cf. Thorsen 1985, 1986, Swerts/Geluykens 1993, also Lehiste 1979, for instance). In the autosegmental-metrical (AM) approach such findings have mainly served to provide evidence for the prosodic hierarchy (cf. Shattuck-Hufnagel/Turk 1996, for a summary). AM studies that take an intonational (rather than a syntax-based, see Nespor/Vogel 1986) approach to phrasing have also found gradual variation of various boundary cues (cf. Redi/Shattuck-Hufnagel 2001, Keating et al. 2003, Dilley et al. 1996, Clifton et al. 2002). De Pijper/Sanderman's (1994) study of the perception of prosodic break strength suggested that naïve, i.e. phonetically untrained, listeners are able

---

[8] To the possible criticism that introducing this fuzzy cesura category may be an easy way out for 'lazy' transcribers, it can be objected that this would be an issue with all sorts of prosodic, and other, notation conventions alike.

[9] There may be interesting correlations between candidate, or weak, cesuras and the accomplishment of multi-unit turns, complex syntactic structure and action formation, among others, though (see Szczepek Reed 2010, in prep., Barth-Weingarten in prep.).

to weigh prosodic breaks in strength auditorily with stronger breaks being realized with more phonetic boundary cues (mainly pause and pitch discontinuity) (also Schafer et al. 2000). As a result, there has been a swing in AM from syntax-based to phonetically oriented research (see, e.g., the "revised ToBI" system RaP (Rhythm and Pitch) in Breen et al. 2010, also Ladd 2008: 288-299, Brugos et al. 2008). However, the major shortcoming of these results is that they have been obtained predominantly with experimental and read-aloud monologic data, 'spoken prose' (Abercrombie 1965). Their relevance for consequential talk-in-interaction still needs to be shown (for a similar argument see Howell/Kadi-Hanifi 1991, Huber 1991, Campbell 2000: 329, Kohler 2006, Schegloff 1998: 247, 254).

In discourse-functional linguistics (DFL), Chafe (1994) reports differences in the extent of unit-final intonation depending on which IU level – these are arranged in a nested manner and coupled with information-structural levels – it occurs with. Also, Barnwell (2011), although with an experimental, statistical approach, arrives at the conclusion that the IU boundaries should be understood as a continuum from clear boundary to clear non-boundary. In a DFL-like approach for German, Peters/Kohler/Wesener (2005: 145) also arrive at the idea of *Zäsuren*. Using a potentially syntax-oriented and quantitative-statistical approach on quasi-monologic data, they claim that varying extensions of the single features and feature bundles result in various degrees of separation between successive prosodic phrases, from strong separation to complete integration.

Finally, interactional-linguistic studies have pointed out the relevance of gradual parameter differences for participants: The notion of degrees of prosodic (dis)integration of intonation phrases has been employed by Schönherr (1997) and Birkner (2006, 2008) for German. They state that it is achieved by the combination of varying numbers of cues (bundles of (dis)continuity cues) and ranges from all cues present (prosodic disintegration) to none of the cues present (prosodic integration). While Schönherr focuses on boundary cues (boundary tones, pausing, lengthening, change of global pitch register, volume and/or speech rate, rhythm etc.), Birkner (2006: 227) explicitly includes the (non)occurrence of a nuclear accent in this. Varying extensions of the relevant parameters themselves are not mentioned, except for a reduction of accent strength in comparison to neighboring phrases as a marker of continuity with Schönherr. Auer (2010) points out that we need to distinguish between pitch movements that project continuation and those that contextualize turn-ending (see also Selting 1995, 2000, Gilles 2005, Kern 2007, but see Ford 2001, Szczepek Reed 2004 for English). Participants regularly treat these differently (compare the production of continuers vs. full turns), so there is some reason to suspect that we are dealing with different kinds of prosodic cesuras here. Preliminary results presented elsewhere (Barth-Weingarten 2009) on the prosodic-phonetic features co-occurring with these different kinds of speaker changes in English even suggest that the difference is one of degree: Turn-endings are marked with the most salient and extensive set of prosodic-phonetic features and this feature set is gradually reduced (both in terms of the extension of the individual prosodic-phonetic parameters and the overall number of parameters involved) the more the cesura occurs 'within' the turn (compare prosodic marking at places of turn-taking and continuers) and the TCU (compare prosodic marking at TCU endings and in complex TCUs) (see also Barth-Weingarten/Couper-Kuhlen 2011, Barth-Weingarten in prep.).

Studies such as these, of course, need more fine-grained prosodic-phonetic analyses. For space considerations these cannot be laid out here in any detail (but see Barth-Weingarten in prep.), but the York approach of Phonetics for Conversation (cf., e.g., French/Local 1983, Local/Kelly/Wells 1986, Kelly/Local 1989a, b, Local/Walker 2005) appears to be a suitable path to take. It advocates, among other things, transcribing what we hear consistently and impressionistically, with the decisions on the relevance of phonetic events ultimately being judged against the participants' relevancies. This is based on the conviction 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.
> [...] 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."
> (Local 2007: 3, cf. also Couper-Kuhlen/Selting 1996: 45-47)

This approach would support the assumption that variation in cesural marking is of potential relevance.

In addition, this approach advocates the understanding of phonetic events as parameters, i.e. as "a variable, an ingredient which is continually present but changing in value" (Abercrombie 1965: 16) parallel to the time-axis (see also Crystal's 1969: 195-196 on a parametric approach to intonation, also the concept of 'flux' in du Bois in prep.). In terms of the prosodic-phonetic cesuring of talk, this translates into notating pitch, loudness and tempo changes, for instance, vs. assigning intonation unit(boundarie)s. What results from this is "a multi-layered approach" that "treats and records variables separately" (Kelly/Local 1989b: 202). Parametric listening deconstructs the auditory impression into its component parts and therefore promotes an impressionistic description of the data (Walker 2004: 32). It helps analysts "to free themselves from prejudices [...] which they might have by virtue of knowing and speaking their own language [...] and which avoid insofar as it is possible any reliance, conscious or not, on preordained phonetic categories" (Kelly/Local 1989a: 30-31). This also seems the road to take if we intend to study what happens at prosodic cesuras in talk in more detail: closely noting the prosodic-phonetic facts (see also du Bois et al. 1992: 104, 110, Schuetze-Coburn 1994, also section 4).

Regardless of whether one takes the layman's phonetically naïve stance or that of a trained phonetician, notation of the phenomena observed is a very real practical issue. Section 4 suggests a notation system for (candidate) cesuras.

4.   Implementing the cesura approach in transcription

The level of granularity (cf. Schegloff 2000, Bittner/Smith 2001, for instance) with which one intends to approach the issue of prosodic cesuring is directly reflected in the technicalities of notation. Most basically, there are two major ways to capture prosodic cesuring. These are, in fact, already inherent in Auer's (2010) way of notation: On the one hand, Auer uses a certain symbol for indicating prosodic cesuras, on the other hand, he notates the actual prosodic facts. Depending on whether one is mainly interested in issues beyond prosodic cesuring, or does not feel competent enough to tackle the individual prosodic parameters, or in what exactly is happening at prosodic cesuras in detail, one could choose the one or the other.

However, in the following I will suggest a system that combines both ways of notation, although to different degrees, and can thus be extended as necessary. Roughly there are three levels of granularity: minimal, basic and fine transcription. These terms are chosen deliberately because the suggestions laid out here aim at compatibility with the granularity levels employed by the transcription system most widely used in Germany, GAT (cf. Selting et al. 1998, 2009, for an adaptation to English see Selting et al. 2011).

In its current use, GAT assigns a very prominent position to IUs (compare *Intonationsphrasen* 'intonation phrases' (IPs) in GAT 2 (Selting et al. 2009), see also the more holistic concept of *Phrasierungseinheiten* ('phrasing units'), in GAT 1 (Selting et al. 1998: 101)). Already in the minimal transcript – the first of altogether three levels of transcription granularity – transcribers are required to insert line breaks/hard returns after each IP boundary. Thus, while phonetic cues to IP boundaries are described (mainly following the British school), this approach requires the polar, categorical decision-taking that has been problematized in this

paper (see section 2). In addition, while pitch movement and other prosodic and even phonetic parameters – depending on the level of transcription granularity, which is expandable according to researchers' needs ("onion-skin principle") – can be notated in relative detail, the minimal notation of one IP per line is carried through to the finest levels of detail. GAT does take note of the fact that turn-prefaces and tags can be separated prosodically from the preceding or following IP to various degrees, though. This is indicated by specific combinations of (lack of) line break and boundary tone marking (cf. Selting et al. 2011: 21-22). The transcription suggestions laid out below are intended to amend GAT with respect to the notation of candidate cesuras in general along these lines without altering the original GAT system more than necessary. In addition, section 4.3 suggests an alternative, parametric notation system which may be useful when the actual location of the cesura is not clear.

It should be noted that the transcription suggestions are indeed just that: suggestions, to be adapted to practical use for the respective research purposes at hand. More intermediary ways may turn out to be better suited to specific objects of study.

4.1 Minimal transcription of weak prosodic cesuras

In GAT, the minimal transcript is the simplest form of discourse representation (see also Selting et al. 2011). It is said to be a working tool, which is usually not sufficiently detailed for CA and IL research purposes. It captures the wording of the speaker contributions and their sequential structure in terms of turns and continuers. Yet, at the same time, it already requires the transcriber to segment the contributions into IPs. Therefore, it is at this transcription level already that the need for notating (candidate) prosodic cesuras may arise.

Against the background of the need to distinguish between at least three kinds of prosodic cesuras (see section 3.2), I suggest the following minimal notation conventions:

notation convention/symbol	technical realization	meaning
no symbol in running text	-	clear non-cesura
… | …	a single vertical bar (U+007C)[10] without line break[11]	candidate prosodic cesura
line break	hard return	clear prosodic cesura

Tab. 1: Minimal transcription conventions for prosodic cesuring

Note that this level very much restricts the notation of degrees of prosodic cesuring (see sections 3.2 and 4.2). It only notates the perception of the relevant cesura as 'clear cesura', 'clear non-cesura', and 'weaker than a clear cesura but stronger than a non-cesura'. At the same time, these conventions avoid presenting a multi-unit turn as one long block of running text (see Auer 2010, also the York approach) because clear cesuras are still marked by clear line breaks. This pays tribute to the fact that, in talk, there are clear (non-)cesuras as well as "fuzzy" ones, which can be made visible by maintaining the system of line breaks for clear cesuras and which may have informational value (see section 3.2). This is also reflected in the line/segment numbering: If the larger segment containing the candidate cesura(s) needs more space than is provided by a single line, the segment is continued on the next line and the latter is indented and not assigned a new segment number.

In addition to the three cesura categories, the GAT minimal transcript also notates some of the prosodic cesural features: in- and outbreaths as well as pausing (see Selting et al. 2011).

---

[10] This provides the symbol's character encoding in Unicode.

[11] The choice of this symbol takes advantage of the fact that it is easily available on the keyboard and relatively stable across fonts. Moreover, the vertical bar has already been associated with segment boundaries, namely as a marker for these in running text transcripts for non-linguistic research areas. The combination of line breaks and the vertical bar in one transcript will, however, disambiguate the symbol's use within the transcript. Moreover, other context features of the transcript itself will also make obvious in which research context the transcript occurs.

Ex. (3') illustrates these with the candidate cesura case from section 1:

(3')

```
897        some kind of a deal | whereby he comes | an'
           says | (0.29) a sunday_((creak))
901        offers a sunday evening liturgy
902        °hh
```

4.2 Basic transcription of weak prosodic cesuras

For the GAT basic transcription, the minimal GAT transcript "is expanded to include prosodic information which is necessary to avoid misinterpretations of the segments in their interactional context" (Selting et al. 2011: 7). This includes the notation of prosodic cesuring features such as (focus) accenting, (the kinds of cesural) pitch movements, latching, lengthening and glottal closure. Moreover, it specifically introduces the notion of IPs.

In terms of prosodic cesuring, we can adapt the minimal conventions to this greater level of granularity by extending them in the following way:

notation convention/symbol	technical realization	meaning
no symbol in running text	-	clear non-cesura
… | …	a single vertical bar (U+007C) without line break	weak(er) candidate prosodic cesura
… || …	a double vertical bar (U+007C) without line break	strong(er) candidate prosodic cesura
… |$_n$ …	n vertical bars (U+007C) without line break	candidate prosodic cesura of nth strength
line break	hard return	clear prosodic cesura

Tab. 2: Basic transcription conventions for prosodic cesuring (with an indication of notation feasibility by means of font colour)

Note that in accordance with the GAT principle of expandability, the symbol for candidate prosodic cesuras does not change its general meaning. Rather, it is adapted to the analytic depth achieved/desirable for a certain research purpose: The number of vertical bars is used to iconically notate the cesura strength observed; the more bars, the more salient/stronger the cesura. As |$_n$ indicates, there are, in principle, no limits to this. Yet, the finer the distinctions, i.e. the more bars there are employed, the harder to notate the perceived differences consistently. Therefore, it seems most recommendable to restrict oneself to just four cesura types with the notation values indicated in Table 2 in black font. As an illustration see the following ex. (3").

(3")

```
897        some kind of a deal | whereby he com:es || an'
           say:s || <<?>(0.29)>_a sunday_((creak))\
901        ?offers a sunday evening LIturgy:;=
902        =°hh?
```

Note that in any case with this system the chosen level of granularity should be employed consistently throughout (the relevant parts of) a transcript, as the signaling values change when vertical bars are added.

4.3 Fine transcription of weak prosodic cesuras

The GAT fine transcript provides a range of suggestions for notating additional prosodic features thought relevant for specific research purposes, including an extended notation of accents and the pitch movements connected with them as well as the notation of (the type of) pitch jumps and register, loudness and tempo changes, if necessary with an indication of the degree and temporal extension of the respective change. Moreover, changes of voice quality are indicated with their temporal extension (cf. Selting et al. 2009, 2011).

All of these features, and more, may be relevant in terms of prosodic cesuring. Hence, they are indicated with the GAT prosodic notation conventions in the excerpt. The vertical bars are maintained at this transcript level to point out correlations between perception and prosodic features. Compare ex. (3''') for illustration:

(3''')

```
897        ↑sOme kind of a ‾dEal | wherebY he ‾cOm:es || an'
           ‾sAy:s || <<ʔ>(0.29)ˀ_a ‾sUnday_((creak))\
901        ↑ˀoffers a ‾sUnd<<creak>ay> Evening ↑‾LITur↓g<<creaky,l,dim>y:>;=
902        =°hhˀ
```

Note that in this way, while notating the relevant features makes the transcript more difficult to read, at the same time, co-occurring feature notations also indicate the cesura strength of a particular point in the excerpt (see, for instance, the bundle of features at the cesura at the end of l. 901).

The attentive reader may have noticed that in fact the fine notation of prosodic cesuras is building up while working on the excerpt and increasing its notation granularity. The notation may also be extended along the GAT principles for further relevant features. Thus, slight pitch jumps, for instance, could be indicated with the employment of super-/subscript arrows ($^{\uparrow}$, $_{\downarrow}$) in contrast to the normal-script single and double arrows already suggested in GAT:

(3'''')

```
897         ↑sOme kind of a ‾dEal | wherebY he ‾cOm:es || an'
            ‾sAy:s || <<ʔ>(0.29)ˀ_a ‾sUnday_((creak))\
901         ↑ˀoffers a ‾sUnd<<creak>ay> Evening ↑‾LITur↓g<<creaky,l,dim>y:>;=
```

As long as the transcriber follows the principle of unambiguousness in adapting existing symbols or introducing new ones, the GAT notation system is, in principle, extendable without limitations. It is, thus, also able to capture very fine degrees in prosodic cesuring.

There is one problem, though, that may occur with amending the GAT transcription with vertical bars: The notation of cases in which it is not clear where exactly the prosodic cesura is located.

4.4 Notation of prosodic cesural areas

As has been stated at the very beginning of this paper, there are not only cases in which the prosodic cesura is weak(er) than in other cases, but also those in which it is not clear where exactly the cesura is located, i.e. where we are dealing with cesural *areas* rather than points.

With GAT-based transcription conventions, these could be indicated with the symbols ⌊ (U+235) for the beginning and ⌋ (U+251) for the end of the cesural area. When more than one cesural area occurs in close vicinity, the symbols can be indexed. If necessary, perceptual salience of the cesural area can be indicated in the basic transcript by multiplying the symbols. Other prosodic notation conventions are added as suggested by the GAT conventions.

In the relevant lines of ex. (2), this would look like this:

(2') minimal transcription of cesural areas

```
566-570    =an' they finally went there ⌊1 _an' ⌋1 stayed there⌊2 _an' ⌋2
           had a wonderful time=
```

(2") basic transcription of cesural areas

```
566-570    =an' thEy finally WENT there; ⌊⌊1 _<<all>an'> ⌋⌋1 STAYED there;
           ⌊⌊2_<<all>an'> ⌋⌋2 had a WONderful ti:m:e,=
```

(2"') fine transcription of cesural areas

```
566-570    =an' ↑thEy finally/ˉWENT th<<creaky>ere;> ⌊⌊1 _<<all>an'>         /
                        /⌋⌋1↑ˉSTAYED th<<creaky>ere;ˉ⌊⌊2 _<<all>an'>⌋⌋2↑had a/
                        /↑WONder<<creaky>ful>                       /
                        /´tI:m:e,=                                  /
```

Note that the rhythmic notation, which is also part of GAT fine transcription (see Selting et al. 2011), points out why the position of the cesuras may be perceived as uncertain: There seems to be some kind of overlay, or "clash" of prosodic cesuras – indicated by final lengthening, non-modal voice quality and accent distribution, among other parameters – and rhythmic feet boundaries (cf. Couper-Kuhlen 1993).

In cases like this one, cesura and feature notation may become rather intricate. Moreover, various parameters on different levels may partially overlap. A possible solution for this may be a parametric notation.

4.5 Parametric notation of prosodic cesuras
The York Phonology for Conversation approach applies such a notation (see Local/Kelly/Wells 1986, Kelly/Local 1989a, for instance, also Selting 1995). In a musical-notation scheme manner, tiers for various prosodic(-phonetic) parameters run parallel with a(n) (modified) orthographic notation (cf. Selting et al. 2011) of the utterance wording.

Schuetze-Coburn (1994: 153) had, in fact, already developed a rather intricate parametric notation scheme for IU boundaries. He discarded it, however, for reasons of feasibility in larger corpora. And indeed a print-out on paper for presentational purposes with restrictions by the paper margins may look quite complicated.

However, it seems to me that PRAAT (Boersma/Weeninck 1992-2010, http://www.praat.org) offers a technical solution with which at least the actual analysis, if not the one used for presentational purposes, can be handled. Apart from a time line as well as pitch and intensity curves, among other features, it offers the spectrogram to also keep track of phonetic and articulatory details in a manner that parametricizes also the notation of phonetic "segments". Fig. 1 presents a screen shot of what this could look like with the relevant part of ex. (2).
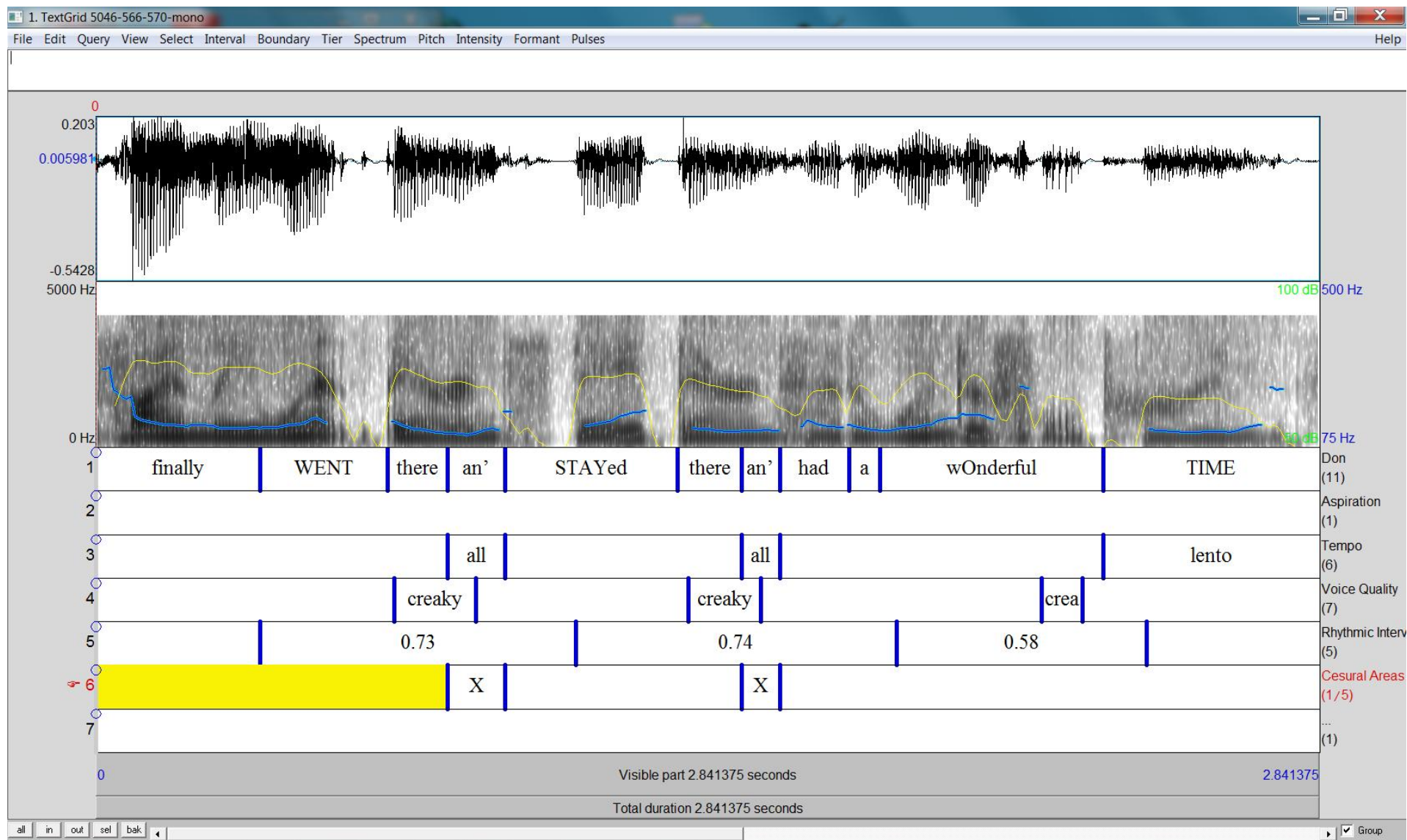
Fig. 1: Parametricized notation of the relevant lines of ex. (2) in PRAAT (screenshot from version 5.1.34)

Note that further tiers for additional parameters, and co-participant speakers, can be added as needed. Video-recordings can be annotated accordingly in ELAN (http://www.lat-mpi.eu/tools/elan/, cf. Wittenburg et al. 2006, for instance).

5 Summary

This paper has dealt with the phenomenon of perceiving "fuzzy" prosodic unit(boundarie)s in analyzing and transcribing talk-in-interaction. In doing so, it first explicitly acknowledged these as a phenomenon transcribers and analysts encounter in the form of weak prosodic boundaries and boundaries whose location is not clear. It then pointed out that most current approaches to prosodic structuring provide little means to handle them in a satisfactory way because they start off from the unit model. The latter has been argued to introduce a number of theoretical assumptions and associations that make it difficult to deal with "fuzzy boundaries" and study the prosodic(-phonetic) phenomena leading to this perception in their own right.

Instead, the paper follows Auer (2010) in advocating an approach that focuses on the unit boundaries, rather than the units themselves. To explicitly dissociate these from the units concept, they are referred to as 'cesuras'. In addition, this paper goes beyond earlier ideas in that it suggests to also notate weak, candidate cesuras and cesural areas, regardless of which prosodic(-phonetic) features lead to this kind of perception. This provides us with a tool to adapt the transcription of talk to language reality, or our perception of it. At the same time, it allows us to pay equal attention to all potentially relevant prosodic(-phonetic) parameters and to study the phenomenon of weak cesuras as well as cesural areas and their potential interactional functions in more detail. As the degree of prosodic cesuring may indeed be relevant in this regard, this paper, in a second part, suggested notation conventions that allow the transcriber and the analyst to notate candidate prosodic cesuras and cesural areas on the level of granularity needed.

It is hoped that the suggestions made resolve some of the problems transcribers face when notating the prosodic structuring of talk and that a (more) careful notation of weak cesuras and cesural areas in talk will provide the basis for new insights into the interplay of prosody and other dimensions of the organization of language and action.

References

Abercrombie, David. 1965. *Studies in phonetics and linguistics*. London: Oxford University Press.

Auer, Peter. 2000. On line-Syntax – Oder: was es bedeuten könnte, die Zeitlichkeit der mündlichen Sprache ernst zu nehmen. In: *Sprache und Literatur* 85, 43-56.

Auer, Peter. 2005. Syntax als Prozess. In: *InLiSt (Interaction and Linguistic Structures) 41*. http://www.uni-potsdam.de/u/inlist/issues/41/index.htm.

Auer, Peter. 2010. Zum Segmentierungsproblem in der gesprochenen Sprache. In: *InLiSt (Interaction and Linguistic Structures) 49*. http://www.inlist.uni-bayreuth.de/issues/49/InLiSt49.pdf (last access: 06.01.2011)

Barnwell, Brendan. 2011. The perceptual basis of intonation units. (manuscript, Febr 14, 2011)

Barth-Weingarten, Dagmar. 2009. "When to say something – some observations on prosodic-phonetic cues to the placement and types of responses in multi-unit turns." In: Barth-Weingarten, Dagmar/ Dehé, Nicole/ Wichmann, Anne (eds.). *Where prosody meets pragmatics*. Bingley: Emerald: 143-182.

Barth-Weingarten, Dagmar (in prep.). "Intonation units", turn-taking and language change – a parametricized approach to cesuring in talk-in-interaction. Postdoc thesis.

Barth-Weingarten, Dagmar/ Couper-Kuhlen, Elizabeth. 2011. Action, prosody and emergent constructions: The case of *and*. In: Auer, Peter/ Pfänder, Stefan (eds.). *Emergent constructions*. Berlin, New York: de Gruyter, 264-293.

Beckman, Mary E./ Ayers Elam, Gayle. 1997. Guidelines for ToBI Labelling (version 3.0). The Ohio State University Research Foundation (http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf, last access: 8 June 2010).

Beckman, Mary E./ Pierrehumbert, Janet B. 1986. Intonational structure in Japanese and English. In: *Phonology Yearbook* 3, 255-310.

Birkner, Karin. 2006. (Relativ-)Konstruktionen zur Personenattribuierung: "ich bin n=mensch der...". In: Günthner, Susanne/ Imo, Wolfgang (eds.). *Konstruktionen in der Interaktion.* Berlin, New York: De Gruyter (manuscript), 205-238.

Birkner, Karin. 2008. *Relativ(satz)konstruktionen im gesprochenen Deutsch. Syntaktische, prosodische, semantische und pragmatische Aspekte.* Berlin, New York: de Gruyter.

Bittner, Thomas/ Smith, Barry. 2001. A Unified Theory of Granularity, Vagueness, and approximation. In: Proceedings of the 1st Workshop on Spatial Vagueness, Uncertainty, and Granularity (SVUG01), 1-39. URL:
http://www.acsu.buffalo.edu/~bittner3/BittnerSmithSVUG01.pdf (last access: March 28, 2011)

Du Bois, John W. 2008. Rhythm and tunes: The intonation unit in the structure of dialogic engagement. Pres. held at the International Conference "Prosody and interaction", Potsdam, 15-17. September 2008.

Du Bois, John W./ Schuetze-Coburn, Stephan/ Paolino, Danae/ Cummings, Susanne. 1992. Discourse transcription, Santa Barbara Papers in Linguistics 4, Dept. of Linguistics, University of California, Santa Barbara.

Bolinger, Dwight. 1964. Around the edge of language: Intonation. In: *Harvard Educational Review* 34 (2. 282-293. Reprinted in: Bolinger, D. (ed.). 1972. *Intonation. Selected readings.* Harmondsworth: Pengiun: 19-29.

Bolinger, Dwight. 1984. Intonational Signals of Subordination. In: Berkeley Linguistics Society (ed.). *Proceedings of the tenth annual meeting of the Berkeley Linguistics Society.* Berkeley, California: University of California, 401-413.

Bolinger, Dwight. 1989. *Intonation and its uses. Melody in grammar and discourse.* Stanford: SUP.

Brazil, David. 1985/1997. *The communicative value of intonation.* Cambridge: CUP.

Brazil, David/ Coulthard, Malcolm/ Johns, Catherine. 1980. *Discourse Intonation and Language Teaching.* London: Longman.

Breen, Mara/ Dilley, Laura C./ Kraemer, John/ Gibson, Edward. (subm.). Inter-transcriber reliability for two systems of prosodic annotation: ToBI (Tones and Break Indices) and RaP (Rhythm and Pitch). http://people.umass.edu/mbreen/pubs/BreenDilleyKraemerGibson.pdf (last access: 18.06.2010) (submitted to *Corpus Linguistics or Linguistic Theory*).

Brown, Gillian/ Currie, Karen L./ Kenworthy, Joanne. 1980. *Questions of intonation.* London: Croom Helm.

Brugos, Alejna/ Veilleux, Nanette/ Breen, Mara/ Shattuck-Hufnagel, Stefanie. 2008. The alternatives (Alt) tier for ToBI: Advantages of capturing prosodic ambiguity. In: Proceedings *of Speech Prosody 2008*, Campinas, Brazil (http://www.people.umass.edu/mbreen/pubs.html, last access: 18 June 2010)

Campbell, Nick. 2000. Timing in speech: A multi-level process. In: Horne, Merle (ed.). *Prosody: Theory and experiment.* Dordrecht: Kluwer, 281-334.

Chafe, Wallace. 1988. Linking intonation units in spoken English. In: Haiman, John/ Thompson, Sandra A. (eds.). *Clause Combining in Grammar and Discourse.* Amsterdam, Philadelphia: Benjamins, 1-27.

Chafe, Wallace. 1994. *Discourse, consciousness and time. The flow and displacement of conscious experience in speaking and writing.* Chicago: Chicago University Press.

Clifton, Charles, Jr./ Carlson, Katy/ Frazier, Lyn. 2002. Informative prosodic boundaries. In: *Language and Speech* 45 (2). 87-114.

Couper-Kuhlen, Elizabeth. 1986. *An Introduction to English prosody.* Tübingen: Niemeyer.

Couper-Kuhlen, Elizabeth. 1993. *English Speech Rhythm. Form and function in everyday verbal interaction.* Amsterdam, Philadelphia: Benjamins.

Couper-Kuhlen, Elizabeth/ Selting, Margret. 1996. Towards an interactional perspective on prosody and a prosodic perspective on interaction. In: Couper-Kuhlen, Elizabeth/ Selting, Margret (eds.). *Prosody in Conversation: Interactional Studies.* Cambridge: Cambridge University Press, 11-56.

Cruttenden, Alan. 1986/1997. *Intonation.* Cambridge: CUP.

Crystal, David. 1969. *Prosodic systems and intonation in English.* Cambridge: Cambridge University Press.

Crystal, David. 1975. *The English tone of voice. Essays in intonation, prosody and paralanguage.* London: Arnold.

Dilley, Laura/ Shattuck-Hufnagel, Stefanie/ Ostendorf, Mari. 1996. Glottalization of word-initial vowels as a function of prosodic structure. In: *Journal of Phonetics* 24: 423-444.

Ford, Cecilia E. 2001. At the intersection of turn and sequence. Negation and what comes next. In: Selting, Margret/ Couper-Kuhlen, Elizabeth (eds.). *Studies in interactional linguistics.* Amsterdam, Philadelphia: Benjamins, 51-79.

Ford, Cecilia E. 2004. Contingency and units in interaction. In: *Discourse Studies* 6: 27–52.

Ford, Cecilia E./ Thompson, Sandra A. 1996. Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns. In: Ochs, Elenor/ Schegloff Emanuel A./ Thompson, Sandra A. (eds.). *Interaction and grammar.* Cambridge: Cambridge University Press, 134-184.

Ford, Cecilia E./ Fox, Barbara A./ Thompson, Sandra A. 1996. Practices in the construction of turns: The "TCU" revisited. In: *Pragmatics* 6, 427-454.

Fox, Anthony. 1984. Subordinating and co-ordinating intonation structures in the articulation of discourse. In: Gibbon, Dafydd/ Richter, Helmut (eds.). *Intonation, accent and rhythm.* Berlin, New York: de Gruyter, 120-133.

French, Peter/ Local, John. 1983. Turn-competitive incomings. In: *Journal of Pragmatics*, 7, 17-38.

Gilles, Peter. 2005. Regionale Prosodie im Deutschen: Variabilität in der Intonation von Abschluss und Weiterweisung. Berlin: de Gruyter.

Gumperz, John/ Berenz, Norine. 1993. Transcribing conversational exchanges. In: Edwards, Jane/ Lampert, Martin D. (eds.). *Talking data. Transcription and coding in discourse research.* Hillsdale, N.Y.: Erlbaum, 91-121.

Halford, Brigitte K. 1994. *The discourse function of intonation.* In: Halford, Brigitte K./ Pilch, Herbert (eds.). Intonation. Tübingen: Narr, 69-88.

Halford, Brigitte K. 1996. *Talk Units. The structure of spoken Canadian English.* Tübingen: Narr.

Halliday, Michael A.K. 1967. *Intonation and grammar in British English.* The Hague: Mouton.

Halliday, Michael A. K. 1985. *An introduction to functional grammar.* London: Arnold.

't Hart, Johan/ Collier, René/ Cohen, Antonie. 1990. *A perceptual study of intonation. An experimental-phonetic approach to speech melody.* Cambridge: CUP.

Heritage, John. 1984. *Garfinkel and ethnomethodology.* Cambridge: Polity Press.

Hermes, Dik J. 2006. Stylization of pitch contours. In: Sudhoff, Stefan/ Lenertová, Denisa/ Meyer, Roland/ Pappert, Sandra/ Augurzky, Petra/ Mleinek, Ina/ Richter, Nicole/ Schließer, Johannes (eds.). *Methods in Empirical Prosody Research.* Berlin, New York: de Gruyter, 29-61.

Howell, Peter/ Kadi-Hanifi, Karima. 1991. Comparison of prosodic properties between read and spontaneous speech material. In: *Speech Communication* 10: 163-169.

Huber, Dieter. 1991. On the discourse function of intonation. In: *Proceedings of the XIIth International Congress of Phonetic Sciences*, Vol. 5: 190-193.

Keating, Patricia/ Cho, Taehong/ Fougeron, Cécile/ Hsu, Chai-Shune. 2003. Domain-initial articulatory strengthening in four languages. In: Local, John/ Ogden, Richard/ Temple, Rosalind (eds.). *Phonetic interpretation. Papers in Laboratory Phonology VI.* Cambridge: CUP, 145-161.

Kelly, John/ Local, John. 1989a. *Doing phonology. Observing, recording, interpreting.* Manchester, New York: Manchester University Press.

Kelly, John/ Local, John. 1989b. *On the use of general phonetic techniques in handling conversational material.* In: Roger, Derek/ Bull, Peter (eds.). Conversation. An interdisciplinary perspective. Clevedon: Multilingual matters, 197-212.

Kern, Friederike. 2007. Prosody as a resource in children's game explanations: Some aspects of turn construction and recipiency. In: *Journal of Pragmatics* 39, 111-133.

Knowles, G. 1991. Prosodic labelling: The problem of tone group boundaries. In: Johansson, Stig/ Stenström, Anna-Brita (eds.). *English computer corpora: Selected papers and research guide.* Mouton de Gryuter: 149-163.

Kohler, Klaus J. 2006. Paradigms in experimental prosodic analysis: From measurement to function. In: Sudhoff, Stefan/ Lenertová, Denisa/ Meyer, Roland/ Pappert, Sandra/ Augurzky, Petra/

Mleinek, Ina/ Richter, Nicole/ Schließer, Johannes (eds.). *Methods in empirical prosody research*. Berlin: de Gruyter, 123-152.

Ladd, Robert D. 2008. *Intonational phonology*. Cambridge: CUP.

Lehiste, Ilse. 1979. Perception of sentence and paragraph boundaries. In: Lindblom, Bjorn/ Ohman, Sven (eds.). *Frontiers of speech research*. London: Academic Press, 191-201.

Lerner, Gene H. 1996. On the 'semi-permeable' character of grammatical units in conversation: conditional entry into the turn space of another speaker. In: Ochs, Elinor/ Schegloff Emanuel A./ Thompson, Sandra A. (eds.). *Interaction and grammar*. Cambridge: Cambridge University Press, 238-276.

Lerner, Gene H. 2003. Selecting next speaker: The context-sensitive operation of a context-free organization. In: *Language in Society* 32, 177-201.

Local, John. 2007. Phonetic detail and the organisation of talk-in-interaction. In: *Proceedings ICPhS 2007* (http:/www.icphs2007.de, accessed: 20-02-2008),

Local, John/ Kelly, John/ Wells, W.H.G. 1986. Towards a phonology of conversation: turn-taking in Tyneside English. In: *Journal of Linguistics* 22, 411-437.

Local, John/ Walker, Gareth. 2005. Methodological imperatives for investigating the phonetic organization and phonological structures of spontaneous speech. In: *Phonetica* 62, 120-130.

Mukherjee, Joybrato. 2001. *Form and Function of Parasyntactic Presentation Structures. A Corpus-based Study of Talk Units in Spoken English*. Amsterdam, Atlanta: Rodopi.

Nespor, Marina/ Vogel, Irene. 1986. *Prosodic phonology*. Dordrecht: Foris.

Palmer, Harold. 1922. English intonation, with systematic exercises. Cambridge: Heffer.

Peters, Benno/ Kohler, Klaus J./ Wesener, Thomas. 2005. Phonetische Merkmale prosodischer Phrasierung in deutscher Spontansprache. In: Kohler, Klaus J./ Kleber, Felicitas/ Peters, Benno (eds.). *Prosodic Structures in German Spontaneous Speech* (AIPUK 35a), Kiel: IPDS, 143-184.

Pierrehumbert, Janet B. 1980. *The phonetics and phonology of English intonation*. PhD, MIT.

de Pijper, Jan Roelof/ Sanderman, Angelien A. 1994. On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues. In: *Journal of the Acoustical Society of America* 96 (4. 2037-2047.

Pike, Kenneth L. 1945. *The intonation of American English*. Ann Arbor, Michigan: University of Michigan Publications.

Redi, Laura/ Shattuck-Hufnagel, Stefanie. 2001. Variation in the realization of glottalization in normal speakers. In: Journal of Phonetics 29, 407-429.

Sacks, Harvey/ Schegloff, Emanuel A./ Jefferson, Gail 1974. A simplest systematics for the organization of turn-taking for conversation. In: *Language* 50, 696-735.

Schafer, Amy J./ Speer, Shari R./ Warren, Paul/ White, S. David. 2000. Intonational disambiguation in sentence production and comprehension. In: *Journal of Psycholinguistic Research* 29 (2):169-182.

Schegloff, Emanuel A. 1987. Analyzing Single Episodes of Interaction: An Exercise in Conversation Analysis. In: *Social Psychological Quarterly* 50 (2): 101-114.

Schegloff, Emanuel A. 1988. Discourse as an interactional achievement II: An excercise in Conversation Analysis. In: Tannen, Deborah (ed.). *Linguistics in context: Connecting observation and understanding*. Norwood, N.J.: Ablex, 135-158.

Schegloff, Emanuel A. 1996a. Turn-organization: one intersection of grammar and interaction. In: Ochs, Elinor/ Schegloff, Emanuel A./ Thompson, Sandra A. (eds.). *Interaction and grammar*. Cambridge: Cambridge University Press, 52-133.

Schegloff, Emanuel A. 1996b. Confirming allusions: Toward an empirical account of action. In: *American Journal of Sociology* 102: 161-216.

Schegloff, Emanuel A. 1998. Reflections on studying prosody in talk-in-interaction. In: *Language and Speech* 41: 235-263.

Schegloff, Emanuel A. 2000. *On granularity*. In: Annual Review Sociol.. 26: 715-720.

Schönherr, Beatrix. 1997. *Syntax – Prosodie – nonverbale Kommunikation. Empirische Untersuchungen zur Interaktion sprachlicher und parasprachlicher Ausdrucksmittel im Gespräch*. Tübingen: Niemeyer.

Schuetze-Coburn, Stephan. 1992. Prosodic phrase as prototype. In: *Proceedings of the IRCS workshop on prosody in natural speech*. Vol. 1, Institute for Research in Cognitive Science, Univ. of Pennsylvania, Philadelphia, PA, 171-180.

Schuetze-Coburn, Stephan. 1994. *Prosody, syntax, and discourse pragmatics: Assessing information flow in German conversation*. PhD dissertation, University of California, Los Angeles.

Selting, Margret. 1993. Phonologie der Intonation. Probleme bisheriger Modelle und Konsequenzen einer neuen interpretativ-phonologischen Analyse. In: *Zeitschrift für Sprachwissenschaft* 11: 99-138.

Selting, Margret. 1995. *Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation*. Tübingen: Niemeyer.

Selting, Margret 2000. The construction of units in conversational talk. In: *Language in Society* 29: 477-517.

Selting, Margret. 2005. Syntax and prosody as methods for the construction and identification of turn-constructional units in conversation. In: Hakulinen, Auli/ Selting, Margret (eds.). *Syntax and Lexis in conversation. Studies on the use of linguistic resources in talk-in-interaction*. Amsterdam, Philadelphia: Benjamins, 17-44.

Selting, Margret/ Auer, Peter/ Barden, Birgit/ Bergmann, Jörg/ Couper-Kuhlen, Elizabeth/ Günthner, Susanne/ Meier, Christoph/ Quasthoff, Uta/ Schlobinski, Peter/ Uhmann, Susanne. 1998. Gesprächsanalytisches Transkriptionssystem (GAT). In: *Linguistische Berichte* 173, 91-122.

Selting, Margret/ Auer, Peter/ Barth-Weingarten, Dagmar/ Bergmann, Jörg/ Bergmann, Pia/ Birkner, Karin/ Couper-Kuhlen, Elizabeth/ Deppermann, Arnulf/ Gilles, Peter/ Günthner, Susanne / Hartung, Martin/ Kern, Friederike/ Mertzlufft, Christine/ Meyer, Christian/ Morek, Miriam/ Oberzaucher, Frank/ Peters, Jörg/ Quasthoff, Uta/ Schütte, Wilfried/ Stukenbrock, Anja/ Uhmann, Susanne. 2009. Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 10: 353-402 (www.gespraechsforschung-ozs.de).

Selting, Margret/ Auer, Peter/ Barth-Weingarten, Dagmar/ Bergmann, Jörg/ Bergmann, Pia/ Birkner, Karin/ Couper-Kuhlen, Elizabeth/ Deppermann, Arnulf/ Gilles, Peter/ Günthner, Susanne/ Hartung, Martin/ Kern, Friederike/ Mertzlufft, Christine/ Meyer, Christian/ Morek, Miriam/ Oberzaucher, Frank/ Peters, Jörg/ Quasthoff, Uta/ Schütte, Wilfried/ Stukenbrock, Anja/ Uhmann, Susanne (in prep., 2011). A system for transcribing talk-in-interaction: GAT 2 *(translated and adapted for English by Elizabeth Couper-Kuhlen and Dagmar Barth-Weingarten)*. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* (www.gespraechsforschung-ozs.de).

Shattuck-Hufnagel, Stefanie/ Turk, Alice E. 1996. A prosody tutorial for investigators of auditory sentence processing In: *Journal of Psycholinguistic Research* 25(2): 193-247.

Swerts, Marc/ Geluykens, Ronald. 1993. The prosody of information units in spontaneous monologue. In: *Phonetica* 50: 189-196.

Szczepek Reed, Beatrice. 2004. Turn-final intonation in English. In: Couper-Kuhlen, Elizabeth/ Ford, Cecilia E. (eds.). *Sound patterns in interaction*. Amsterdam, Philadelphia: Benjamins: 97-117.

Szczepek Reed, Beatrice. 2010. Intonation phrases in natural conversation: A participants' category? In: Barth-Weingarten, Dagmar/ Reber, Elisabeth/ Selting, Margret (eds.). *Prosody in interaction*. Amsterdam: Benjamins, 191-212.

Szczepek Reed, Beatrice. (in prep.). Prosody, syntax and action formation: Intonation phrases and 'action components'. In: Bergmann, Pia/ Brenning, Jana/ Pfeiffer, Martin/ Reber, Elisabeth (eds.). Prosody and visual signals. Approaches towards usage-based grammar theories. (Series "Linguae et litterae"). Berlin, New York: de Gruyter.

Tench, Paul. 1990. *The roles of intonation in English discourse*. Frankfurt, M.: Lang.

Thorsen, Nina G. 1985. Intonation and text in standard Danish. In: *Journal of the Acoustical Society of America* 77, 1205-1216.

Thorsen, Nina G. 1986. Sentence intonation in textual context. In: *Journal of the Acoustical Society of America* 80, 1041-1047.

Thompson, Sandra A. 2011. Grammar and social interaction: the question of 'units'. Presentation at the workshop "Grammar and interaction revisited", FiDiPro project "Grammar and interaction: The linking of actions in speech and writing", Helsinki, 10-12 March 2011.

Trager, George L./ Smith, Henry Lee Jr. 1951/1975. *An outline of English structure. Studies in Linguistics: Occasional Paper* 3, American Council of Learned Societies, Washington D.C..

Trim, J.L.M. 1959. Major and minor tone-groups in English. In: Le Maître Phonétique, 3rd series 112: 26-29. Reprinted in: Jones, W.E./ Laver, J. (eds.) 1973. *Phonetics in linguistics. A book of readings*. London: Longman, 320-323.

Walker, Gareth. 2004. The phonetic design of turn endings, beginnings, and continuations in conversation. PhD thesis, University of York, Department of language and linguistic sciences. (manuscript).

Wennerstrom, Ann. 2001. *The music of everyday speech. Prosody and discourse analysis*. Oxford: OUP.

Wichmann, Anne. 2000. *Intonation in text and discourse. Beginnings, middles and ends*. London: Longman.

Wittenburg, Peter/ Brugman, Hennie/ Russel, Albert/ Klassmann, Alex/ Sloetjes, Han. 2006. ELAN: a professional framework for multimodality research. In: *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*, 1556-1559 (available via http://www.mpi.nl/publications/escidoc-60436/@@popup, last access: 23 May 2011).

Yule, George. 1980. Speakers' topics and major paratones. In: *Lingua* 52: 33-47.

Appendix: Transcription conventions

In this paper, the GAT2 transcription conventions (Selting et al., 2009, in press) are employed. The most important symbols used here are listed below. For reasons of readability, mainly standard orthography is chosen. Line/segment numbers refer to time stamps of the original LDC recordings in a simplified manner.

Minimal transcript

`[  ]` `[  ]`	simultaneous talk
`°h,°hh/h°,°hh`	in-/outbreaths according to duration
`(.),(0.5)`	micro pause; measured pause of the duration stated
`an'`	modified orthography (*and*)
`and_uh`	cliticizations within units
`<<creaky>    >`	change in voice quality as stated
$\rightarrow$	refers to a line of transcript relevant in the argument
one line	represents roughly one segment/intonation phrase

Basic transcript

`=`	latching
`:, ::, ::`	lengthening according to duration
`ʔ`	glottal closure
`\`	cut-off without glottal closure
`SYLlable`	primary accent
`?`	phrase-final pitch rising to high
`,`	phrase-final pitch rising to mid
`–`	level phrase-final pitch
`;`	phrase-final pitch falling to mid
`.`	phrase-final pitch falling to low

Fine transcript

`sYllable`	secondary accent
↑, ↑↑, ꜛ	smaller, larger, or very small pitch upstep
↓, ↓↓, ꜜ	smaller, larger, or very small pitch downstep
`<<l>  >`	low pitch register, with scope
`¯SO`	level accent pitch movement
`<<all>   >`	allegro, fast, with scope
`<<dim>   >`	diminuendo, increasingly softer, with scope